The Hubris and Humility Effect and the Domain-Masculine Intelligence Type: Exploration of Determinants of Gender Differences in Self-Estimation of Ability.

Josephine Storek

University College London

A thesis submitted to University College London

for the degree of Doctor of Philosophy in Psychology

March, 2011

Thesis Declaration

I, Josephine Storek confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: ____

Date: ____

To my children

May you unleash your talents with the right amount of hubris and humility, but most of all…Carpe Diem!

Abstract

This thesis explores the potential determinants of gender differences in self-estimated intelligence. In particular, it addresses the determinants of gender differences in the *'domain-masculine intelligence type'* that is expected to yield the most significant gender differences in the self-estimated intelligence model (SEI). Equally, it sets to confirm the occurrence of the *'hubris-humility effect'* (HHE), i.e. male overestimation and female underestimation of cognitive abilities, specifically in the above intelligence type.

The thesis contains eight chapters, ten correlational studies and five experimental studies. The thesis is divided in two sections. Section one contains the ten correlational studies and section two the five experimental studies. All studies are independent but related.

Chapter one contains a review of the relevant literature. It is divided into three sub-sections: overview, intelligence and hubris-humility effect (HHE) and domain-masculine intelligence type (DMIQ): gender differences in self-estimated intelligence.

Chapter two (Studies 1 and 2) introduces the domain-masculine intelligence type and demonstrates it is the most sensitive indicator of gender differences in the SEI model. HHE is shown to be the most pronounced and confined to occurring on DMIQ. Equally, gender is shown as the best predictor of DMIQ, over and above a number of other demographic variables.

Chapter three (Studies 3 to 5) sets to validate the occurrence of HHE on DMIQ, while it introduces psychometric intelligence ('*g*') and implicit beliefs about intelligence as possible determinants of DMIQ. Studies 3 and 4 examine the role '*g*', as measured by fluid (Gf) and crystallised (Gc) intelligence tests, play in DMIQ. Results confirm the occurrence of HHE on DMIQ and reveal significant gender

differences in Gf and Gc, with medium and large effect sizes. Gender is shown to influence the relationship between '*g*' and DMIQ. Contrary to prediction, a psychometric intelligence measure (Gf), and not gender, is the best predictor of DMIQ. Implicit beliefs about intelligence play no role in the prediction of DMIQ. Study 5 adds gender identity variables, i.e. masculinity and femininity, and self-construct measures, i.e. self-esteem and self-control, to Gf and Gc, as possible predictors of DMIQ. Results validate the existence of HHE on DMIQ and confirm gender as the best predictor DMIQ, over and above '*g*', gender identity variables and self-construct measures.

Chapter four (Studies 6 and 7) examines the role gender identity, i.e. masculinity and femininity, affect measures, i.e. positive and negative affect, and self-constructs, i.e. self-esteem and self-control, play as potential determinants of DMIQ. Both studies confirm the existence of HHE on DMIQ. Study 6 confirms gender as the best and only predictor of DMIQ. Study 7 affirms masculinity as the best predictor of the intelligence type, followed by gender.

Chapter five (Studies 8 and 9) examines the role of culture in DMIQ and its impact on the existence of HHE on DMIQ. Gender identity variables are also included to validate the earlier findings and to explore the role masculinity plays as a predictor of DMIQ, in three distinct cultures. Study 8 was conducted in Czech Republic and Study 9 in Colombia and United Kingdom. Results confirm the occurrence of HHE on DMIQ in all three cultures, with medium effect size for the Czech sample and large effect sizes for the Colombian and British samples. Gender is shown to influence the relationship between gender identity variables and DMIQ. Contrary to prediction, masculinity and not gender, is the best predictor of DMIQ in the Czech Republic sample. In the Colombian sample, none of the entered variables significantly contributes to the prediction of DMIQ. In the British sample, gender is affirmed as the best predictor of DMIQ, followed by masculinity. The results suggest that culture influences the composition of DMIQ determinant(s).

Chapter six (Study 10) explores the role of DMIQ in a precocious sample, i.e. members of Mensa UK. It also sets to validate the occurrence of HHE prevails on DMIQ in a population that is knowledgeable about intelligence as well as aware of its own intellectual superiority. Beliefs about intelligence and gender identity variables are also included to explore whether they will play a role in the prediction of the intelligence type. The results confirm the existence of HHE on DMIQ in this precocious population, providing additional evidence for the degree of embeddedness and impact of HHE on highly gifted individuals. Gender is confirmed as the only and best predictor of DMIQ.

Chapter seven (Studies 11 to 15) contains five independent experimental studies. Study 14 was conducted with three independent samples to test three varying task-confidence conditions. The results of the three individual conditions are reported in the Appendix, while the combined total results are reported in Study 14. The five experiments consist of repeated measurement of DMIQ and a psychometric task (TCAP) that also includes task-success probability probes (TSP). Participants are asked to estimate DMIQ before and after the task. The task contains numerical, reasoning, and crystallised intelligence items as well as task-success or task-confidence probes. The number of the psychometric items and probes are manipulated per experiment to assess their impact on the results. As such, the task is expected to be gender-stereotype inducing.

As in the correlational studies, HHE is predicted to occur in the pre- and post-task DMIQ conditions. Results of all five studies validate the existence of HHE on

DMIQ1 and DMIQ2, with medium to very large effect sizes. Likewise, a significant decrease in the DMIQ estimates is observed in all five studies, with small to medium effect sizes.

In addition, male advantage is confirmed on the psychometric task and the task-success probes. Gender differences in TCAP are observed in Studies 11, 12 and 15, with males correctly solving significantly more psychometric problems than females. Equally, gender differences in TSP occur in Studies 11, 12 and 13, with males providing significantly higher task-confidence answers than females.

To validate the earlier results, gender is expected as the best predictor of DMIQ1 and DMIQ2. Results reveal that gender is the best predictor of DMIQ1 in three out of five studies and in two out of five studies in DMIQ2. Unexpectedly, task-success probes are twice the best predictor of DMIQ1 and three times the best predictor of DMIQ2.

Moreover, gender influences the relationship between TPS and DMIQ1 and DMIQ2 in all five studies. Equally, gender influences the relationship between TCAP and DMIQ1 and DMIQ2, in all but one analysis. Surprisingly, the DMIQ1 and DMIQ2 estimates that are provided by participants in the three task-success probability groups, i.e. low, average and high, are startlingly accurate, with the exception of Study 14. That is, low DMIQ estimates are provided by participants with low task-success confidence, average estimates are provided by participants with average task-success confidence and the highest DMIQ estimates by individuals with highest task-success confidence. Results for TCAP are complex and less accurate. Yet, for both TSP and TCAP, males provide significantly higher DMIQ1 and DMIQ2 estimates than females, providing further evidence for the occurrence of male hubris in the self-estimation of ability process.

Chapter eight presents a brief summary of results and conclusions of this research. Equally, limitations of this research are discussed and a number of future research recommendations provided.

The appendix includes the three individual condition studies of Study 14; that is Studies 14A, 14B and 14C. The TCAP and TSP overviews for Studies 11 to 15 are also included. Finally, Study 16 that uses the combined sample made of the fifteen individual study samples (N = 2292) is integrated. Study 16 tests the main objectives of this thesis through previously used hypotheses and as such provides a summary overview of the results. All main objectives of this thesis are corroborated.

Acknowledgements

I would like to express my gratitude to Professor Dr. Adrian Furnham for agreeing to supervise this dissertation. I could not have wished for a better mentor. I was given freedom, encouragement and sufficient guidance. Rabbi Abraham Joshua Heschel said: *"When I was young I admired clever people; now that I am older I admire kind people"*. In Adrian, I was fortunate enough to meet a '*clever and kind*' person.

I would also like to thank Dr. Tomas Chamorro-Premuzic for agreeing to be my secondary supervisor.

I am eternally grateful to my two very special Ds for unreservedly believing and supporting me on this journey. Without your support, this dream wouldn't be a reality. I am indebted to you both forever.

Contents

List of Figures

List of Tables

Chapter 1: Literature Review

1.1. Overview

Intelligence research has been a key interest area of research for psychologists for over a century. Despite a lot of progress in the field of intelligence research, several areas still lack clarity and necessary conceptualisation (cf. Neisser et al., 1996).

Since intelligence research focuses on individual and group differences in mental abilities and competencies, it has, unsurprisingly, generated more controversy than any other psychological concept. Disagreements concern the definition, measurement and distribution of intelligence.

An area within the intelligence research arena that has been prominent in the last three decades concerns multiplicity theories of intelligences, including the self-assessed (SAI) or self-estimated intelligence(s)[1] (cf. Furnham, 2001; Gardner, 1983, 1993, 1999). This thesis concerns itself with the 'subjective' or self-assessed multiple intelligence(s) (e.g. Furnham, 2001; Gardner, 1983, 1993, 1999) and not with a single general ability '*g*' (cf. Cattell, 1943; Gottfredson, 1997a,b, 2000a,b; Jensen, 1998; Spearman, 1904).

Considerable evidence from the self-assessed intelligence research programme shows that universal gender differences exist in general population (cf. Furnham, 1999, 2001; Furnham & Budhani, 2002; Furnham, Clark, & Bailey, 1999; Furnham,

[1] The terms self-assessed and self-estimated intelligence(s) (SAI/SEI) have been used inconsistently and interchangeably within the self-assessed ability research programme. The two terms/labels are referring to the same construct and have same meaning. Both terms are used throughout this thesis. It must be noted that the SAI/SEI construct has been shown to be an independent construct from self-concept, self-efficacy, academic self-beliefs and personality trait Intellect (Peterson & Whiteman, 2007).

Crawshaw, & Rawles, 2006). This is particularly true for mathematical/logical and spatial intelligences, with males overestimating and females underestimating their ability (e.g. Furnham & Budhani 2002; Furnham & Rawles, 1995; Marsh & Yeung, 1998). This so-called *'hubris-humility effect'* (HHE) (e.g. Beloff, 1992; Furnham, 2001, Furnham, Hosoe, & Tang, 2003) is at the centre of this thesis. Specifically, this thesis addresses the determinants of gender differences that occur on the *'domain-masculine'* intelligence type (DMIQ) by introducing gender role identity, self-concept variables and affect measures in ten correlational studies. The second part contains five experimental studies that were designed to assess the impact of gender on the self-estimation condition(s) by introducing varying psychometric tasks and confidence assessments as well some variables from section one.

This review will start with a brief overview of intelligence, followed by a description of the most pertinent findings about sex differences in test and self-assessed intelligence. The multiple intelligence theories and the self-estimated ability research will be addressed next. In particular, the review will summarise the most relevant findings from the

self-, other-, parental and cultural self-estimates of ability research programme.

The *'hubris-humility effect'* and the *'domain-masculine'* intelligence type will be discussed next. Subsequently, the possible causes of gender differences in mathematics achievement, attitudes and affect, including the gender gap in education and role of paternal attitudes will be addressed. An overview of the most relevant findings about gender differences in cognitive biases, self-perceptions and self-concept as well as the accuracy of self-estimates of performance will follow. The review will then address the most pertinent evidence about self-confidence and stereotypical beliefs and the roles they potentially play in the hubris-humility effect. Sections on gender identity, self-concept, and affect will conclude the overview.

The aim of this review is to summarise the most influential findings in each area, rather than provide an exhaustive overview.

1.2. Intelligence

Academics have struggled to agree on a universal definition and standardized form of intelligence measurement since before the 20th century (cf. Eysenck, 1998; Gottfredson, 1997a, 2000; Neisser et al., 1996; Weinberg, 1989). Though academics are yet to agree a common definition of intelligence, the term itself was already used in the Old Testament and by early Greeks and Romans.

1.2.1. Meaning of Intelligence

The two currently most accepted definitions of intelligence are results of an organised scientific reaction to the publication of *The Bell Curve* (Herrnstein & Murray, 1994). The first one was first published in the Wall Street Journal on December 13, 1994, as a statement of reactions by fifty-two leading figures in intelligence research (Gottfredson, 1994, 1997a): *"Intelligence is a very general mental capability that, among others things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-making smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings – 'catching on', 'making sense' of things, or 'figuring out' what to do."*

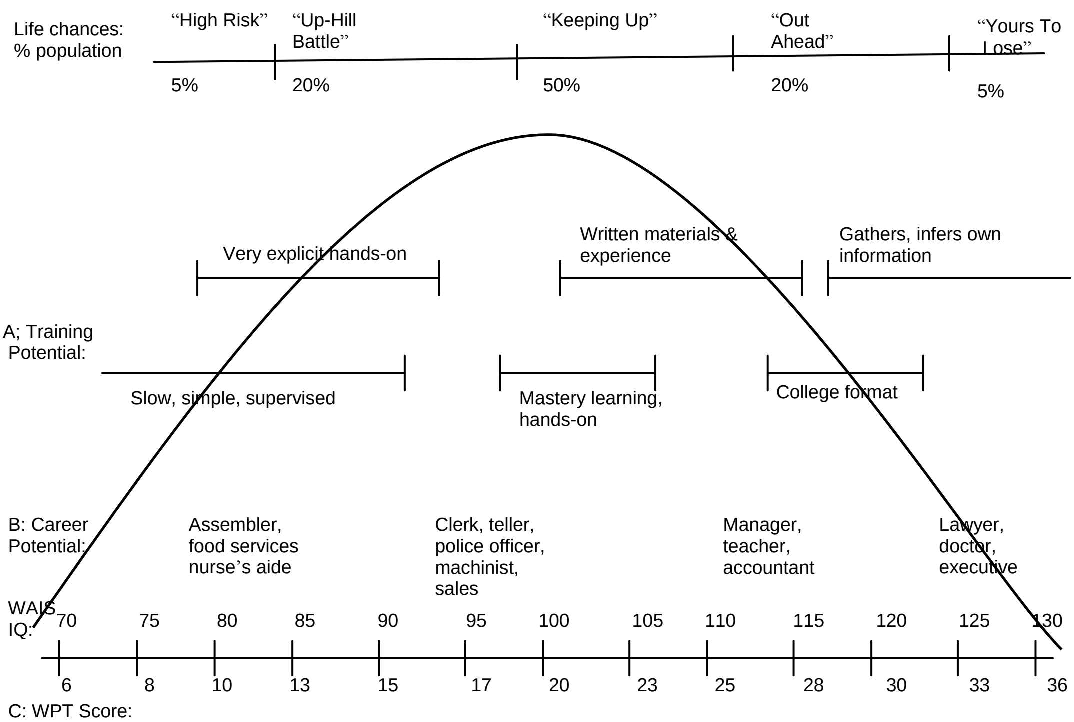
The second definition was published by the task force of the American Psychological Association (APA) (Neisser et al., 1996), also in reaction to the emotional debate that inflamed after The Bell Curve's publication: *"Individuals differ from one another in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought. Although these individual differences can be substantial, they are never entirely consistent: a given person's intellectual performance will vary on different occasions, in different domains, as judged by different criteria."*

Interestingly, the implicit theories or the ideas people hold about what constitutes intelligence, that is, logical reasoning, verbal ability and social intelligence do not significantly differ from the views of the experts (cf. Furnham, 2001; Sternberg et al., 1981; Weinberg, 1989). The general public rightfully associates intelligence with the use and outcomes of widely available psychometric or '*intelligence quotient'* (IQ) tests that are known to predict many social outcomes (Eysenck, 1990, 1994). Public beliefs and implicit intelligence theories are very powerful as they can bring about social and educational changes, such as banning intelligence testing from education, training and recruitment processes (Furnham, 2001) and influence individual performance expectations and appraisals (Pomerantz & Ruble, 1997).

Expert or explicit theories of intelligence generally stem from the psychometric measurement tradition, defining intelligence as the ability to deal with complexity, learn, reason, remember, solve problems, and plan (cf. Deary, 2001; Gottfredson, 1997a, 2000; Jensen, 1998; Wikipedia, Retrieved 19 November 2010). The experts can be further divided into *'lumpers'* or '*splitters'* (cf. Furnham, 2004; Mayr, 1982), with '*lumpers*' defining intelligence as a single or general mental capacity (cf. Binet & Simon, 1905; Gottfredson, 1997a, 2000; Spearman, 1904; Stankov, 2000) and '*splitters*' seeing intelligence as many separate mental abilities (e.g. Cattell, 1943; Carroll, 1993; Gardner, 1983; Sternberg, 1985; Thurstone, 1919).

About 95% of the population has IQs between 70 and 130, i.e. 4 standard deviations from the mean (100). Whilst IQ does not determine one's fate, many studies have shown that IQ is highly and positively correlated with a number of desirable life outcomes, such as high-level education, high status jobs, personal income, health, life success, job performance and job choice and negatively with school drop-out rates, crime rates, incarceration, teen pregnancy rates and crime (see Figure 1; Gottfredson, 2000, 2005; Kuncel et al., 2004).

Figure 1.2.1: Overall life chances at different ranges of the IQ bell curve. (Gottfredson, 1997b, p.117; 2000, p. 1364).



Legend: A: Wonderlic (1992, p. 26); B: Wonderlic (1992, pgs. 20, 26, 27); C: Wonderlic (1992, p. 20).

1.2.2. Intelligence Models

The first scientist to propose a theory of general intelligence was Darwin's cousin, Sir Francis Galton. Galton asserted that intelligence is a biological mental ability, measurable through individual's reaction time to mental tasks. Charles Spearman, Galton's student, demonstrated through the statistical technique of factor analysis (1904) that a general intelligence factor or '*g*' can be calculated from any set of cognitive tests. His Two-Factor Theory of Intelligence states that every mental test can be divided into a general or.'*g*' factor that measures the common function across all ability tests and a specific or '*s*' factor that is unique to each ability test. In fact, psychometric '*g*' and not IQ is used as the research definition of intelligence (Gottfredson, 1997a).

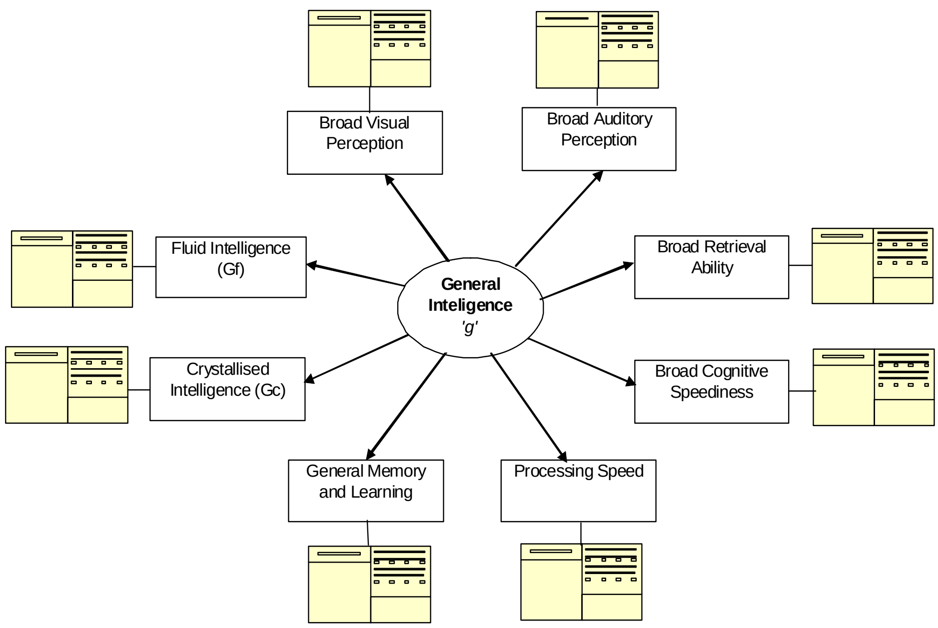
Alfred Binet's and Theophile Simon's development of the first usable mental ability test (1905) marked the beginning of the psychometric testing tradition. The instrument was initially developed to separate children with learning difficulties from normal children. The test was later refined and renamed the Stanford-Binet Test. The creation and use of the Army Alpha and Beta Tests with the United States Army recruits signalled the beginning of adult ability testing (Ackerman, 1996).

Louis Thurstone (1919) argued that there were about seven primary mental abilities, e.g. memory, spatial visualisation, number facility, verbal comprehension instead of a single factor '*g*' (cf. Neisser et al., 1996).

Raymond Cattell (1943, 1963) proposed two types of cognitive abilities, '*fluid*' (Gf) and '*crystallized'* (Gc), that account for differences between adolescents and adults. Fluid intelligence was defined as the capacity to think logically, reason abstractly and solve problems independently, all of which are abilities necessary for mathematical and scientific problem solving. Fluid intelligence or the 'mechanics' of our intellectual functioning, peaks in adolescence and declines slowly with age (Baltes & Staundinger, 1993, 2000). Crystallized intelligence or the ability to use knowledge, skills and experiences, usually increases with age as a result of acquired knowledge and life experience (cf. Ackerman et al., 2000; Baltes & Staundinger, 1993, 2000; Baltes & Schaie, 1976; Deary, 2001; Deary et al., 2003; Hunt, 2000; McArdle et al, 2000).

John Carroll (1993) produced the most comprehensive overview of cognitive abilities with his Three-Stratum Theory of Intelligence (see Figure 2), where '*g*' is the general factor or the third and highest stratum, accounting for about half of individual differences in scores of groups of people. The second stratum consists of eight group factors that represent narrower abilities (Deary, 2001). Highly specific skills make up the first stratum (Gottfredson, 1997a,b). The model provides evidence for individual differences in intelligence.

Figure 1.2.2: Hierarchical Representation of Mental Ability Test Scores Based on John B. Carroll's Three Stratum Theory of Intelligence (Carroll, 1993 in Deary, 2001, p.14).



1.2.2.1. Intelligence and Age

Studies investigating the relationship between age and intelligence (cf. Baltes & Schaie, 1976; Beier & Ackerman, 2001, 2003; Deary ,2001; Facon, 2006; Hartmann, 2006; Matthews,et al., 2000; McArdle et al., 2000; Stankov et al., 1995) tend to conclude that fluid ('*Gf*') and crystallised ('*Gc*') intelligence (Cattell, 1963) differ from each other in terms of growth and decline. *Gf* peaks in late adolescence and declines with age, while *Gc* increases throughout life and remains stable in older age (e.g. Beier & Ackerman, 2001). Thus, middle aged and older adults have been shown to be more knowledgeable than younger adults in almost all domains (cf. Ackerman, 1996, 2000; Ackerman & Rolfhus, 1999; Beier & Ackerman, 2001, 2003). While fluid intelligence is an important predictor of learning outcomes, crystallised intelligence is a key predictor of knowledge and the best predictor of higher educational success (Ackerman, 2006; Beier & Ackerman, 2001, 2003). Most intelligence tests measure fluid and crystallized intelligences (McArdle et al., 2000).

1.2.2.2. Gifted and Highly Precocious Individuals

For the past century researchers associated *'giftedness'* or '*intellectual precocity'* with high IQ (Terman, 1925). However, the terms *'profound giftedness'*, *'high potential',* or *'talent'* have recently been broadened to include multidimensional elements, such as rapid learning, attention control, memory efficiency, desire to develop one's gifts, task commitment, and ability to self-regulate (cf. Halpern et al., 2007; Lubinski & Benbow, 2000; Reis & Renzulli, 2010, p. 308; Sternberg & Davidson, 2005). Equally, gifted individuals are found in all walks of life and among all socio-economic, ethnic, racial and language groups. Yet, more males than females are classified as mathematically gifted, with data suggesting that this is due to a bigger male mean and more male variability on advanced math tests (Feingold, 1988, 1996).

Despite the interest into gifted or talented individuals, and in particular children, concerns have been raised in the United States about the lack of tailor-made education and talent development programmes for the most gifted (Benbow & Stanley, 1996). U.S. data shows that the most precocious students are falling behind in terms of achievement level and potential in comparison with other developed nations (Benbow & Stanley, 1996). So, whilst the overall IQ has increased steadily with each subsequent generation since 1932 (cf. Flynn, 1987) and better SAT test results are being achieved by average students (e.g. Cole, 1997; Herrnstein & Murray, 1994), the academic results of America's most gifted have been progressively decreasing since the 1960s. Coercive egalitarianism, (Schroeder-Davis, 1993), anti-intellectualism, and dumbing-down of the curriculum have all been blamed for the decline in academic performance of the most talented students (Benbow & Stanley, 1996). International data on gifted education shows rather varied approaches and results in each individual country (Kim, 2006).

Studies comparing gifted and normal populations found little differences between the two groups. Equally, differences between high ability males and females appear to emulate gender differences observed in normal ability groups (Roznowski, Reith, & Hong, 2000), with two notable differences. Firstly, precocious students are less stereotyped in their beliefs about typical feminine and masculine abilities. Secondly, gifted girls interests resemble those of normal males (Lubinski & Humphreys, 1990).

In addition, gifted (high) school students participate in more preparatory courses, follow more math/sciences classes, believe that math/English courses are important for their future, have more professional aspirations, receive more guidance from parents, are from families with above average SES, achieve highest academic grades, work harder at school and on homework, like school more, have higher self-esteem, watch less TV and have more positive attitudes toward female careers than the average students (Halpern et al., 2007; Roznowski, Reith, & Hong, 2000, p. 108). Also, precocious adolescent students who excel in verbal skills are attracted towards humanities and social sciences, whilst those with superior mathematical and spatial skills lean towards physical sciences and engineering (Shea, Lubinski, & Benbow, 2001). These choices resemble the choice patterns seen in the normal population.

Yet, the most striking gender difference in the high ability population is the fact that gifted females appear to have lower educational and career expectations, despite regularly academically outperforming males (Lubinski & Humphreys, 1990; Roznowski, Reith, & Hong, 2000). In fact, highly gifted females tend to have better grades than highly gifted males, but lower academic self-concepts influence their career choices (Preckel et al., 2008). Ferriman et al. (2009) affirmed in their longitudinal study of profoundly gifted math/science graduates that differences in career and life choices are mediated by parenthood. These findings were supported by Benbow et al. (2000) who found that women were torn between family-childcare and career achievement conflicts, leading to less successful careers. These findings provide further evidence as to why highly educated and profoundly gifted females, upon becoming parents, often choose family and community roles over careers. Hence while gifted males choose linear career paths that bring about status, gifted females embrace multiple social roles that lead to cyclical career paths (Benbow et al., 2000; Roznowski, Reith, & Hong, 2000; Xie & Shauman, 2003). Again, these findings mirror the choices of normal population.

1.2.3. Sex Differences in Intelligence

Research on sex differences in cognitive abilities has been the subject of passionate debate. Overall, the research programmes are either concerned with the existence of sex differences or their causes (Feingold, 1996, p. 25). Studies investigating the existence of sex differences typically focus on meta-analyses of former studies (cf. Halpern, 2000; Halpern et al., 2007; Hyde, Fennema, & Lamon, 1990a for mathematical abilities; Hyde & Linn, 1988 for verbal abilities; Linn & Peterson, 1985 and Voyer, Voyer, & Bryden, 1995 for spatial abilities), observed effect sizes and their moderators, and standardised test norms (Feingold, 1996, p. 25). Studies concerned with the causes are divided into nature-nurture advocates which often provide contrasting theoretical explanations for the observed sex differences (cf. Ackerman et al., 2001; Baenninger & Newcombe, 1995; Casey, 1996; Crawford, Chaffin, & Fitton, 1995;. Dykiert et al., 2009; Hyde, 1996; Lynn, 1999; Spelke, 2005; Steele & Ambady, 2006; Steinmayr et al., 2010).

It should be noted that to date no definitive answers as to the origins and causes of sex differences in intelligence are available. However, recent statistical advancements have made it possible to establish that diverse methodological approaches influence the degree of the observed cognitive sex differences (Brunner, Krauss, & Kunter, 2008; Dykiert et al., 2009; Johnson & Bouchard, Jr., 2007; Steinmayer et al., 2010; van der Sluis et al., 2008). For instance, Steinmayr, Beauducel, & Spinath (2010) found male advantage on verbal, numerical and figural intelligences when they were measured as manifest sum scores and female advantage on verbal intelligence when it was measured as factor score estimates or latent variables. These two methods also reduced the observed effect sizes. Interestingly, sex differences in fluid and crystallised intelligences were not influenced by either method, with reported male advantage on both. Brunner, Krauss, & Kunter (2008) reported significant increase in effect size for the male math advantage when the nested-factor model was used (.94 < *d* <.1.16) compared to the small effect size (*d* =.35) observed with the standard model. Similarly, it has been proposed that some of the male advantage on general cognitive ability ('*g'*) is related to the greater male variance at the higher and lower ends of IQ distribution as well as sample restriction (Arden & Plomin, 2006; Carr et al., 2008; Dykiert et al., 2009; Halpern et al., 2007; Johnson & Bouchard, Jr., 2007)

Furthermore, some researchers have proposed that sex differences in cognition are disappearing thanks to better educational practices and opportunities, socio-cultural changes and policies, experience and training (Baenninger & Newcombe, 1995; Crawford et al., 1995; Feingold, 1988, 1996; Halpern et al., 2007; Hyde, 1996; Spelke, 2005; Voyer, Voyer, & Bryden, 1995; Xie & Shauman, 2003). This proposition is supported by some cross-cultural data, but only for highly gifted populations (e.g. Deary et al., 2003).

The most accepted view on cognitive sex differences at the moment is that some differences decreased whilst others remained unchanged, with observable sex differences in the narrower areas of each stratum, e.g. female advantage in speech production, spelling and verbal fluency compared to the small sex differences in verbal abilities (cf. Cole, 1997; Feingold, 1988, 1996; Halpern & Wright, 1996; Halpern et al., 2007; Hyde & Linn, 1988). Besides, Feingold (1988) demonstrated that robust sex differences are least likely to decrease over time.

In summary, this section will provide a concise overview of the most pertinent findings about sex differences in intelligence as they are relevant to the subsequent sections on gender differences in self-estimated intelligence. Table 1.2.1. provides an overview of observed sex differences in specific intelligences.

1.2.3.1. Sex Differences in General Intelligence ('g')

Most studies on sex differences in intelligence have been conducted about differences in general intelligence ('*g*'), with no uniform agreement about male advantage (e.g. Colom & Garcia-Lopez, 2002; Deary et al., 2003; Halpern et al., 2007; Lynn, 1999; Spelke, 2005; van der Sluis et al., 2008). Historically, it has been accepted that no sex differences in general intelligence exist, as evidenced by the standardised intelligence test norms (Ackerman, 2006; Terman & Merrill, 1937; Wechsler, 1944). However, these assertions have been contested with claims of male advantage in general intelligence as supported by data on tests of fluid and crystallised intelligence, the General Knowledge Test (GKT), Naglieri Nonverbal Ability Test (NNAT), Raven's Standard and Advanced Progressive Matrices (SPM and APM), Scholastic Assessment Test (SAT), and Wechsler Adult Intelligence Scale (WAIS), (cf. Ackerman, Bowen, Beier, & Kanfer, 2001; Arden & Plomin, 2006; Deary et al., 2003; Jackson & Rushton, 2006; Lynn, Irwing, & Cammock, 2002; Lynn & Irwing, 2002; Lynn & Irwing, 2004; Lynn, Allik, & Irwing, 2004; Lynn, Wilberg, & Margraf-Stiksrud, 2004; Novell & Hedges, 1998; Rojahn & Naglieri, 2006).

It should be noted that the majority of the observed sex differences occur in early adolescence and adulthood, usually from the age of 14 years onwards and that no sex differences in '*g*' are observed in children (e.g. Arden & Plomin, 2006; Lynn & Irwing, 2004). In fact, the reported male advantage in '*g*' is 4 IQ points which equals to a small effect size (d =.12) (Jackson & Rushton, 2006).

1.2.3.2. Sex Differences in Verbal Abilities

Females tend to excel in verbal abilities but the differences are small (Cole, 1997; Feingold, 1996; Halpern & Wright, 1996; Halpern et al., 2007; Maccoby & Jacklin, 1974). Females have a strong advantage over males in verbal fluency, spelling/writing ($.50 < d < .60$), and language ($.40 < d < .50$). On the contrary, data from Differential Aptitude Test (DAT), Iowa Tests of Basic Skills (ITBS) and California Achievement Tests (CAT) reveal a small male advantage on verbal analogies (Halpern & Wright, 1996; Hyde & Linn, 1988). The female verbal advantage was confirmed across cultures (Ogle et al., 2003).

1.2.3.3. Sex Differences in Visuo-Spatial Abilities

Males outperform females in most areas of visuo-spatial abilities, where the biggest sex differences, as measured by effect sizes, are found (Halpern et al., 2007; Maccoby & Jacklin, 1974; Voyer, Voyer, & Bryden, 1995). The male advantage is presumed to contribute to the observed male advantage in mathematics and sciences, as documented by male superiority on standardised exams (e.g. SAT-M, GMAT).

Spatial sex differences are most pronounced in spatial perception, mental rotation and spatial visualisation (Linn & Petersen, 1985), with strong male advantage on the Mental Rotation Test (MRT) (Cooke-Simpson & Voyer, 2007; Vandenberg & Kuse, 1978; Voyer & Doyle, 2010; Voyer, Voyer, & Bryden, 1995). Female spatial advantage was documented in areas of fine motor skills, picture/face identification and perceptual/processing speed and writing (Caramata & Woodcock, 2006; Cole, 1997; Vlachos, Andreou, & Andreou, 2003). Since sex differences in visuo-spatial abilities appear in infancy, biological theories of brain lateralization, hormonal and genetic influences have been generally accepted as causes (cf. Caramata & Woodcock, 2006; Halpern et al., 2007; Pinker, 2002; Vlachos et al., 2003). However, some researchers argue that the observed decrease in sex differences in spatial abilities is caused by environmental,socio-cultural and educational changes (Baenninger & Newcombe, 1995; Casey, 1996; Crawford et al., 1995; Spelke, 2005; Vlachos et al., 2003).

1.2.3.4. Sex Differences in Mathematical Abilities

Girls outperform boys in math throughout elementary education, but the advantage is equalised by high school and reversed by college (e.g. Benbow et al., 2000; Duckworth & Seligman, 2005; Halpern et al., 2007). Some studies (e.g. Levine et al., 1999; Penner & Paret, 2008) have shown that the male math advantage appears much earlier than college. Data suggests that the male advantage appears as early as kindergarten, and this is particularly true for children with high parental education. On the other hand, other researchers (e.g. Spelke, 2005; Xie & Shauman, 2003) argue that no real sex differences in math abilities exist and that boys and girls show equal aptitude for mathematics.

Females excel at computing, algebra and speed, while males surpass in geometry, calculus and reasoning (Hyde, Fennema, & Lamon, 1990a). Girls do well on math tests when they are directly related to the taught curriculum, but this is not the case in boys (Halpern, 2000). Nonetheless, the female math disadvantage emerges when the taught content becomes more complex, i.e. from counting to multiplication/division and fractions. Therefore it is likely that matter complexity is responsible for girls' math underachievement (Gibbs, 2010).

Equally, boys achieve better math grades and higher scores on math achievement tests than do females (cf. Crawford et al., 2003; Halpern, 2000; Halpern et al., 2007; Hyde, 1996; Hyde, Fennema, & Lamon, 1990a, N = >1 million children; d =. 29). The reversal of the sex differences occurs when (advanced) math classes become optional, with significantly more males than females taking college level math classes (Cole, 1997; Hyde, 1996). Similarly, no sex differences were found in basic math skills, but strong male advantage was reported in advanced math skills (Carr et al., 2008; Crawford et al., 2003; Feingold, 1988; Halpern et al., 2007). The existing grade-test inequality in math and sciences is evidenced by female grade advantage and male advantage on the standardised tests, such as the Scholastic Aptitude Test (SAT) and Graduate Record Examination (GRE). These results are confirmed by the observed male advantage at the highest end of distribution on SAT-M and other standardised math tests (cf. Benbow et al., 2000; Deary et al., 2003; Halpern et al., 2007). It should be noted that the observed sex differences are small, between ($.16 < d < .32$) (Halpern et al., 2007; Hyde, Fennema, & Lemon, 1990a).

Table 1.2.1: Summary of the Findings of Five Meta-Analyses about Sex Differences in Specific Intelligences

Type of Intelligence	*d*
Adults and children	
Linn & Petersen (1985)	
Mental rotation	**.73**
Spatial perception	.44
Spatial visualisation	.15
Hyde & Linn (1988)	
Speech production	-.33
Reading comprehension	-.03
Vocabulary	-.02
Voyer, Voyer, & Bryden (1995)	
Mental rotation	**.56**
Spatial perception	.44
Spatial visualisation	.19
Adolescents	
Feingold (1988)	
Mechanical reasoning	**.76**
Spelling	-.45
Language	-.40
Perceptual speed	-.34
Spatial relationships	.15
Numerical ability	-.10
Abstract reasoning	-.04
Verbal reasoning	-.02
Hedges & Novell (1995)	
Science	.32
Perceptual speed	-.28
Spatial ability	.19
Mathematics	.16
Reading comprehension	-.09
Vocabulary	.06

Legend: Negative d value represents higher female scores and positive d value stands for higher male value; d =.20 is small, d =.50 is medium and d =.80 is large (Cohen, 1988, 1992). Medium and large effect sizes are in bold.

1.2.4. Multiple Intelligence Theories

Critics of the psychometric approach questioned the value of IQ tests and '*g*' (cf. Ceci, 1990; Gardner, 1983; Sternberg & Grigorenko, 2002). Most questioned is whether intelligence is much more complex and a wider concept and as such can not be successfully captured by a single measure. Another point of criticism is that although individuals differ in their talents and weaknesses, psychometric measures fail to capture these differences. Advocates of the multiple intelligence theory argue that '*g*' is a measure of academic ability (Gottfredson, 2004b p.174) and that other, non-*'g'* related intelligences are more important in day-to-day functioning (cf. Gardner, 1983, 1993, 1999; Sternberg et al., 2000). *'Flynn's effect'* (Flynn, 1984, 1987) or the unexplained generational increase in intelligence, was also used by the critics of the psychometric tradition as supporting evidence for the existence of multiple intelligences.

Whilst the initial intelligence debate centred around issues of heritability, genetics and environmental effects on intelligence (cf. Bouchard, 1998; Bouchard et al., 1990; Crawford et al., 1995; Feingold, 1988, 1996; Loehlin et al, 1997; Plomin, 1999; Plomin & Spinath, 2004; Spelke, 2005), it was the research of group differences, and in particular sex[2], gender, and race differences in intelligence that caused the most controversy (cf. Benbow, 1988; Benbow et al., 2000; Flynn, 1987; Furnham, 2001; Halpern, 2000, 2002; Herrnstein & Murray, 1994; Hyde et al.,

---

[2] Social science studies differentiate between the terms sex and gender. The term *'sex'* is used when referring to a biological construct and/or genetic traits, classifying individuals as males and females. The term *'gender'* refers to a social role or gender identity and was introduced by sexologist John Mooney in 1955; the term was not generally used until the feminist movement started to use gender to differentiate between biological sex and social identity/role. Thus, the English language and literature now differentiates between biological sex, psychological gender and social sex role (Wikipedia, Retrieved 17 November 2010).

1990a,b; Jensen, 1998; Lynn, 1999; Neisser et al., 1996; Pinker, 2002; Rushton & Jensen, 2005a,b).

Then almost three decades ago, the theories of multiple intelligences (Gardner, 1983, 1999), the triarchic theory of intelligence (Sternberg, 1985) and emotional intelligence (Goleman, 1995) have made their entry and had an immediate impact on the scientific debate. These *'novel'* theories shaped the public opinion about what intelligence is but also what behaviour(s) should be included in the measurement of intelligence.

Howard Gardner's theory of multiple intelligences was most warmly received by educational professionals. Nonetheless, it immediately became a subject of critical discussion by fellow psychologists (cf. Barnett & Ceci, 2002; Eysenck, 1994; Gottfredson, 2004a, b; Waterhouse, 2006; White, 2005).

Gardner, a developmental psychologist at Harvard University's Graduate School of Education, rejected the single intelligence factor concept as insufficient and incomplete. He also argued against the traditionalist psychometric measurement of cognitive skills. In fact, Gardner proposed that the existing psychometric tests are only capable of verbal/linguistic, mathematical/logical and spatial intelligence assessment and leave *'other'* intelligences, such as musical, body-kinaesthetic and interpersonal, un-assessed. Consequently, he argued that in day-to-day life individuals must deploy multiple skills in order to manifest their intelligence and proposed that our understanding of human intelligence should come from the understanding of the person-environment interaction and not from artificial intelligence tests (Gardner, 2006).

Gardner's suggestion that *'everybody is smart in some way'* was instrumental in gaining wide public and professional (pedagogues) acceptance and popularity (cf. Gottfredson, 2004a,b). In his book *Frames of Mind* (1983) that is based on neuro-psychological studies and cross-cultural surveys of gifted individuals, savants, musical virtuosos, Gardner published an overview of seven independent and equal intelligences: verbal/linguistic, mathematical/logical, spatial, musical, body-kinaesthetic, interpersonal and intrapersonal. Two more intelligences – naturalist and existential – were added in 1999. Gardner also wanted to add spiritual/moral intelligence but did not find enough supporting evidence. Each intelligence is supposedly governed by its own perception, learning and memory (Weinberg, 1989, p. 99), as seen in top athlete's superior body-kinaesthetic intelligence or laureate poets' enhanced verbal/linguistic intelligence.

Critics of Gardner's multiple intelligence theory argued that the majority of proposed multiple intelligences are not intelligences but instead talents, particular cognitive qualities, forms of accomplishment or traits (cf. Barnett & Ceci, 2002; Eysenck, 1994; Gottfredson, 1997a, 2004a,b; Klein, 1997; Lohman, 2001; Mathews et al., 2000; Scarr, 1985; Sternberg, 1985, 1991; Visser et al., 2006a, 2006b; Waterhouse, 2006; White, 2005). For example, inter- and intrapersonal intelligences were shown to be more related to personality traits than intelligence (Lohman, 2001). Gardner's refusal to acknowledge the hierarchical structure of intelligence was also heavily criticised (e.g. Barnett & Ceci, 2002; Waterhouse, 2006; Weinberg, 1989; White, 2005). Yet, Gardner's verbal/linguistic, mathematical/logical, spatial and musical intelligences correlated with the general intelligence factor *'g'* and were intercorrelated with each other (Gottfredson, 2004b; Visser et al., 2006a, 2006b). However, Gardner came under most attack for the lack of formal measures and data supporting his theory. In effect, the theory was never tested or validated by Gardner or peers (Sternberg & Grigorenko, 2002; Visser et al, 2006b; Waterhouse, 2006a,b).

Robert Sternberg's (1985) triarchic theory postulates three fundamental aspects of intelligence – analytic, creative and practical. Creative intelligence is needed when a person has to respond innovatively. Practical intelligence involves selecting and adopting to a particular environment. Only the analytic aspect of Sternberg's theory can be measured by existing psychometric tests.

Sternberg argued that intelligence is an ability to learn from experience and capability to apply this knowledge in novel situations (Weinberg, 1989). Sternberg's theory was subjected to rigorous criticism by scientific community for its unempirical nature and lack of supporting data (e.g. Brody, 2003; Gottfredson, 2003a,b).

Daniel Goleman's book *'Emotional Intelligence'* (EI) (1995) was an immediate bestseller. As with Gardner (1983) and Sternberg (1985), Goleman's theory argues that *'g'* is only one of co-equal cognitive abilities that are necessary for successful functioning and survival. Emotional intelligence is described by Goleman as a combination of numerous traits, skills and competencies that facilitate leadership performance. The model consists of four components: self-awareness, self-management, social-awareness, and relationship management, with a number of emotional competencies within each component (Goleman, 1995, 1998). These competencies are acquired and mastered through practice. The EI theory postulates that individuals are born with a general emotional intelligence that predicts their aptitude to develop emotional competencies (Boyatzis et al., 2000). Some researchers (cf. Petrides et al., 2007; Waterhouse, 2006) believe that EI is an amalgamation of *'g'* with personality traits Agreeableness, Extraversion, and Neuroticism (Costa &

McCrae, 1980). Despite the development of two EI measures, the theory was labelled *'pop psychology'* (Mayer, Roberts, & Barsade, 2008).

Table 1.2.2: Gardner's Multiple Intelligences Model and Definitions (Gardner, 1983, 1999).

Gardner's Multiple Intelligences	*Gardner's Definitions (1983, 1999)*	*SAI Measure Definitions (Furnham & Gasson,1998) 1998)*
Verbal/ Linguistic	This area has to do with words, spoken or written. People with high verbal-linguistic intelligence display a facility with words and languages. They are typically good at reading, writing, telling stories and memorizing words along with dates. They tend to learn best by reading, taking notes, listening to lectures, and discussion and debate. Those with verbal-linguistic intelligence learn foreign languages very easily as they have high verbal memory and recall, and an ability to understand and manipulate syntax and structure. Careers that suit those with this intelligence include writers, lawyers, policemen, philosophers, journalists, politicians, poets, and teachers.	Ability to use words
Mathematical/ Logical	This area has to do with logic, abstractions, reasoning, and numbers. While it is often assumed that those with this intelligence naturally excel in mathematics, chess, computer programming and other logical or numerical activities, a more accurate definition places less emphasis on traditional mathematical ability and more on reasoning capabilities, abstract patterns of recognition, scientific thinking and investigation, and the ability to perform complex calculations. It correlates strongly with traditional concepts of "intelligence" or IQ. Careers which suit those with this intelligence include scientists, physicists, mathematicians, logicians, engineers, doctors, economists and philosophers.	Ability to reason logically, solve a number problem
Spatial	This area deals with spatial judgment and the ability to visualize with the mind's eye. Careers which suit those with this type of intelligence include artists, designers and architects. A spatial person is also good with puzzles.	Ability to find your way around in the environment and form mental images
Musical	This area has to do with sensitivity to sounds, rhythms, tones, and music. People with a high musical intelligence normally have good pitch and may even have absolute pitch, and are able to sing, play musical instruments, and compose music. Since there is a strong auditory component to this intelligence, those who are strongest in it may learn best via lecture. Language skills are typically highly developed in those whose base intelligence is musical. In addition, they will sometimes use songs or rhythms to learn. They have sensitivity to rhythm, pitch, meter, tone, melody or timbre. Careers that suit those with this intelligence include instrumentalists, singers, conductors, disc-jockeys, orators, writers and composers.	Ability to perceive and create pitch and rhythm
Body-Kinaesthetic	The core elements of the bodily-kinaesthetic intelligence are control of one's bodily motions and the capacity to handle objects skilfully (206). Gardner elaborates to say that this intelligence also includes a sense of	Ability to use bodily functions or motor movements

	timing, a clear sense of the goal of a physical action, along with the ability to train responses so they become like reflexes. In theory, people who have bodily-kinaesthetic intelligence should learn better by involving muscular movement (e.g. getting up and moving around into the learning experience), and are generally good at physical activities such as sports or dance. They may enjoy acting or performing, and in general they are good at building and making things. They often learn best by doing something physically, rather than [by] reading or hearing about it. Those with strong bodily-kinaesthetic intelligence seem to use what might be termed muscle memory - they remember things through their body such as verbal memory. Careers that suit those with this intelligence include: athletes, dancers, musicians, actors, surgeons, doctors, builders, police officers, and soldiers. Although these careers can be duplicated through virtual simulation, they will not produce the actual physical learning that is needed in this intelligence.	
Inter-personal	This area has to do with interaction with others. In theory, people who have a high interpersonal intelligence tend to be extroverts, characterized by their sensitivity to others' moods, feelings, temperaments and motivations, and their ability to cooperate in order to work as part of a group. They communicate effectively and empathize easily with others, and may be either leaders or followers. They typically learn best by working with others and often enjoy discussion and debate. Careers that suit those with this intelligence include sales, politicians, managers, teachers, and social workers.	Ability to understand other people
Intra-personal	This area has to do with introspective and self-reflective capacities. People with intrapersonal intelligence are intuitive and typically introverted. They are skilful at deciphering their own feelings and motivations. This refers to having a deep understanding of the self; what are your strengths/ weaknesses, what makes you unique, you can predict your own reactions/ emotions. Careers which suit those with this intelligence include philosophers, psychologists, theologians, lawyers, and writers. People with intrapersonal intelligence also prefer to work alone.	Ability to understand yourself and develop sense of your own identity
Naturalistic	This area has to do with nature, nurturing and relating information to one's natural surroundings. Careers which suit those with this intelligence include naturalists, farmers and gardeners.	Ability to identify and employ many dimensions in the natural world, e.g. classifying animals and plants
Existential	Ability to contemplate phenomena or questions beyond sensory data, such as the infinite and infinitesimal. Ideal careers: cosmologist, philosopher.	Ability to understand the significance of life, the meaning of death, and the experience of love
[Spiritual]		[Ability to engage in thinking about cosmic issues, the achievement of a state of being and the ability to have spiritual effects on others]

Legend: SAI measure definitions (Furnham and Gasson, 1998) were used in this thesis and include definition of spiritual intelligence.

1.2.5. Objective and Subjective Self-Assessments

Whilst intelligence research is considered '*objective*', the self-estimated intelligence literature was subjected to considerable criticism (DeNisi & Shaw, 1977; Mabe & West, 1982; Paulhus et al., 1998) for being '*subjective*' and unreliable, thanks to low mean validity coefficients. In fact, the entire correlational literature, i.e. literature of individual differences, has been accused of being unreliable and subjective, in particular by situationalists (cf. Mischel, 1968; Mischel & Shoda, 1995).

In the last two decades, a number of social, cognitive and applied psychologist research programmes focused on objectivity/subjectivity, accuracy, validity and biases in self-assessments and self-ratings (cf. Ackerman et al., 2002; Ackerman & Wolman, 2007; Alicke et al., 1995; Allik et al., 2010; Borkenau & Liebler, 1993; Critcher & Dunning, 2009; Forbes & Schmader, 2010; Guenther & Alicke, 2010; Hall & Carter, 1999; Kim et al., 2010; Krueger & Mueller, 2002; Moore & Small, 2007; Swim, 1994).

Conscious and unconscious self-assessments of abilities are made by individuals on a daily basis (cf. Ackerman et al., 2002; Ackerman & Wolman, 2007; Borkenau & Liebler, 1993). Interestingly, self-ratings stay stable whether they are made independently or in reference to a peer group and are moderated by self-enhancement motives (Guenther & Alicke, 2010; Klar & Giladi, 1999). Evidence also suggests that strangers' ratings of one's personality and intelligence are rather accurate (Borkenau & Liebler, 1993; Allik et al., 2010), with medium rank-order correlation r =.43 (both studies).

Thus, *'subjective'*, i.e. self-assessed, and *'objective'*, i.e. psychometrically tested abilities are concurrent, partly because individuals have a fair understanding of their strengths and weaknesses in various aptitude domains (Ackerman et al., 2002). Correlations between self-assessed and actual abilities are reported to range between $.26 < r < .55$ for general ability, $.41 < r < .50$ for mathematical, and $.39 < r < .52$ for spatial abilities (Ackerman & Wolman, 2007; Bailey & Lazar, 1976; DeNisi & Shaw, 1977; Mabe & West, 1982).

The current working hypothesis proposes that self-estimates of intelligence should not substitute psychometric tests, since self-assessed and objective intelligences only share a little variance, with correlation coefficients, usually around $r =.30$ (Furnham, 2001; Furham, von Stumm, Makendrayogam, & Chamorro-Premuzic, 2009; Holling & Preckel, 2005; Mabe & West, 1982; Paulhus et al., 1998).

However, a meta-analysis of self-estimates of intelligence (Poropat, 2010, unpublished manuscript) that analysed 149 effect sizes reported an overall effect size ranging between $.32 < r < .67$, with standard estimation error of 14.21 to 11.14 IQ points, concluded that self-estimates can be used as valid proxy measures of intelligence test scores. Equally, Kornilova et al. (2009) demonstrated that subjective estimates of intelligence (SEI) and academic self-concept were the strongest predictors of academic achievement, over and above ‘real’ IQ measures. Similarly, Kim et al. (2010) reported a strong positive relationship ($r = .81$), between actual and self-reported math task performance, but only after subjects were instructed to focus on accurate estimation. These findings are in line with Holling and Preckel’s study (2005), where participants gave more accurate self-estimates than expected.

Finally, areas of psychology where self-estimates are regarded to be as valid as psychometric measures are occupational psychology and career counselling (Gati et al., 2006; Gottfredson, 2003c; Prediger, 1999).

1.2.5.1. Accuracy of Self-Assessments of Performance

The accuracy of self-estimates debate is more than 100 years old (Ackerman & Wolman, 2007). Social psychologists have long argued that people hold inflated and overly favourable self-views and ability beliefs. As a result, their performance estimations are subjected to systematic estimation errors (e.g. Kruger & Dunning, 1999; Harrison & Shaffer, 1994; Krueger & Mueller, 2002; Lichtenstein & Fischhoff, 1977).

The inclination to predictably over- or underestimate one's intelligence, also called the '*Downing effect'* was demonstrated in various cultures (Davidson & Downing, 2000). Thus, individuals with below average intelligence overestimate their abilities, while individuals with above average capabilities underestimate theirs. Moreover, researchers established that the ability to accurately estimate one's intelligence was correlated to one's IQ, i.e. the higher one's IQ, the more correct the estimation.

In fact, overinflated beliefs are more likely to occur on easy tasks, the so-called *'Better-Than-Average effect'* (BTAE). Equally, when tasks are difficult or success unlikely, people tend to believe that they are worse than others. This outcome is called *'Worse-Than-Average effect'* (WTAE) (Alicke et al., 1995; Guenther & Alicke, 2010; Krueger & Mueller, 2002; Moore & Small, 2007).

Data shows that the level of task abstraction and ambiguity shapes the comparison bias. In other words, if the comparison target is personally known to the person making the comparison or the task is specific, the BTAE is reduced (Ackerman et al., 2002; Alicke et al., 1995). Data also suggests that individuals believe they are better than others when the task is easy and worse than others when the task is difficult. BTAE and WTAE effects are more pronounced in situations when personal performance feedback is accessible but improve when peer performance feedback is available (Moore & Small, 2007).

This in turn leads to overestimation of own abilities when peer comparisons are made, i.e. the better-than-average bias (e.g. Guenther & Alicke, 2010; Kruger & Dunning, 1999). However, the ability to correctly evaluate one's aptitude vis-à-vis others, is vital for academic and professional success (Burson et al., 2006). Thus, the ability to correctly appraise one's self-perceptions is vital for successful self-regulation (Beyer, 1998).

Evidence shows that the worst performers tend to overestimate their performance (cf. Burson et al., 2006; Ehrlinger & Dunning, 2003; Kruger & Dunning, 1999), while competent performers estimate accurately. Kruger & Dunning (1999) demonstrated that not only are incompetent individuals inclined to overestimate their skills, they also fail to acknowledge their own incompetence.

It is also true that individuals tend to overestimate their performance on easy tasks, but underestimate their accomplishments on difficult tasks (Kruger & Dunning, 1999, 2002; Moore & Small, 2007). What's more, Ehrlinger & Dunning (2003) demonstrated that predefined self-perceptions of ability mediate one's estimates of performance. In fact, these self-perceptions correlate more strongly with one's self-estimates of performance than do the actual achievements (Ehrlinger & Dunning, 2003).

Other cognitive biases such as *'anchoring'* and *'availability heuristic'* (cf. Ariely, 2008; Kahneman & Tversky, 2000;) were shown to influence the accuracy of people's judgements. In *'anchoring'* individuals *'anchor'* previous experience and use the *'anchor'* to approximate their future performance. Guenther and Alicke (2010) demonstrated that self-estimates serve as anchors and are moderated by self-enhancement motives. '*Availability heuristic',* or the particular judgement criteria that become most quickly available during self- and peer judgement situations (Kahneman & Tversky, 2000), was also shown to distort judgements (Dunning et al., 1989).

Nonetheless, several studies (e.g. Ackerman et al., 2002; Chamorro-Premuzic et al., 2009) demonstrated that individuals were capable of accurate self-estimates of ability. Further support was provided by Chamorro-Premuzic et al. (2010) in their extensive study of 5957 British children that revealed that self-perceived abilities are valid measures of intellectual competence. Equally, contrary to popular belief about inaccuracy of gender stereotypes, individuals were more likely to be accurate or to underestimate gender differences than to overestimate them (Hall & Carter, 1999; Swim, 1994).

1.2.6. Self-Estimated Intelligence

Over the last 25 years a great number of studies investigated self-estimated intelligence[3]. Self-assessments of abilities are expressions of person's self-concept that reveal a degree of self-insight (Harrington & Schafer, 1996; Kornilova et al., 2009). Applied psychologists have been interested in lay theories of intelligence for numerous reasons, such as gaining insight into peoples' understanding of individual differences in intelligence, the development and stability of the theories over time, the

---

[3] A number of compatible terms that relate to one's self-assessed/estimated abilities have been used by various researchers, adding to the confusion about the meaning and relevance of the phenomenon, i.e. self-estimates of intelligence, self-assessed/estimated/perceived/evaluated intelligence, perceptions of intelligence, academic self-confidence/competence/competencies/esteem/concept/efficacy/ belief/ beliefs. Chamorro-Premuzic et al. (2010) proposed to use a collective term self-perceived abilities (SPA). As noted in Footnote 1 this thesis uses SAI/SEI and self-assessed abilities interchangeably throughout.

effects of popularising academic theories on society, and the impact of the self-estimates on performance expectations and evaluations (Furnham, 2001; Furnham et al., 2009). Many studies have shown that lay beliefs about intelligence are time- and culture-specific (Furnham, 2001, p. 1401). Consequently, major findings about self-, others'-, and parental estimates of ability and the role of culture will be summarised in the following sections.

1.2.7. Self-, Others'- and Parental Estimates of Ability

The initial self-estimation of intelligence studies were conducted by Hogan (1978). He asked American students to estimate theirs and their parents' overall intelligence as well as the intelligence of men and women in general. Hogan found that men gave higher self-estimates than females. Many studies followed (e.g. Beloff, 1992; Bennett, 1996, 1997; Rammstedt & Rammsayer, 2000 a, b, 2002a,b), including a systematic research programme by Furnham and his collaborators, replicating Hogan's results.

The above-mentioned programme was extended to include Gardner's (1983, 1999) multiple intelligence model. Thus, Bennett (1996) used six multiple intelligences, Furnham and his collaborators (e.g. Furnham, Clark, & Bailey, 1999; Furnham, Fong, & Martin, 1999) initially seven, and later ten multiple intelligences (Furnham, 2002; Neto, Furnham, & Paz, 2007).

Studies examining IQ self-estimates of grandparents, parents, siblings and childrens' intelligence are interesting because they also provide insight into gender stereotypical beliefs about intelligence. Data suggests that parental gender, age and socio-economic status, i.e. education and income, influence self-estimates for sons and daughters (cf. Furnham, 2001; Furnham, Clark & Bailey, 1999; Furnham, Mkhize, & Mndaweni, 2004; Furnham, Rakow, & Mak, 2002; Rammstedt & Rammsayer, 2002b).

The pattern of gender inequality in IQ estimation was replicated for parents, siblings and grandparents, with mothers estimated as inferior to fathers by both sexes (cf. Hogan, 1978; Beloff, 1992; Byrd & Stacey, 1993; Furnham & Bunclark, 2006; Petrides, Furnham, & Martin, 2004; Furnham, Rakow, Sarmany-Schiller, & De Fruyt, 1999; Furnham, Rakow & Mak, 2002; Furnham & Rawles, 1995). Among siblings, brothers provided higher self-estimates than did their sisters, (cf. Byrd & Stacey, 1993; Furnham & Rawles, 1995) and parents rated their sons as more intelligent than their daughters (e.g. Furnham, 2000; Furnham & Budhani, 2002; Furnham, Fong, & Martin, 1999; Furnham & Gasson, 1998; Furnham, Hosoe, & Tang, 2002; Furnham, Reeves & Budhani, 2002; Furnham & Thomas, 2004; see also Table 1.2.3.).

Equally, grandfathers thought themselves brighter than grandmothers (Furnham & Rawles, 1995, 1999). Where possible, the calculated Cohen's *d* effect sizes (Cohen, 1988, 1992) ranged from medium to large (Furnham, 2001, p.1389). Table 1.2.3. gives an overview of self, parental, other's and relatives' overall, mathematical and spatial self-estimate studies.

Table 1.2.3: Results of Studies with Overall (g), Mathematical-Logical and Spatial IQs Rating of Self and Others

Study	*Country*	*Sample*	*Men*	*Women*	*Difference*
Overall IQ					
Beloff (1992)	Scotland		n = 502	n =265	
Self			126.90[a]	120.50	6.40
Mother			118.70[a]	119.90	-1.20
Father			125.20[a]	127.70	-2.50
Bennett (1996)	Scotland		n = 96	n = 48	
Self			117.10[a]	109.40	7.70
Byrd & Stacey (1993)	N. Zealand		n = 105	n = 112	
Self			121.50	121.90	-0.40
Mother			106.50[a]	114.50	-8.00
Father			122.30[a]	127.90	-5.60
Sister			110.50[a]	118.20	-7.70
Brother			116.00[a]	114.10	1.90
Furnham and Gasson (1998)	England	Normal population	n = 112	n = 72	
Self			107.99[a]	103.84	4.15
Male child (1st child)			109.70[a]	107.69	2.01
Female child (1st child)			102.36[a]	102.57	-0.21
Furnham & Petrides (2004)	England	Normal population	(n =82)	(n = 138)	
Overall IQ			108.90[a]	106.60	2.30
Emotional IQ			105.40	111.20[a]	-5.80
Furnham and Rawles (1995)	England	University	n = 84	n = 161	
Self		Students	118.48[a]	112.31	6.17
Mother			109.42[a]	108.70	0.72
Father			116.09[a]	114.18	1.91
Furnham and Rawles (1999)	England		n = 140	n = 53	
Self			120.50[a]	116.64	3.86
Psychometric IQ (WAIS)			6.94[a]	4.47	2.47
Furnham, Reeves, and Budhani (2002)	England		n = 84	n = 72	
Self			110.15[a]	104.84	5.31
Male child (1st son)			114.32	116.09	-1.77
Female child (1st daughter)			104.32	110.66[a]	-6.34
Reilly and Mulhern (1995)	Ireland		n =80	n = 45	

Self Estimated IQ			113.90[a]	105.30	8.60
Psychometric IQ (WAIS)			106.10	106.90	-0.80
Mathematical and Spatial IQs			*Men*	*Women*	*Difference*
Furnham (2000)	England	Normal population	(n = 46)	(n = 66)	
Mathematical IQ			110.54[a]	102.66	7.88
Spatial IQ			111.84[a]	104.81	7.03
Furnham (2004)	England	Normal population	(n = 94)	(n = 141)	
Numerical/Mathematical IQ			110.10[a]	101.90	8.20
Spatial IQ			111.40[a]	103.50	7.90
Furnham & Budhani (2002)	England	School	(n =136)	(n = 149)	
Self (G+B) Overall (O) IQ		children & parents	102.01[a]	99.74	2.27
Mathematical IQ			101.54[a]	95.55	5.99
Spatial IQ			105.63[a]	101.97	3.66
			(n = 45)	(n = 52)	
Mother's est. of ds. & sons (O)			105.34	105.33	0.01
Mathematical IQ			104.89	101.87	3.02
Spatial IQ			103.93	104.65	-0.72
			(n = 25)	(n = 35)	
Father's est. of ds. & sons (O)			111.20	108.14	3.06
Mathematical IQ			106.80	104.06	2.74
Spatial IQ			108.72	103.71	5.01
Furnham & Bunclark (2006)	England	British parents & school children	(n = 61)	(n = 84)	
Mathematical IQ			111.46[a]	100.48	10.98
Spatial IQ			110.97[a]	100.36	10.61
Furnham, Clark & Bailey (1999)	England	Normal population	(n = 89)	(n = 91)	
Mathematical IQ			112.11[a]	104.04	8.07
Spatial IQ			110.50	107.94	2.56
Furnham & Thomas (2004)	England	Normal population	(n =84)	(n = 138)	
Overall IQ			110.65[a]	106.15	4.50
Mathematical IQ			113.95[a]	103.02	10.93
Spatial IQ			114.42[a]	105.04	9.38
Holling & Preckel (2005)	Germany	High school students	(n = 37)	(n = 51)	
Number Series (Abstract Reasoning)			121.53[a]	117.01	4.52
Figure Matching (A. reasoning)			114.99[a]	105.49	9.50
Spatial Orientation			118.20[a]	108.93	9.27
Memory			117.42[a]	110.80	6.62

Rammstedt & Rammsayer (2000b)	Germany	University Students	(n = 54)	(n = 51)	
Mathematical IQ			119.10[a]	107.00	12.10
Spatial IQ			119.10[a]	104.50	14.60
Overall IQ			114.10	111.90	2.20
Rammstedt & Rammsayer (2002a)	Germany	University & college students	(n = 135)	(n = 132)	
Mathematical IQ			113.80[a]	105.60	8.20
Spatial IQ			114.90[a]	106.30	8.60
Rammstedt & Rammsayer (2002b)	Germany	University & college students	(n = 121)	(n = 107)	
Low Education Level					
Mathematical IQ			110.30[a]	99.90	10.40
Spatial IQ			115.20	108.40	6.80
High Education Level					
Mathematical IQ			115.10[a]	111.00	4.10
Spatial IQ			113.90	105.90	8.00

Legend: [a] = Indicates significant gender differences in that cell; provided where reported by the authors.

1.2.8. Self-Estimated Intelligence and Culture

Various international studies about the cross-cultural understanding of intelligence (e.g. Favia & Fontane, 1997; Segall, Dasen, Berry, & Poortinga, 1999; Swami et al., 2008; Sternberg, 1990; Yang & Sternberg, 1997; Wober, 1973) reveal that laymen's concept of intelligence is broader than that of experts. As such, Asians incorporate speed of thinking and judgment into their definition of intelligence, whereas Africans include broader social factors, such as co-operation and wisdom (Furnham & Baguma, 1999; Furnham & Akande, 2004; Segall et al., 1999; Yang & Sternberg, 1997; Wober, 1973).

Secondly, there are clear cultural differences in lay definitions of intelligence. In addition, the least educated individuals seem most prone to culturally specific stereotypical beliefs (Furnham, Shahidi, & Baluch, 2002).

The main findings in the cross-cultural SEI literature that are based on studies from more than 20 countries, are that self-estimates vary across cultures, with lower estimates provided by Asian, North African and Latin American participants compared to Americans and European (cf. Furnham & Chamorro-Premuzic, 2005; Furnham & Fong, 2000; Furnham, Fong, & Martin, 1999; Furnham, Hosoe, & Tang, 2001; Furnham & Mottabu, 2004; Furnham, Rakow, Sarmany-Schiller, & De Fruyt, 1999; Furnham, von Stumm, Makendrayogam, & Chamorro-Premuzic, 2009; Neto, Furnham, & Paz, 2006). Secondly, across all cultures men rate themselves higher on overall (‘*g*’), mathematical/logical, and spatial, intelligences (e.g. Furnham & Chamorro-Premuzic, 2005; Furnham & Fukumoto, 2008; Furnham, Mkhize, & Mndaweni, 2004; Furnham & Mottabu, 2004; Swami et al., 2006; Yuen & Furnham, 2005; also see Table 1.2.4.), with exception of few studies from Africa (Furnham & Akande, 2004; Furnham, Callahan, & Akande, 2004; Furnham & Mkhize, 2003) and Eastern Europe (Furnham, Rakow, Sarmany-Schiller, & De Fruyt, 1999). The African results led Furnham and Akande (2004) propose that intelligence is perceived as female normative in traditional African cultures. Lastly, gender differences in SEI across cultures are more robust for self than for others (Furnham & Chamorro-Premuzic, 2005).

Thus, clear and consistent cultural differences were found, with Americans providing the highest self-estimates, Asians, i.e. Japanese, Hong Kong Chinese, Singaporeans, Hawaiians, the lowest, and Britons in between (e.g. Furnham & Fong, 2000; Furnham & Fukumoto, 2008; Furnham, Fong, & Martin, 1999; Furnham, Hosoe, & Tang, 2001; Yuen & Furnham, 2005). Equally, participants from Singapore and Egypt held more gender and cultural beliefs than the British and/or the Americans (Furnham, Fong, & Martin, 1999; Furnham & Fong, 2000; Furnham & Mottabu, 2004). In addition, studies in South African countries showed clear cultural differences, with Namibians and isiZulus providing the lowest and Zambians, Indian South Africans, and Nigerians the highest self-estimates of intelligence (cf. Furnham & Akande, 2004; Furnham, Callahan, & Akande, 2003; Furnham, Mkhize, & Mndaweni, 2004). Cultural differences were also observed between Middle Eastern and Sub-Continent Asians, with Lebanese participants providing significantly higher self-estimates than Indian participants (Nasser & Singhal, 2006).

These results were contributed to the diverse cultural norms and values (Furnham, Rakow, & Mak, 2002), such as '*hubris-humility*' (Beloff, 1992; Bond, 1991) or the individualist vs. collectivist culture norms (cf. Hofstede, 1998, 2003; McSweeney, 2002). However, the observed cross-cultural SEI differences are in stark contrast with the psychometric empirical evidence that shows Asian advantage in '*g*' over Europeans/Americans (e.g. Jackson & Rushton, 2006; Lynn, 1997; Lynn et al., 2004).

Similarly, participants from poorer or developing countries provided higher self-estimates of intelligence compared to participants from developed countries. Several reasons for these skewed self-views were suggested, such as lack of exposure to IQ testing during the education process, limited performance feedback, and fiercer academic competition (e.g. Furnham & Baguma, 1999; Furnham, Rakow, Sarmany-Schiller, & De Fruyt, 1999; Nasser & Singhal, 2006).

What's more, the majority of the cross-cultural SEI studies replicated observed gender and generational differences, providing further evidence that these beliefs are pan-cultural and universal (e.g. Furnham & Baguma, 1999; Furnham & Chamorro-Premuzic, 2005; Furnham, Fong, & Martin, 1999; Furnham & Fong, 2000). Likewise, largest gender differences were affirmed to universally occur on mathematical/logical and spatial intelligences (e.g. Furnham & Baguma, 1999; Furnham, Fong, & Martin, 1999; Furnham & Fong, 2000; Furnham & Fukumoto, 2008; Furnham, Hosoe, & Tang, 2001; Furnham & Mottabu, 2004; Furnham, Rakow, & Mak, 2002; Furnham, Shahidi, & Baluch, 2002; Yuen & Furnham, 2005).

Table 1.2.4. summarises main findings in the cross-cultural SEI research.

Table 1.2.4: Results of Cross-Cultural Studies of Self-Estimates of Intelligence

Study	*Country*	*Sample*	*Men*	*Women*	*Difference*
Furnham & Akande (2004)	Namibia	Normal	(n = 54)	(n = 74)	
Mathematical IQ		population	91.90[a,b]	111.70	-19.80
Spatial IQ			95.50[a,b]	109.10	-13.60
	SAR		(n = 53)	(n = 73)	
Mathematical IQ			100.80[a,b]	116.20	-15.40
Spatial IQ			101.40[a,b]	113.30	-11.90
	Zambia		(n = 52)	(n = 37)	
Mathematical IQ			106.00[a,b]	108.90	-2.90
Spatial IQ			112.40[a,b]	118.20	-5.80
	Zimbabwe		(n = 25)	(n = 53)	
Mathematical IQ			100.60[a,b]	121.20	-20.60
Spatial IQ			104.00[a,b]	122.50	-18.50
Furnham & Baguma (1999)	UK	University	(n = 37)	(n = 63)	
Overall IQ	(n = 100)	students	111.24[b]	107.98	3.26
Numerical IQ (M/L+S)[1]			112.21[a,b]	105.19	7.02
Cultural 2 IQ (M +B-K)[2]			106.49[b]	103.48	3.01
	USA		(n =54)	(n = 30)	
Overall IQ	(n = 84)		113.72[b]	113.64	0.08
Mathematical/Spatial IQ			114.85[a,b]	111.91	2.94
Cultural IQ			106.87[b]	108.80	-1.93
	Uganda		(n = 51)	(n =35)	
Overall IQ	(n = 86)		110.12[b]	109.00	1.12
Mathematical/Spatial IQ			111.39 [a,b]	103.64	7.75
Cultural IQ			96.48[b]	101.93	-5.45
Furnham, Callahan, &Akande (2004)	SAR	University	(n = 28)	(n = 70)	

Mathematical IQ	White	students	102.31[c]	99.23	3.08
Spatial IQ			107.89	104.44	3.45
	Black		(n = 39)	(n =44)	
Mathematical IQ			94.87[c]	92.50	2.37
Spatial IQ			104.[61a,b]	101.13	3.48
	Nigeria		(n = 55)	(n = 80)	
Mathematical IQ			100.00[b]	92.20	7.80
Spatial IQ			108.73[a,b]	97.87	10.86
Furnham & Chamorro-Premuzic (2005)	Argentina	University	(n = 83)	(n = 134)	
Overall IQ		Students & normal population	110.40[a]	105.10	5.3
Mathematical IQ			108.30[a,b]	96.70	11.6
Spatial IQ			107.80[a,b]	98.70	9.1
	UK		(n =56)	(n = 129)	
Mathematical IQ			116.40[a,b]	105.40	11.0
Spatial IQ			113.60[a,b]	104.30	9.3
Furnham & Fong (2000)	UK	University	(n =31)	(n = 51)	
Overall IQ Estimated	(n = 84)	students	110.66[b]	108.59	2.07
Overall IQ Actual (RSPM)			54.55[a]	55.64	-1.09
	Singapore		(n = 37)	(n = 51)	
Overall IQ Estimated	(n = 88)		107.81[b]	105.86	1.95
Overall IQ Actual (RSPM)			53.81[a]	56.02	-2.21
Furnham, Fong, & Martin (1999)	UK	University	(n = 94)	(n =133)	
Overall (*g*) IQ	(n = 227)	Students	110.64 [a,b]	108.05	2.59
	Singapore		(n = 37)	(n =51)	
	(n = 88)		107.80 [a,b]	105.85	1.95
	Hawaii		(n =26)	(n = 27)	
	(n = 53)		105.85 [a,b]	103.62	2.23
Furnham & Fukumoto (2008)	Japan	Normal population	(n = 74)	(n = 124)	
Numerical/mathematical IQ	(n = 198)		101.72 [a]	92.40	9.32
Spatial IQ			102.16 [a]	97.08	5.08
Furnham, Hosoe, & Tang (2001)	UK	University	(n = 96)	(n = 133)	
Overall IQ	(n = 229)	Students & parents	110.64[a,b]	108.06	2.58
Numerical IQ			112.71[a,b]	105.94	6.77
Cultural IQ			103.89[b]	102.94	0.95
Overall IQ	USA		(n = 102)	(n = 111)	
Numerical IQ	(n = 213)		112.00[a,b]	110.24	1.76
Cultural IQ			113.51[a,b]	109.01	4.50

Overall IQ Numerical IQ Cultural IQ	Japan (n = 164)		105.93[b] (n = 62) 102.09[a,b] 104.68[a,b] 99.21[b]	105.11 (n = 102) 98.58 96.24 98.67	0.82 3.51 8.44 0.54
Furnham, Mkhize, & Mndaweni (2004) Mathematical IQ Spatial IQ Mathematical IQ Spatial IQ	SAR Indians isi-Zulu	Normal population	(n = 63) 105.61[a,b] 105.28[a,b] (n = 55) 94.72[b] 104.03[a,b]	(n = 41) 118.34 118.56 (n = 46) 99.50 104.54	 -12.73 -13.28 -4.78 -.0.51
Furnham & Mottabu (2004) Overall IQ (Self) Numerical Ability (Cattell) Spatial Ability Mechanical Ability Overall IQ (Self) Numerical Ability (Cattell) Spatial Ability Mechanical Ability	UK Egypt	University Students	(n = 59) 118.69[a] 111.42[a] 114.91[a,b] 108.91[a,b] (n = 54) 113.27[a] 114.42[a] 106.00[a,b] 103.50[a,b]	(n = 92) 110.70 103.09 105.23 96.63 (n = 64) 110.33 102.14 108.47 91.73	 7.99 8.33 9.68 12.28 2.94 12.28 -2.47 11.77
Furnham. Rakow & Mak (2002) Mathematical IQ Spatial IQ	Hong Kong	Parents of School children	(n = 79) 105.43[a] 104.20[a]	(n = 114) 101.61 97.91	 3.82 6.29
Furnham, Shahidi, & Baluch (2002) Mathematical IQ Spatial IQ Mathematical IQ Spatial IQ	UK (n = 212) Iran (n = 154)	University students	(n = 92) 112.43[a,b] 113.01[a,b] (n = 62) 107.04[a,b] 114.40[a,b]	(n = 132) 105.21 106.70 (n = 92) 103.90 112.02	 7.22 6.31 3.14 2.38
Nasser & Singhal (2006) Mathematical IQ Spatial IQ Mathematical IQ Spatial IQ	Lebanon India	University students	(n =401) 115.47[a,b] 104.30[a,b] (n = 142) 104.13[a,b] 108.56[a,b]	(n = 247) 110.78 113.78 (n = 110) 106.19 108.53	 4.69 -9.48 -2.06 0.03

Neto, Furnham, & Paz (2007)	Macao	University	(n = 90)	(n = 107)	
Mathematical IQ		Students & parents	100.33	103.42	-3.09
Spatial IQ			104.62	106.96	-2.34
	Portugal		(n = 139)	(n = 192)	
Mathematical IQ			107.66	99.21	8.45
Spatial IQ			114.25	104.43	9.82
Swami, Furnham, Kannan (2006)	Malaysia	Normal population	(n = 112)	(n = 118)	
Overall IQ			108.99[a]	105.08	3.91
Mathematical IQ			108.67[a]	105.32	3.35
Spatial IQ			107.30[a]	103.96	3.34

Legend: [a] = Indicates significant gender differences in that cell; provided where reported by the author(s). [b] == Indicates significant cultural/nationality differences in that cell; provided where reported by the author(s). [c] == Indicates significant race differences in that cell; provided where reported by the author(s).[1] = M/L = mathematical/logical intelligence, S = spatial intelligence. [2] = M = musical intelligence, B-K = Body-kinaesthetic intelligence.

1.3. Hubris and Humility Effect & Domain-Masculine Intelligence Type: Gender Differences in Self-Estimated Intelligence

Probably the most fascinating finding in the self-estimated intelligence research programme was the fact that males significantly overestimate and females significantly underestimate their abilities (cf. Beloff, 1992; Bennett, 1996, 1997; Betsworth, 1999; Hogan, 1978; Furnham, 1999; Holling & Preckel, 2005; Furnham, 2000; Furnham & Budhani, 2002; Furnham & Bunclark, 2006; Furnham, Crawshaw, & Rawles, 2006; Furnham & Gasson, 1998; Furnham & Rawles, 1995, 1999; Furnham, Reeves & Budhani, 2001; Furnham & Thomas, 2004; Furnham, von Stumm, Makendrayogam, & Chamorro-Premuzic, 2009; Furnham, Wytykowska, & Petrides, 2005; Pallier, 2003; Rammstedt & Rammsayer, 2000a,b, 2001, 2002; Reilly & Mulhern, 1995; Stieger et al., 2010). The phenomenon of male overestimation and female underestimation was named the *'hubris-humility effect'* (HHE) (cf. Beloff, 1992; Furnham, Clark, & Bailey, 1999; Furnham, Fong, & Martin, 1999; Furnham & Ward, 2001; Hogan, 1978).

Evidence shows that intelligence (gender) beliefs influence self-evaluations and can in turn act as self-fulfilling prophecies that directly impact performance and encourage institutionalisation of those beliefs (cf. Ackerman & Wolman, 2007; Beyer & Bowden, 1997; Chamorro-Premuzic & Arteche, 2008; Furnham & Thomas, 2004; Dweck, 1999). The observed gender differences in self-estimated abilities are stable and consistent across cultures (e.g. von Stumm, Chamorro-Premuzic, & Furnham, 2009).

It is unclear whether HHE correctly depicts male and female understanding of their cognitive abilities or whether the inflated and deflated self-perceptions impact one's behaviour and performance. Equally, it remains to be answered whether the female *'humility'* is a reflection of an accurate female self-estimation or whether it is a direct outcome of negative female self-assessments, performance expectancies, stereotypical self- beliefs or low self-confidence. In fact, female self-estimates were shown to be significantly more accurate than were males'. Male self-estimates were significantly inflated compared to their actual psychometric scores (e.g. Rammstedt & Rammsayer, 2002a,b; Reilly & Mulhern, 1995). These findings were further substantiated by Carr et al. (2008) who reported that girls were more accurate in assessing their mathematical skills and knowledge, despite low math ability confidence. Unsurprisingly, boys were overconfident, but their actual mathematical performance was poor.

On the other hand, self-enhancement beliefs were shown to be psychologically beneficial (cf. Kwan et al., 2008). However, self-enhancement bias was found to correlate with low resilience, inferior social skills, poorer GPA and high levels of defensiveness and narcissism (Kwan et al., 2008; Kim et al., 2010, p.396). It is likely that the self-enhancement bias plays a role in the *'hubris'* element of HHE.

While the cause(s) and working mechanisms of HHE remain to be identified, the following causes have been suggested to play a role: diverse child rearing and socialisation practices (Beloff, 1992), social and gender-role normative stereotyping and self-stereotyping (Guimond et al., 2006), self-enhancement and self-derogatory evaluation biases (Beyer, 1990, 1998, 1999; Furnham, 2001; Kwan et al., 2008), lack of confidence and/or overconfidence (Sleeper & Nigro, 1987), gender differences in self-concept and inaccurate self-estimates (Pallier, 2003; Roberts, 1991), personality traits and male superiority in certain areas of cognition (Chamorro-Premuzic & Furnham, 2005; Furnham & Rawles, 1995; Hyde et al., 1990a,b; Lynn et al., 2002; Maccoby & Jacklin, 1974; Voyer et al., 1995).

Thus, it remains to be answered what role the aforementioned causes play in HHE. Equally important is to ascertain on which multiple intelligences HHE is most pronounced, and why. The next sections will address these possible causes in greater detail.

1.3.1. Domain-Masculine Intelligence Type (DMIQ)

Studies that used the multiple self-assessed intelligences model (e.g. Furnham, Clark & Bailey, 1999; Furnham & Gasson, 1998; Furnham, 2000; Furnham & Bunclark, 2006; Furnham & Mkhize, 2003; Rammstedt & Rammsayer, 2002a) found that gender differences were strongest on the mathematical/logical and spatial intelligences, followed by overall and verbal intelligences, with males significantly overestimating and females significantly underestimating their abilities (see Tables 1.2.3. and 1.2.4. for an overview).

A meta-analytical study investigating the magnitude of gender differences in mathematical/logical, spatial, overall and verbal self-assessed intelligences (Szymanowicz, Chamorro-Premuzic, & Furnham, 2011, unpublished manuscript), found that the biggest weighted mean effect sizes were for mathematical/logical, (d = .44), followed by spatial (d = .43), overall (d =.37) and verbal (d =.07) intelligence, with males providing higher estimates in all but verbal intelligence.

Unsurprisingly, mathematical, spatial and verbal intelligences were the best predictors of self-estimated overall intelligence as demonstrated through numerous multiple regression analyses (e.g. Furnham, 2001). This finding led Furnham (2000) to conclude that gender differences in self-estimates of intelligence reflect laymen's view of intelligence, i.e. an amalgamation of overall, mathematical and spatial intelligences.

Thus, individuals anchor their strongest cognitive ability (Kahneman & Tversky, 2000) and use it when evaluating their own and others' intelligence (Furnham & Bunclark, 2006). Hence, males draw on mathematical/logical and spatial intelligences as their '*point fort*' while females rely on verbal and personal/emotional intelligences (Furnham & Petrides, 2004; Petrides, Furnham, & Martin, 2004).

Furnham (2000) proposed that people view intelligence as '*male-normative*', since mathematical/logical and spatial intelligences are areas where males are believed to excel. This particular claim is explored in this thesis with the introduction of the '*domain-masculine intelligence type*' (DMIQ), a composite of mathematical/logical and spatial intelligences. Accordingly, the investigation of the relationship between DMIQ and HHE and DMIQ's role in the prediction of HHE as well as the confirmation of DMIQ as the most sensitive predictor of gender differences in SEI and identification of HHE determinants are central to this thesis.

1.3.2. Gender Differences in Mathematics Achievement, Attitudes and Affect

In order to explain gender differences in math achievement, attitudes and affect, as well as the female underrepresentation in science, technology, mathematics, and engineering (further as STEM), researchers proposed a number of socio-cultural, biological, attitudinal, and environmental influences that could underlie gender differences in math achievement, attitudes and affect (Else-Quest, Hyde, & Linn, 2010; Halpern et al., 2007).

The two principal socio-cultural models that attempt to explain the gender gap in math achievement, affect and attitudes are 1) the *gender stratification hypothesis* (Baker & Jones, 1993), and 2) *gender expectancy-value model* (Eccles, 1994).

The first model argues that girls have less opportunity to achieve in math, develop more negative affect towards math and perform less well because they do not perceive math as useful. According to the later model, individuals take on challenges when they value the task and because they believe they can be successful (Eccles, 1994; Else-Quest et al., 2010). Whilst gender and cultural stereotypes, parental, peer and teacher attitudes, and individual's goals give task its perceived value, relevant past experiences, degree of task difficulty, and ability self-concepts shape success expectations (Eccles, 1994). It follows that if males and females have different success expectations of math and science tasks, gender differences will occur (Halpern et al., 2007).

Hyde and colleagues (1990a,b) meta-analysed gender differences in math attitudes and affect, and found that females held more negative attitudes towards math, had lower self-confidence and stereotyped math as domain-masculine, with comparable effect sizes for math attitudes and math affect (*d* =-.90). Moreover, meta-analytical studies of gender differences in math achievement revealed decreased gender gap (Linn & Hyde, 1989; Hyde, Fennema, & Lamon, 1990a).

Thus, distinct mechanisms govern math attitudes and achievement. In fact, attitudinal causes such as math self-confidence, gender stereotypes, and perceived relevance and enjoyment of math, influence math achievement and participation in (advanced) math courses (Crombie et al., 2005; Hyde, Fennema, & Lamon, 1990a; Meelissen & Luyten, 2008, p. 83). For girls, self-confidence and 'liking' of math were significant determinants of math achievement and involvement in math and sciences (Meelissen & Luyten, 2008). But, females were most self-confident about their verbal and language abilities (Meece et al., 2006; van der Sluis et al., 2010), whilst self-confidence in boys was positively correlated with domain-masculine perception of mathematics (Meelissen & Luyten, 2008). Boys also valued math and sciences more than girls and attributed success in these domains to ability, while girls contributed math and science success to effort and hard work. (Meece et al., 2006). These findings uphold typical gender role stereotypical beliefs.

1.3.2.1. Gender Gap in Education and Examination

Females earn higher grades than males in all major subjects throughout elementary school. This pattern changes when girls and boys enter high school and college (e.g. Hyde et al., 1990a,b; Kessel & Linn, 1996). Some recent data suggests that boys outperform girls in math and close the reading gap before they reach ten years (Husain & Millimet, 2009). Evidence also suggests that in high school, females get better math grades, but boys take more (advanced) math courses and outperform girls on high ability math tests (cf. Crombie et al., 2005; Gallagher & De Lisi, 1994; Hyde, Fennema, & Lamon, 1990a; The College Board, 1998). These gender differences increase with age (e.g. Hyde, Fennema, Lamon, 1990a).

A meta-analysis of 493,495 14-16 year olds across 69 nations revealed small effect sizes (d <.15) in gender differences in mathematics achievement, although cross-nationally the effect sizes varied significantly (-.42 < ds <.40) (Else-Quest, Hyde, & Linn, 2010). While only small differences were observed in math achievement, boys were significantly more confident, less anxious, more extrinsically motivated, and had higher math self-concept and self-efficacy than girls (.10 < ds < .33; national -.61 < ds < .89), which is in line with previous research (e.g. Hyde, Fennema, Ryan, et al., 1990b). The cross-national gender gap in math achievement was moderated by socio-cultural factors in each nation, such as parity in math course enrolment and female public life presence (Else-Quest et al., 2010). Bedard and Cho (2010) reported similar results of early male advantage in math and sciences in almost all developed OECD countries.

Data also suggests that the course taking pattern might be partially responsible for the observed math gender gap in college (e.g. Ayalon, 2003; Crombie et al., 2005). The fact that high school males enrol in more advanced math classes positively correlates with their future college GPA and leads to more science and math careers compared to females (The College Board, 1998). Thus, taking mathematics courses in high school narrows the college gender gap and increases STEM career applications (Ayalon, 2003).

In addition, boys outperform girls by 0.4 SDs on standardised tests such as SAT-M that are vital for admission to higher education (e.g. Benbow et al., 2000; Cole, 1997; Feingold, 1988, 1996; Halpern et. al., 2007; Hyde, Fennema, & Lamon, 1990a; Kessel & Linn, 1996; Maccoby & Jacklin, 1974; The College Board, 1998). This asymmetry contrasts with the female grade advantage. Hence, while tests measure specific skills at a particular point in time, grades denote a broad range of skills, such as motivation, achievement, scholastic attendance, and participation.

It was suggested that girls' reliance on conservative and classroom-taught problem solving strategies hurts their standardized math test performance, where more abstract and unconventional techniques are necessary for successful performance (Gallagher & De Lisi, 1994; Stumf & Stanley, 1996). Additionally, Kessel and Linn (1996) proposed that low self-confidence, heightened susceptibility to stereotypical beliefs and inability to cope with SAT's speed cause female math underperformance.

Yet in 2010, for the first time, American women earned more doctorates than did men (de Vise, 2010). American women also earned the majority of undergraduate and graduate degrees in the last decade (e.g. Spelke, 2005; Xie & Shauman, 2003). In the U.S. women now earn 70% of doctorates in health science, 67% in education and 60% in social and behavioural sciences (de Vise, 2010). However, 80% of doctorates in engineering and the majority of math and physical science doctorates go to men.

Whilst the diminishing doctorate gender gap is seen by some as confirmation of the changed higher educational status of women, gender differences in occupational preferences remain unchanged. In fact, gender differences in career goals and occupational choices emerge before primary school, with boys showing more interest in science careers than girls (Weisgram & Bigler, 2006). Those differences are affirmed in adolescence, when males and females make different life and career choices (Sax & Harper, 2007). Thus, it is likely that educational and career choices result from distinct gender-role socialization and reflect dissimilar life goals of males and females (Eccles, 1987). Unsurprisingly, Linver and Davis-Kean's (2005) findings revealed that gender differences in self-concepts and not grades, were the strongest determinants of gender differences in vocational choices. Moreover, decline in female math self-confidence was responsible for avoidance of STEM careers and led to more college major changes (Sax, 1994; Sax & Harper, 2007). Yet, Kawakami, Steele, et al. (2008) reported that women who disliked or feared math could be successfully trained to engage in math; thus, the math gender gap could be further reduced with appropriate intervention techniques.

1.3.2.2. Parental Attitudes, Expectations and Stereotypical Beliefs

Parental expectations, beliefs and stereotypical biases have been shown to influence child-rearing and socialising practices as well as impact children's educational achievement, career choices and self-perceptions (Beyer, 1990, 1998, 1999; Bleeker & Jacobs, 2004; Jacobs & Eccles, 1992; Lytton & Romney, 1991). Interestingly, a meta-analytical study of gender differences in child-rearing practices revealed that despite similarities in raising boys and girls, fathers significantly encourage sex stereotyped behaviours in boys (Lytton & Romney, 1991)

Data also suggests that parental sex-stereotypical views and expectations of children's abilities and academic accomplishment influence children's ability self-concept and predict performance (e.g. Bleeker & Jacobs, 2004; Halpern et al., 2007; Jacobs & Eccles, 1992; Linver & Davis-Kean, 2005; Lytton & Romney, 1991; Maccoby & Jacklin, 1974). Equally, parental level of education and the degree of involvement with their children's education were powerful determinants of children's math achievement (Halpern et al., 2007; Roznowski et al., 2000).

Jacobs and Eccles (1992) reported that parents, who believed in male math advantage thought their sons to be significantly better in math than their daughters. In particular, mothers' stereotypical beliefs biased their children's perceptions and had a bigger impact on their ability than actual grades. Bleeker and Jacobs (2004) have shown in their longitudinal study that mothers' childhood gender stereotypes influenced adolescents' self-perceptions of math ability and significantly predicted career choices. Mothers' expectations were the most powerful predictor of academic performance through high school, responsible for math choice in girls and outstanding grades in boys (Linver & Davis-Kean, 2005). Equally, teachers' student ability perceptions predicted actual test scores at a later stage. Thus, parents, teachers as well as peers, are powerful forces in shaping children's ability self-concepts and academic career choices (Halpern et al., 2007).

1.3.3. Gender Differences in Cognitive Biases, Self-Perceptions, and Self-Beliefs of Ability and Performance

Research in mainstream gender differences is founded on the premise of male behavioural normativeness and female behavioural 'deviancy' (Roberts, 1991). Equally, the behavioural consequences of self-perceptions are most paramount in regards to gender. The majority of studies investigating gender differences in biases, self-beliefs, stereotypes, over-confidence, and over- and underachievement (see also sections 1.2.5, 1.2.5.1, 1.3.4 and 1.3.5), concluded that individuals are not very effective in judging their own performance (cf. Burson et al., 2006; Dunning et al., 1989; Epley & Dunning, 2000; Guenther & Alicke, 2010; Harrison & Shaffer, 1994; Kruger & Dunning, 1999, 2002; Krueger & Mueller, 2002; Moore & Small, 2007). These findings are grounded in Festinger's social comparison theory (1954) and self-categorisation theory (Turner, 1999), with individuals lacking accurate insight in their skills and responses.

Ehrlinger and Dunning (2003) established that women held more negative and self-handicapping beliefs about their scientific ability than men, which led to lower performance estimates. No significant gender differences in the actual performance were observed. Beyer (1990, 1998) reported similar results, with females significantly underestimating their performance, being less confident about their answers, and more prone to negative self-perception biases than males. However, this only held true when the task was '*masculine*'. Similarly, Betsworth's (1999) study with well-educated women found that women significantly underestimated their performance on six out of nine abilities on the General Aptitude Test and had lower career expectations. Comparable results were obtained with stereotyped groups (e.g. Steele & Aronson, 1995).

These results imply that low pre-task expectancies on gender-typed tasks result in female performance underestimation. Equally, good performance does not automatically lead to positive self-evaluation. Societal and parental pressures, gender-role stereotypical beliefs and learned helplessness (Diener & Dweck, 1980) were proposed to cause the female under-performance on masculine tasks (Beyer, 1990, 1999). These findings imply that self-beliefs are stronger determinants of future behaviour than objective feedback (Critcher & Dunning, 2009). Thus, a talented female student who assumes that her math assignment is below par will provide humbler self-estimates of ability than an average male student who firmly believes in his skills. It seems likely that similar mechanisms may play role in HHE.

Men and women also react differently to behavioural and performance feedback. Women are more influenced by the emotive meaning of evaluative feedback than men (Roberts, 1991). In fact, men tend to be more self-confident about their performance and dismiss the 'unfitting' elements of feedback. Women, on the other hand, regard evaluative feedback as an opportunity to gain more self-insight (Roberts, 1991). However, both genders react equally during negative feedback situations. Kim et al. (2010) demonstrated that inaccurate feedback leads to self-handicapping that affects actual performance. Yet, accurate self-assessors were less likely to self-handicap and did better academically than did self-enhancers and self-effacers. In fact, Kim, Chiu, and Zou (2010) found that self-enhancers inflated their performance in order to uphold positive self-belief, whist self-effacers were driven by self-doubt about their performance. It is likely that several of the above-mentioned mechanisms play a role in HHE.

1.3.4. Role of Self-confidence and Overconfidence

Self-confidence, overconfidence, overconfidence bias as well as *'hubris'*, fall within the interest areas of positive psychologists. Self-confidence refers to beliefs in one's abilities to perform and succeed that are acquired through previous experiences. A subtle difference exists between hubris and overconfidence. Whilst hubris refers to undeserved confidence, overconfidence refers to disproportionate belief in achievement. Overconfidence bias occurs when subjective confidence in one's abilities or judgement is significantly greater that the actual performance, e.g. providing 90% certainty rating while being wrong 35% of time.

Overconfidence has been proposed to be responsible for the tendency to overestimate one's abilities, which seems to especially occur in male populations (Burson, Larrick, & Klayman, 2006; Carr et al., 2008; Dunning et al., 1990; Moore & Small, 2007; Pallier, 2003). In general, individuals are most overconfident on difficult tasks (Dunning et al., 1990; Jonsson & Allwood, 2003; Lichtenstein & Fischhoff, 1977). Carr, Hettinger-Steiner, Kyser, and Biddlecom, (2008) found that boys were overconfident but inaccurate about their math skills, whilst girls were accurate but reported lower math confidence. Equally, overconfident estimations are made because people fail to sufficiently analyse potential uncertainties in their assessments (Griffin et al., 1990). Thus, the *'confidence bias'* resembles Downing, Dunning-Kruger's, Better-Than-Average (BTAE) and Worse-Than-Average (WTAE) effects.

Self-confidence was shown as one of the key predictors of gender differences in achievement, with females reporting lower self-confidence than males, despite no gender differences in actual performance. In fact, females outperformed males on the actual task (cf. Eccles-Parsons, Adler & Meece, 1984; Pallier, 2003; Sleeper & Nigro, 1987). Based on the above, it seems plausible that self-confidence also plays a role in HHE.

1.3.5. Role of Stereotypes

Stereotypes are commonly held persistent beliefs about particular social groups or individuals that are based on assumptions or past experiences. They affect the way we act, feel, and think about others and are usually triggered by unconscious cognitive processes. Commonly held stereotypes are presumed to be persistent, stable over time, and difficult to change. Nevertheless, Garcia-Marques, Santos, and Mackie (2006) provided evidence that stereotypes change since they are contextually dependent. Furthermore, positive and negative self-stereotypes have been found to affect one's performance, with negative self-stereotypes resulting in more accurate and risk-averse performance and positive self-stereotypes in enhanced creativity and speedy response (Seibt & Forster, 2004).

Contrary to popular belief, individuals are rather accurate in gender stereotypes (cf. Diekman, Eagly, & Kulesa, 2002; Hall & Carter, 1999; Swim, 1994). Some researchers asserted that gender stereotypes are responses to social and biological gender roles (e.g. Biernat, 1991; Davies & Shackelford, 2006; Halpern et al., 2007; Hoffman & Hurst, 1990; Jacklin, 1989). Hoffman & Hurst (1990) proposed that gender stereotypes are influenced by the different gender role preferences, i.e. males are *'agentic'*, e.g. hunter-gatherer, and females are *'communal'*, e.g. mother/carer. This assertion is in line with Maccoby and Jacklin (1974) who found male preference for Things and female inclination for People/Relations.

Stereotypes also seem stable across cultures but cross-cultural differences exist in the degree men and women use gender-group stereotypes to define themselves, with the most pronounced gender differences observed in Western countries, e.g. USA and Europe (Guimond et al., 2007).

Gender stereotypes seem to be strongest in areas that are stereotypically associated with masculinity and femininity, such as mathematics, sciences, and arts. Brown and Josephs (1999) reported that women who believed math tests would reveal their inferior math skills, performed significantly worse, than did women who thought the assessment would authenticate their robust math skills. Interestingly, opposite results were found for males, with worst performance observed on tests that were supposed to confirm their math superiority. These findings led to the conclusion that math ability gender stereotypes induce different performance anxiety in men and women. Bonnot and Croizet (2007) reported that low math ability self-concept and math test underperformance in female participants resulted from female math inferiority stereotype. Finally, Kray et al. (2001) in their study of gender stereotypes in negotiations found that when women were told a task will reveal their intrinsic abilities, their performance declined. The same instructions caused improvement in male performance.

The self-stereotyping process might be responsible for priming social, racial and gender role stereotypical behaviours (cf. Ambady, Shih, Kim, & Pittinsky, 2001; Brown & Josephs, 1999; Chatard et al., 2007; Kiefer & Sekaquaptewa, 2007; Nosek, Banaji, Greenwald, 2002; Steele & Ambady, 2006; Wheeler & Petty, 2001). Although females consciously reject math gender stereotypes (Hyde, Fennema, Ryan, et al., 1990b), their math attitudes are negatively impacted by their own implicit stereotypical math beliefs (Nosek, Banaji, & Greenwald, 2002). Likewise, priming a social category, e.g. blondes, can evoke stereotype-consistent behaviours and increase stereotype susceptibility (cf. Shih et al., 1999; Wheeler & Petty, 2001).

Unsurprisingly, women who were primed with category *'female'* demonstrated significantly more stereotypical attitudes towards math and arts than did women who were not primed (Kiefer & Sekaquaptewa, 2007; Steele & Ambady, 2006). Implicit math gender stereotypes also negatively impacted math perceptions and performance in females (Dar-Nimrod, 2007; Kiefer & Sekaquaptewa, 2007). In fact, Dar-Nimrod and Heine (2006) established that women who were primed with beliefs that the female math underperformance is genetically encoded, had significantly worse performance than did women who were primed with other beliefs. Accordingly, Chatard et al. (2007) found that when gender stereotypes were activated prior to a stereotype-reinforcing task (e.g. school math grade recollection), more stereotypical answers were produced. Also, when females were stereotypically primed after the task, their statements confirmed male math superiority and underestimation of own math grades. The same results were found for males and the arts.

1.3.5.1. Stereotype Threat

What role do stereotypes play in HHE? Is it possible that men and women experience the self-estimation situation, i.e. the experiment setting, the questions, the fellow participants and the experimenter(s) differently? If so, are their respective gender-roles affected differently?

A possible explanation has been proposed by Steele and Aronson (1995), the so-called *'stereotype threat'*. Stereotype threat (ST) is a widely researched and documented phenomenon. It implies a situational threat to any member of a group about whom negative stereotypes exist, such as the elderly, single parents, or football fans. The threat arises when identification with a negative stereotype that exists about a particular group becomes pertinent to the individual (Steele, 1997). This is most likely to occur when an individual is in a situation that evokes that stereotype (Steele, 1997). Stereotype threat was shown to have a direct impact on performance, with most evidence about women's underperformance in math, academic underperformance of African Americans, and reduced working memory in Latinos and women (cf. Aronson et al., 1999; Brown & Pinel, 2003; Good, Aronson, &

Harder, 2008; Nosek, Banaji, & Greenwald, 2002; Schmader & Johns, 2003; Steele, 1997; Steele & Aronson, 1995; Wheeler et al., 2001). Evidence also shows that stereotyped individuals are likely to fall victim of stereotype threat as early as elementary school, with susceptibility peaking in early adulthood. As such ST was shown to impact academic, career and life choices of stereotyped individuals (Good et al., 2008).

It also appears that the degree to which ST is experienced depends on the extent of an individual's identification with the stereotype. Evidence shows that the threat has the biggest impact on those who are more confident about their abilities and those who have not yet started to doubt their abilities because the threat situation evokes fear of stereotype confirmation about their particular group (e.g. Brown & Pinel, 2003; Good, Aronson, & Harder, 2008; Schmader & Johns, 2003; Steele, 1997). Moreover, stereotype threat has a lasting impact on those affected because it affects their coping mechanisms and reduces self-control (Inzlicht & Kang, 2010).

Women's susceptibility to negative stereotypes often results in underperformance and adversely affects learning (Rydell et al., 2010; Steele & Aronson, 1995). Good, Aronson, and Harder's (2008) study with high-level math students confirmed that ST suppresses math test performance even in highly talented women. A simple intervention – gender role neutralising task instructions – was sufficient to eliminate ST and improve performance by female participants (Good et al., 2008).

Despite being one of the gloomiest discoveries by social psychologists, evidence now exists that attitudes and self-stereotypical beliefs evoked by ST can be changed with appropriate training. As demonstrated in a study by Forbes and Schmader (2010), women who were successfully retrained to have a positive attitude to math, also reported increased motivation towards math (see also Kawakami, Steele et al., 2008). Men were irresponsive to these manipulations, presumably because of their math superiority beliefs. This finding implies that with the right intervention, it might be possible to alter the underestimating ability pattern that occurs in the hubris-humility effect.

1.3.6. Role of Gender Identity Roles

Gender role refers to a set of generally accepted and expected behavioural traits for males and females that comply with social norms and vary across cultures (Arrindel et al., 2003; Wilcox & Francis, 1997). Historically, gender role constructs were assumed to be the opposites of a single dimension. Bem (1974, 1981b) has changed this view with her gender schema theory that postulates that *'masculinity'* (M) and *'femininity'* (F) are independent from each other and used by individuals to organise their life in terms of *'masculine'* and *'feminine'*. As such, cognitive concepts are believed to be culturally determined and not as previously thought, behavioural traits (Lippa, 2001). The Bem Sex Role Inventory (BSRI((Bem, 1974, 1981a) categorises individuals as masculine, feminine, androgynous, i.e. being both masculine and feminine, and undifferentiated, i.e. no clear gender role preference. Lack of supporting data for the *'androgyny'* concept led Bem to shift away and focus on M and F (Lippa, 2001).

Nonetheless, laymen's understanding of the masculinity and femininity concepts is broader. It incorporates personality traits, social roles, occupations, interests, physical appearance, and sexual preferences (Lippa, 2001). In fact, masculinity was shown to positively correlate with Extraversion and

Conscientiousness and negatively with Neuroticism and Agreeableness, whilst femininity correlated positively with Agreeableness, Conscientiousness, and Neuroticism (Marusic & Bratko, 1998). Francis and Wilcox's (1998) study that used Eysenck's personality dimensions reported a relationship between masculinity, high Extraversion and low Neuroticism. Femininity correlated with high Neuroticism and lie scale scores and low psychoticism.

Is there a relationship between gender stereotypes and gender identity variables? Evidence shows that pre-school children correctly use the feminine gender labels to help them make sense of the world. With age, children come to rely on own cognitive and behavioural references and their gender schemas become more differentiated (Biernat, 1991). Adults appear to associate masculinity with *'agentic'* and femininity with *'relational, communal'* gender stereotypical roles that were previously shown to influence gender stereotypical beliefs and behaviour (Hirschy & Morris, 2002).

Finally, Rudman and Phelan (2010) demonstrated that priming women with traditional (e.g. male pilot and female flight-attendant) as well as non-traditional (e.g. female pilot and male flight-attendant) gender roles increases gender stereotyping and decreases interest in masculine occupations. Similar to stereotype threat, gender role beliefs influenced performance on a spatial ability test. Massa, Mayer, and Bohon (2005) demonstrated that female performance depended on provided instructions, with masculine women performing better with spatial instructions and feminine women with empathy instructions. Interestingly, these results were not replicated with male participants, suggesting that females are more susceptible to societal and behavioural stereotypical beliefs and expectations. These findings represent additional insight for the working of HHE.

1.3.7. Role of Self-Concept and Self-Construct

Self-concept is a multidimensional concept that explains how individuals perceive the impact of their behaviour on their environments as well as how they are perceived by others (Marsh, 1990). Thus, self-perceptions are key to self-concept. Self-concept is a more generic term that incorporates more specific concepts of self-esteem (Rosenberg, 1965), self-efficacy (Bandura, 1997), and self-control/discipline (Duckworth & Seligman, 2005, 2006; Tangney et al., 2004); often the term self-construct is used as a higher order core construct for the individual concepts (Judge et al., 2002).

Self-concept was shown to be an important predictor of one’s ability and performance (cf. Eccles, 1987; Halpern et al., 2007; Marsh, 1990). Gender differences in self-concept appear to follow stereotypical patterns, with men having higher self-perceptions of math, problem-solving, emotional stability and physical abilities and women higher verbal, social, moral and artistic self-perceptions (Vispoel & Forte Fast, 2000, p.92). Equally, gender differences in self-concept of ability are better predictors of individual’s career choices than is actual performance (Eccles et al., 1984; Eccles, 1987).

The current gender difference research programmes are focused on differences in male and female self-concept (e.g. Ehrlinger & Dunning, 2003; Maddux & Brewer, 2005) and little attention is paid to the causes of gender differences in self-construals (Guimond et al., 2006). Yet, gender differences in self-construals (Cross & Madson, 1997), i.e. males believing themselves as independent and females as interdependent, were proposed as causes of many sex differences in social behaviour. Indeed, Guimond, Martinot, Chatard, Crisp, and Redersdorff (2006) concluded in their extensive study of gender differences in self-construals that the female self-construal is relationally orientated, while the male self-construal is self-directed.

1.3.7.1. Self-Control and Self-Discipline

Self-control or self-discipline is the ability to employ one's willpower over other desires and motives in order to achieve a certain goal or exhibit certain behaviour. Duckworth and Seligman (2005) reported that high self-discipline or self-control leads to enhanced academic performance, better final grades and school attendance. Self-discipline has also been shown to be gender sensitive, with girls' higher self-discipline responsible for better grades and scholastic test results (Duckworth & Seligman, 2005, 2006). These findings are further supported by Csikszentmihalyi et al. (1997) who argue that self-discipline, hard-work, and dedication are key to superior academic achievement, more so than high intelligence.

1.3.7.2. Self-Esteem

A vast psychological literature exists on (global and social) self-esteem which refers to individual's assessment of own worth, self-liking, self-respect, and self-competence. Self-esteem differs from self-confidence and self-efficacy since it does not specifically focus on one's ability or future performance.

Historically, self-esteem research focused on relationship with well-being, with low self-esteem associated with depression and high self-esteem with life satisfaction and happiness (e.g. Hirschy & Morris, 2002).

High self-esteem individuals have been shown to associate success with internal causes and failure with external causes compared to low self-esteem individuals. Unsurprisingly, child-rearing practices seem directly responsible for high or low self-esteem, with hostile mothering leading to low self-esteem and maternal approval and emotional responsiveness to high self-esteem (Keltikangas-Jaarvinen et al., 2003).

Male attributional behaviour appears to mirror high self-esteem individuals, whilst females tend to exhibit a self-derogatory attributional style that resembles that of low self-esteem individuals (Hirschy & Morris, 2002; Petiprin & Johnson, 1991). An extensive meta-analysis of gender differences in self-esteem (N = 97,121) revealed a small effect size favouring men (d =.21) with peak in late adolescence (d =.33) (Kling et al., 1999). Data in support of the male advantage in self-esteem show that (global) self-esteem correlates positively with (cultural) masculinity but not femininity (Hirschy & Morris, 2002; Schmitt & Allik, 2005).

Moreover, self-esteem is a cross-cultural phenomenon that has been shown to correlate negatively with Neuroticism and positively with Extraversion (Schmitt & Allik, 2005; Pullman & Allik, 2000). Thus, it seems plausible that self-esteem acts as a self-defence mechanism against negative affect and emotional instability.

1.3.8. Role of Affect

Affect, or the experience of emotion, is usually assessed through self-rated mood assessment. A two-dimensional model of affect - positive and negative – has been affirmed universally (Watson & Tellegen, 1985; Watson et al., 1988; Watson et al., 1999). Positive affect (PA) describes the degree to which an individual feels active and enthusiastic, with high PA associated with energy and involvement and low PA with unhappiness and sluggishness. Negative affect (NA) describes a degree of general upset, with high NA associated with anger, fear, guilt, loathing and anxiety and low NA with composure and calmness (Watson et al., 1988).

Predictably, affect correlates with anxiety and depression, with low PA positively correlated with depression and NA correlated with anxiety (Watson & Tellegen, 1985). Equally, NA is strongly correlated with Neuroticism and PA strongly with Extroversion (Costa & McCrae, 1980). This confirms the hypothesis that positive and negative affect correspond with Extraversion and anxiety/Neuroticism (Watson et al., 1988). These findings were further given support by research into gender difference in affect, with women scoring higher on negative affect than men (Smith & Reise, 1998).

1.4. Aims of Thesis

The present research has the following objectives. The first main objective is to corroborate the existence of the '*hubris-humility effect*' (HHE), or gender differences in self-estimation of intelligence, on the numerical/logical-spatial factor of the SEI model (e.g. Beloff, 1992; Bennet, 1997; Bond, 1991; Furnham, 2001; Furnham & Baguma, 1999; Holling & Preckel, 2005; Pallier, 2003; Rammstedt & Rammsayer, 2002a,b), i.e. on the '*domain-masculine intelligence type'* (DMIQ). DMIQ is a novel term introduced in this thesis.

HHE comprises of an over-estimation of ability or *'hubris'* that is observed in males and an under-estimation of ability or *'humility'* observed in females (Furnham, 2001; von Stumm et al., 2009). HHE has been shown to be most profound on the mathematical/logical and spatial factor of the SEI model (e.g. Furnham & Fukumoto, 2008; Swami et al., 2006; Yuen & Furnham, 2005) and has been confirmed to exist across cultures and geographies (e.g. Furnham & Chamorro-Premuzic, 2005; Furnham & Fong, 2000; Furnham, Hosoe, & Tang, 2001; Furnham & Mottabu, 2004; Furnham, von Stumm, et al., 2009).

The second main objective of this thesis is to validate the fact that gender is the best determinant of DMIQ, over and above a number of potential determinants of gender differences in SEI. To that end, a number of possible determinants are introduced that have been shown (e.g. Beyer, 1998, 1999; Chamorro-Premuzic and Arteche, 2008; Chamorro-Premuzic, Furnham, Moutafi, 2004; Duckworth & Seligman, 2005, 2006; Ehrlinger & Dunning, 2003; Gottfredson, 2000; Guimond et al., 2006 ; Halpern et al., 2007; Hirchy & Morris, 2002; Kwan et al., 2008; Lippa, 2001; Petiprin & Johsnon, 1991) or are expected to play a role in the intelligence type, based on literature in the field (e.g. Ackerman & Wolman, 2007; Ambady et al., 2001; Carr et al., 2008; Dar-Nimrod, 2007; Dunning et al., 1990; Feingold, 1988, 1996; Nosek, Banaji, & Greenwald, 2002; Pallier, 2003; Sleeper & Nigro, 1982; Steele & Aronson, 1995; Watson & Tellegen, 1985), such as general intelligence ('*g*'), gender identity variables, self-construct, affect measures and the experimental task containing psychometric problems and task confidence probes. Within this context, the role of age in the prediction of the intelligence type will also be further examined, given that age has been previously shown to moderate self-estimates of intelligence and fluid intelligence (cf. Beier & Ackerman, 2001, 2003; Rammstedt & Rammsayer, 2002b).

Consequently, to further corroborate the occurrence of HHE in DMIQ and the role of gender in the prediction of DMIQ, two studies will be conducted with three distinct cultures (Hofstede, 1998, 2003), Czech Republic, Colombia and the United Kingdom. Despite the extensive cross-cultural research of gender differences in SEI, no previous studies were conducted in the Czech Republic and Colombia. Likewise, the two main objectives will be tested with a precocious population in order to ascertain whether precocity, awareness about intelligence as well as beliefs about intelligence impact the occurrence of HHE in DMIQ and the role of gender in the prediction of the type. The first ten correlational studies will be reported in Chapters 2 to 6, which constitute the first part of this thesis. The to-be-tested hypotheses will be formulated and presented in each chapter.

The second part of this thesis will comprise of five experimental design studies that are reported in Chapter 7. To date, no experimental studies have been conducted as part of the SEI research programme. The experimental studies will introduce repeated measurement of DMIQ and a specially designed psychometric task (TCAP) that will also include task-success probes (TSP).

The objective of the second part of the thesis is to ascertain whether the repeated measurement and the task will 1) impact the occurrence of HHE on DMIQ, 2) facilitate size reduction in HHE from the initial task (T1) to the post-task (T2) estimation condition, 3) assist explanation of DMIQ's best predictor, 4) enable an in-depth investigation of gender's role in the relationships between DMIQ and TCAP and DMIQ and TSP, and 5) facilitate understanding of the role gender plays in TCAP and TSP.

Repeated measurement of DMIQ estimation will be used to examine whether HHE can be manipulated or reduced following a gender-stereotype inducing task, i.e. TCAP (Bartsch & Nesselroade, 1973). TCAP will be included to determine whether individuals are incapable of accurate self-assessments of ability or performance (Burson et al., 2006; Guenther & Alicke, 2010; Moore & Small, 2007) or whether the provided post-task DMIQ estimates will be accurate (e.g. Ackerman et al., 2002; Chamorro-Premuzic et al., 2010; Hall & Carter, 1999; Swim, 1994) as well as investigate literature claims of male advantage in mathematics achievement, attitudes and affect (cf. Crombie et al., 2005; Beyer, 1998; Hyde et al., 1990a,b; Halpern et al., 2007; Meelissen & Luyten, 2008; Sax & Harper, 2007). Equally, the reported male over self-confidence, in particular in math achievement and domain-masculine abilities, (Carr et al., 2008; Meece et al., 2006; Meelissen & Luyten, 2008; van der Sluis et al., 2010) will be tested through the task-success probability probes. However, validation of sex differences in psychometrically assessed intelligence is not an objective of this thesis.

DMIQ will be measured as a combination of mathematical/logical and spatial self-estimated intelligences from the SEI model (Furnham & Gasson, 1998; Gardner, 1983). Potential determinants will be assessed using valid and generally utilised measures of fluid and crystallised intelligence (Baddeley, 1968; Bryon, 2006; Lynn, Irwing, & Cammock, 2002; Wonderlic, 1992) masculinity and femininity (Bem, 1981a), self-esteem (Rosenberg, 1965), self-control (Tangney et al., 2004), and positive and negative affect (Watson et al., 1988). The psychometric task comprises of numerical, reasoning, spatial and crystallised intelligence items that will vary per experimental study (Bryon, 2006; Irwing, Cammock, & Lynn, 2001; University of Kent, 2007) to investigate its impact on the results. The number of task-success estimation probes that are incorporated in the psychometric task, will also be varied per study, in order to determine whether it will impact the results. As in part one, specifically formulated hypotheses will be tested and reported in each study.

A summary of the findings and conclusions of the thesis as well as limitations and ideas for future research will be given in chapter 8.

Chapter 2: Hubris-Humility Effect and Domain-Masculine Intelligence Type: Gender Differences in Self-Estimated Intelligences

2.1. General Introduction

Considerable previous research (e.g. Beloff, 1992; Furnham, 2000, 2004; Furnham & Bunclark, 2006; Furnham, Clark, & Bailey, 1999; Furnham & Rawles, 1995; Furnham & Thomas, 2004; Rammstedt & Rammsayer, 2002a,b) demonstrated the existence of the *'Hubris-Humility Effect'* (HHE). The effect holds that females underestimate their abilities, whereas males overestimate theirs. The *'hubris-humility'* is strongest on mathematical/ logical and spatial intelligences but also occurs on verbal and overall intelligences. A meta-analytical study of gender differences in self-estimated intelligences found the biggest effect sizes for mathematical/logical (d =.44) and spatial (d =.43) intelligences (Szymanowicz, Chamorro-Premuzic, & Furnham, 2011, unpublished manuscript), further confirming the view of male normativeness of intelligence by laymen (Furnham, 2000, 2001). Accordingly, a novel type of intelligence, the *'Domain-Masculine Intelligence Type'* (DMIQ), that is a composite of mathematical/logical and spatial intelligences (Furnham & Gasson, 1998; Gardner 1983) is introduced and proposed to be the best predictor of HHE as well as gender differences in SEI.

Data confirming HHE's existence was found across geographies and in many socio-economic climates and cultures (e.g. Furnham, Fong, & Martin, 1999; Furnham, Hosoe, & Tang, 2003; Furnham, Shahidi, & Baluch, 2002; Swami & Furnham, 2010; von Stumm et al, 2009). Thus, HHE seems to represent a commonly held view that men are better in maths and sciences than are women (e.g. Bennett, 1996, 1997;

Bethsworth, 1999; Beyer, 1990, 1998; Halpern et. al., 2007; Hyde, Fennema, & Lamon, 1990; Sax, 1994; Sax & Harper, 2007).

The exact causes of HHE are yet to be identified but a number of factors was suggested to play a role, such as social and gender-role stereotypes and self-stereotyping (Aronson & Steele, 2005; Dar-Nimrod, 2007; Guimond et al., 2006), self-enhancement and self-derogatory biases (Beyer, 1990, 1999; Kwan et al., 2008), over-confidence and lack of confidence (Beyer & Bowden, 1997; Sleeper & Nigro, 1982), inaccurate self-estimates (e.g. Pallier, 2003; Roberts, 1991), parental influences (e.g. Beloff, 1992), socially desirable responding (Vispoel & Forte Fast, 2000), gender differences in gender-role (Bem, 1974) and self-constructs (Eccles, 1987; Rosenberg, 1965) and male dominance in the narrower intelligence strata (e.g. Chamorro-Premuzic & Furnham, 2005; Lynn et al., 2002) were suggested to play a role in the effect (see Chapter 1, section 1.3 for more information).

This chapter and the studies contained herein seek to demonstrate that ability belief systems are powerful determinants of human behaviour that tend to be most extreme in areas susceptible to widely-held gender stereotypes.

Equally, it should be clarified from the onset that this thesis does not aspire to validate Gardner's theory of multiple intelligences (1983) or to contribute to the ongoing discussion about sex differences in cognitive abilities ('*g*') (e.g. Colom et al, 2002; Halpern et al, 2007; Lynn & Irwing, 2002; Spelke, 2005; van der Sluis et al., 2008; Voyer et al, 1995).

The aim of this chapter is to confirm the existence of HHE, in particular on DMIQ. Equally, it aims to establish that DMIQ is the best predictor of gender differences in the ten self-estimated intelligences. Finally, it seeks to determine whether gender is the best predictor of DMIQ.

2.2. Study 1

Hubris-Humility Effect and the Validity of the Domain-Masculine Intelligence Type

2.2.1. Introduction

This study sets out to confirm the existence of the Hubris-Humility Effect (HHE) in self-estimated intelligences (SEI). Secondly, it aims to establish that Domain-Masculine Intelligence Type (DMIQ) is the most '*sensitive predictor*' of gender differences in SEI. DMIQ is central to this thesis as it tests out the premise that the general population perceives intelligence as male-normative and the assertion that gender differences in SEI are most pronounced on the numerical factor (e.g. Furnham, 2001; Furnham, Clark, et al., 1999; Furnham, Fong, et al, 1999; Swami & Furnham, 2010).

Thus, it was hypothesised that HHE will be most pronounced on DMIQ (H1). It was also expected that HHE will be most pronounced on DMIQ compared to the mathematical/logical and spatial intelligences separately (H2). HHE was not expected to occur on the eight remaining self-assessed intelligences, i.e. verbal, musical, body-kinaesthetic, interpersonal, intrapersonal, existential, spiritual, and naturalistic (H3). Gender was expected to be the best predictor of DMIQ over and above age, ethnic background and highest educational qualification (H4).

2.2.2. Method

Participants

A total of one hundred and thirty participants took part in this study. There were 77 (59%) females and 53 males. Their ages raged from 17 to 70 (M = 25.95, SD = 11.65). 49% of the participants reported their ethnic background as Caucasian, 27% as Far-East Asian, 16 % as Subcontinent Asian, and 2% as African. 58% of participants have completed their education to A-level, 5% achieved non-university level of education, 24% achieved BA/BSc level, 9% achieved MA/MSc/MBA level and 2% earned PhD/Doctorate as their highest level of education. 75% of the participants were single, 19% were married or living with a partner, 1% was divorced and 1% was widowed. 36% of participants were the oldest child, 32% the youngest, 18% the middle child and 15% the only child. 54% of the participants were native English speakers, 15% were native Chinese speakers, 12% were native Russian speakers, 9% were native Danish speakers, 2% were Italian and 2% were Yoruba native speakers. All participants were fluent in English and no problems were reported during completion assessment session.

Measures

Self-estimated Intelligence (SEI) (Furnham & Gasson, 1998)

This is a simple half-page questionnaire based on that developed by Furnham and Gasson (1998). The measure was used in all self-estimated intelligence programmic studies by Furnham and his collaborators (e.g. Furnham & Akande, 2004; Furnham & Chamorro-Premuzic, 2005; Furnham & Mottabu, 2004; Furnham & Rawles, 1995, 1999; Furnham, Shahidi, & Baluch, 2002; Swami & Furnham, 2010). The measure consists of a normal IQ score distribution (M = 100, SD = 15) with descriptive labels and a normal distribution IQ curve figure. The average score is 100, a score of 55 is labelled ‘mild retardation’, a score of 75 a ‘borderline retardation’, a score of 85 ‘low average’, score of 115 ‘high average’, score of 130 ‘superior’, and that of 145 ‘gifted’. Thereafter, a table with the ten labelled and briefly described intelligence types and the overall- estimated IQ score was provided, e.g. ‘*Verbal/Linguistic Intelligence: the ability to speak fluently along with understanding of grammar (syntax) and meaning (semantics)’*. The ten intelligences were based on Gardner (1983) and comprise of verbal, mathematical, spatial, musical, body-kinaesthetic, interpersonal, intrapersonal, existential, spiritual, and naturalistic intelligences. The participants were asked to estimate their ten own actual intelligences as well as their overall IQ scores by providing an actual IQ score estimate. Alpha for Domain-Masculine Intelligence Type was .62 and the inter-item correlation r =.45.

Procedure

Participants were recruited among first year undergraduate students, who were participating in an introductory psychology class at University College London. Pilot study revealed that it took approximately 15 minutes to complete the survey. Participants were given hard copies of the survey with detailed instructions. Detailed feedback about the purpose of the study was provided at the end of the session. Participants were aware that they were able to withdraw from the study at any time and to leave any questions unanswered. In accordance with the Ethics requirements of the Psychology Department as well as BPS ethical procedures, informed consent was sought from all participants before the surveys were handed out.

2.2.3. Results

2.2.3.1. Hubris and Humility Effect, the Domain-Masculine Intelligence Type and Mathematical and Spatial Intelligences

An independent samples t-test, t(127) = -5.18, *p* =. 00, two-tailed, confirmed significant differences between males (M = 117.72, *SD* = 13.72) and females (*M* = 106.41, *SD* = 11.01) in the Domain-Masculine Intelligence Type. The magnitude of the differences in the means (mean difference = -11.31, 95% CI:-15.64 to -6.99) was large (η^2[4] = .17; Hedge's Adjustment *d* =.90). Hypothesis 1 was confirmed.

In order to confirm the incremental predictive power of DMIQ over mathematical/logical and spatial intelligences, two independent samples t-tests were computed for mathematical/logical and spatial intelligences independently. For mathematical/logical intelligence the results were significant, t (128) = -4.22, *p* = .000, two-tailed (M_{male} = 119.25, SD_{male} = 16.53, vs. M_{female} = 108.17, SD_{female} = 13.34; Mean Difference =.-11.08, 95% CI: -16.27 to -5.88; η^2 = .12, Hedge's Adjustment[5] *d* =.73). For spatial intelligence, significant differences in scores were also found t (127) = -4.37, *p* = .000, two-tailed (M_{male} = 116.12, SD_{male} = 15.60, vs. M_{female} = 104.65, SD_{female} = 13.91; Mean Difference =.-11.47, 95% CI: -16.66 to -6.28; η^2 = .13, Hedge's Adjustment =.77). Thus, the effect sizes revealed that DMIQ was a better predictor of HHE than the two intelligences independently. Hypothesis 2 was confirmed.

---

[4] η^2 s the proportion of the total variance that is attributed to an effect. It is calculated as the ratio of the effect variance to the total variance $\eta^2 = SS_{effect} / SS_{total}$ (Field, 2005).

[5] Hedge's Adjustment is a Cohen's *d* measure based on sample size (Devilly, 2010). Similar to Hedge's *g* (Hedges, 1981) and like the other effect size measures, it is based on a standardised difference $g = \bar{x}1 - \bar{x}2 / s^*$. But its pooled standard deviation s^* is computed differently from Cohen's *d* and the bias for the population effect size (*θ)* is corrected. Hedges and Olkin (1985) refer to the unbiased estimator g^* as *d,* but it is not the same as Cohen's *d*. $J(a) = \Gamma(a/2) / \sqrt{a/2}\Gamma((a-1)/2)$ (Wikipedia, November, 2010).

In order to determine whether HHE is confined to occur on DMIQ and mathematical/logical and spatial intelligences, a number of independent samples t-tests with the multiple SEI were conducted. Results are presented in Table 2.2.1.

Table 2.2.1: Summary Statistics and Effect Sizes for DMIQ and 10 Multiple Self-Assessed Intelligences- Total Sample and Per Gender

Intelligence Type	Total M (SD) n	Males M (SD) n	Females M (SD) n	t(df)	Mean Difference	95% CI Lower to Upper²	η^2	*d*
DMIQ	110.97 (13.34) 129	117.72 (13.72) 52	106.41 (11.01) 77	-5.18(127)***	-11.31	-15.64 to -6.99	**.17**	**.90**
Math/Log	112.68 (15.64) 130	119.25 (16.53) 53	108.17 (13.34) 77	-4.22(128)***	-11.08	-16.27 to -.5.88	**.12**	**.73**
Spatial	109.27 (15.61) 129	116.12 (15.60) 52	104.65 (13.91) 77	-4.37(127)***	-11.47	-16.66 to -6.28	**.13**	**.77**
Verbal	110.23 (12.67) 130	109.70 (13.22) 53	110.60 (12.36) 77	0.40(128)	0.90	-3.59 to 5.39	.00	.07
Musical	100.68 (17.07) 130	99.23 (18.95) 53	101.69 (15.70) 77	0.81(128)	2.46	-3.58 to 8.50	.01	.14
Body-kinaesthetic	109.26 (12.55) 129	111.25 (13.24) 52	107.92 (11.96) 77	-1.48(127)	-3.33	-7.76 to 1.11	.02	.26
Inter-personal	113.16 (12.61) 130	112.89 (13.79) 53	113.35 (11.82) 77	0.21(128)	0.46	-4.01 to 4.93	.00	.04
Intra-personal	112.49 (13.14) 130	112.68 (14.91) 53	112.36 (11.88) 77	-0.13(128)	-0.32	-4.98 to 4.34	.00	.02
Existential	109.39 (13.95) 129	108.87 (15.84) 52	109.74 (12.62) 77	0.35(127)	0.88	-4.10 to 5.85	.00	.06
Spiritual	102.33 (15.06) 129	101.58 (18.73) 52	102.83 (12.08) 77	0.43(80)	1.25	-4.60 to 7.11	.00	.08
Naturalistic	105.09 (13.07) 129	106.92 (14.32) 52	103.84 (12.10) 77	-1.32(127)	-3.08	-7.71 to 1.55	.01	.23

* p < .05. ** p < .01. *** p < .001. *d* = is Hedge's Adjustment , i.e. Cohen's *d* measure based on sample size. *Note:* Large effect sizes are in bold.

Confirming the hypothesis, significant gender differences were only observed on DMIQ, mathematical/logical and spatial intelligences and not on the remaining eight SEI. Hypothesis 3 was confirmed.

2.2.3.2. Gender, Age, and Ethnic and Educational Background as Predictors of the Domain-Masculine Intelligence Type

The relationship between Domain-Masculine Intelligence Type, gender, age, ethnic background and highest educational qualifications was explored. Age was included because it has been shown to be correlated with psychometric intelligence and possibly influence SEI (e.g. Ackerman, 2006; Beier & Ackerman, 2003). The results of the correlational analysis are presented in Table 2.2.2. The only significant relationship between Domain-Masculine Intelligence Type was with gender (r = .42 p=.00), with males providing higher scores than males (M_{Male} = 117.72, SD_{Male} = 13.72; M_{Female} = 106.41, SD_{Female} = 11.01). A negative correlation was observed between age and ethnic background (r =-.33, p =.00) and a positive correlation between age and educational qualifications (r =.33, p=.00). No other significant relationships were noted.

Table 2.2.2: Correlations, Means and Standard Deviations between DMIQ, Gender, Age, Ethnic and Educational Background

		DMIQ	G	A	E	EQ
	X	110.97	1.41	25.95	2.06	2.86
	(SD)	(13.34)	(.49)	(11.65)	(1.36)	(1.37)
Domain-masculine IQ	(DMIQ)					
Gender	(G)	.42***				
Age	(A)	.08	-.10			
Ethnicity	(E)	-.02	.03	-.33***		
Educ. Qualifications	(EQ)	-.04	-.16	.33***	-.03	

* p < .05 ** p < .01 *** p < .001 (2-tailed). *N between 129 and 130.*

2.2.3.3. Gender as the best predictor of Domain-Masculine Intelligence Type

To investigate whether the correlational pattern for DMIQ differed for males and females, the data was split per gender and correlations recomputed. Results are presented in Table 2.2.3. No significant relationships were observed between the intelligence type and the entered variables.

Table 2.2.3: Correlations, Means and Standard Deviations between DMIQ, Age, Ethnic and Educational Background – Per Gender

Variables	DMIQ Males	DMIQ Females
M	111.72	106.41
(SD)	(13.72)	(11.01)
n	53	77
Age	.10	.18
Ethnicity	-.05	-.02
Educational Q.	.08	-.01

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

Table 2.2.4. shows the hierarchical regression results. Gender, age, ethnic background and highest educational qualifications were the predictor variables and the Domain-Masculine Intelligence Type was the criterion variable. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. Stepwise method was used for each block. Gender ($\beta = .42$, $p =.00$, $r_{part} = .42$) was entered in Step 1, explaining 18% of the variance in the Domain-Masculine Intelligence Type. When the remaining variables, i.e. age, ethnic and educational background, were added at Step 2, gender ($\beta = .43$, $p =.00$, $r_{part} = .42$) continued to be the only significant predictor of the intelligence type, explaining 18% of variance. The overall regression was significant, $F(4,124) = 7.32$, $p = .00$, $f^2 =.23$, with the overall model explaining 19% of total variance in DMIQ. Gender was the best predictor of DMIQ. Hypothesis 4 was confirmed. Thus, hypotheses 1, 2, 3 and 4 were confirmed.

Table 2.2.4: Hierarchical Regression of Gender, Age, Ethnic and Educational Background onto DMIQ

Regression Models	***Domain-Masculine IQ*** *Standardised β*	***t***	r_{part}
Step 1:			
Gender	.42	5.18***	.42
Regression Model[1]		F(1, 127) = 26.81***	
R^2		.17	
R^2 Change		.17	
Adj. R^2		.17	
f^2[6]		.21	
Step 2:			
Gender	.43	5.24***	.42
Age	.14	1.52	.12
Ethnicity	.02	.18	.01
Educational Q.	-.02	-.19	-.02
Regression Model²		F(4,124) = 7.32***	
R^2		.19	
R^2 Change		.02	
Adj. R^2		.17	
f^2		.23	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). *Note:* Significant values are in bold.

2.2.4. Discussion

The first aim of this study was to corroborate the existence of Hubris-Humility Effect on DMIQ. The existence of HHE on overall, mathematical, and spatial self-estimated intelligences, with males reporting significantly higher values than females, is extensively documented (e.g. Furnham, 2001; Furnham, Hosoe, and Tang, 2003; Swami & Furnham, 2010; von Stumm, et al., 2009). The male "*hubris*" was shown to occur in various performance estimation conditions and on various instruments, while the actual male performance was significantly lower than the estimates (cf. Reilly &

[6] Cohen's f^2 is an effect size measure that is used for multiple regressions, simple and hierarchical. The f^2 effect size measure for multiple regression is defined as: $f^2 = R^2 / 1 - R^2$, where R^2 is the squared multiple correlation. The f^2 effect size measure for hierarchical multiple regression is defined as: $f^2 = R^2_{AB} - R^2_A / 1 - R^2_{AB}$, where R^2_A is the variance accounted for by a set of one or more independent variables A, and R^2_{AB} is the combined variance accounted for by $_A$ and another set of one or more independent variables $_B$. f^2_A effect sizes of 0.02, 0.15, and 0.35 are referred to as *small, medium* and *large* respectively (Cohen, 1988; Field, 2005;Wikipedia, November, 2010).

Mulhern, 1995). The opposite was true for female *'humility'*, which (non-significantly) provided performance estimates lower than the actual performance (Reilly & Mulhern, 1995). Thus, it seems likely that other factors influence the male and female estimation processes.

In fact, a growing body of evidence (e.g. Aronson & Steele, 2005; Dar-Nimrod & Heine, 2006; Dar-Nimrod, 2007; Ehrlinger & Dunning, 2003; Good et al., 2008; Halpern et al., 2007; Meelissen & Luyten, 2008; Steele, 1997) revealed that the female lack of confidence, avoidance and underperformance on numerical tasks and disciplines is due to females being more susceptible to societal gender-role stereotypes about mathematical and scientific performance. Situations that evoked female math underperformance stereotypes caused a sharp decline in female math performance (Dar-Nimrod & Heine, 2006; Dar-Nimrod, 2007; Ehrlinger & Dunning, 2003). As with stereotype threat (Steele, 1997), gender-role stereotypes undermined math performance and learning motivation in females (Good et al., 2008). Equally, stereotypical and normative biases led to gender-specific self-fulfilling prophecies (Beyer, 1990, 1998), which might have inherently fed HHE.

Furnham's assertion (2001) that laymen view intelligence as male normative and as a composite of mathematical and spatial intelligences was also tested in the first hypothesis. The results validated the hypothesis and confirmed that HHE was most pronounced on the Domain-Masculine Intelligence Type. The observed effect size for DMIQ was large (η^2 =.17, d =.90) and the mean difference between male and female self-estimates was -11.31 IQ points, which is considerably bigger than the reported sex differences in mathematical and spatial abilities (cf. Halpern et al., 2007; Hyde, Fennema, & Lamon, 1990a; Lynn & Irwing, 2002; Voyer et al., 1995).

Secondly, this study aimed to establish that Domain-Masculine Intelligence Type (DMIQ) was a stronger predictor of HHE than were mathematical/logical and spatial intelligences individually. The results confirmed DMIQ as the best predictor of HHE ($\eta^2 = .17$, $d = .90$), compared to spatial ($\eta^2 = .13$, $d = .77$) and mathematical/ logical ($\eta^2 = .12$, $d = .73$) intelligences. The fact that DMIQ was a more powerful predictor of gender differences in SEI than were the two individual intelligences confirmed Aristotle's holistic notion that the whole is more than the sum of its parts.

The observed effect sizes were also bigger than the reported effect sizes for sex differences in mathematics (e.g. $.16 < d < .32$, Halpern et al., 2007) and spatial abilities ($d = .15 < d < .73$, Feingold, 1988; Linn & Petersen, 1985; Voyer, Voyer, & Bryden, 1995). While the observed effect sizes in SEI mirror the observed effect sizes in *'real'* intelligence, they are considerably bigger. Thus, individuals seem to over-emphasize gender differences in SEI and believe their estimates are a correct representation of sex differences in specific intelligences. Swim (1994) argued that this is due to the fact that individuals over-attribute gender differences in stereotypes. An example of an over-emphasized stereotype is aggression in males. This finding also provides further evidence that gender-role stereotypical beliefs play a role in HHE.

The results also confirmed the third hypothesis, since HHE was only observed on DMIQ, mathematical and spatial intelligences and not on the eight remaining self-estimated intelligences, providing further support for the laymen's *'numerical'* perception of intelligence and the male-normativeness of intelligence hypothesis. Finally, gender was confirmed as the best predictor of DMIQ over and above age, ethnic background and educational qualification, accounting for 18% of explained variance.

2.3. Study 2

Hubris-Humility Effect and the Domain-Masculine Intelligence Type

2.3.1. Introduction

This study set out to validate the findings of Study 2.1. Thus, it was hypothesised that the Hubris-Humility Effect will be most pronounced on the Domain-Masculine Intelligence Type (H1) and more pronounced than on the mathematical/logical and spatial intelligences independently (H2). HHE was not expected to occur on the eight remaining self-estimated intelligences (H3). Conclusively, gender was expected to be the best predictor of DMIQ over and above age (H4).

2.3.2. Method

Participants

A total of one hundred and fifteen University College London undergraduate psychology students took part in this study. There were 77 females (67%) and 38 males. Their age raged from 17 to 46 (M = 19.46, SD = 4.06) years. All participants were fluent in English, with 68% native English speakers, 14% native Chinese speakers, 4% native Russian, 3% native Persian, and 2% native Swedish speakers. 47% of participants claimed they held neutral political convictions, 14% held right-wing and 39% left-wing political convictions. No problems were reported during the testing session.

Measures

Self-estimated Intelligence (SEI) (Furnham & Gasson, 1998)

See Study 1 (section 2.2.2). Alpha for Domain-Masculine Intelligence Type was .34 and the inter-item correlation r =.21.

Procedure

Participants were second year psychology students, who took part in this study as part of their coursework. Hard copies of the survey with detailed instructions were handed out. Participants were aware that they were able to withdraw from the study at any time and to leave any questions unanswered. In accordance with the Ethics requirements of the Psychology Department as well as BPS ethical procedures, informed consent was sought from all participants before the surveys were handed out. Participants were debriefed at the end of the session.

2.3.3. Results

2.3.3.1. Hubris and Humility Effect, the Domain-Masculine Intelligence Type and Mathematical and Spatial Intelligences

An independent samples t-test, $t(113) = -3.49$, p =. 00, two-tailed, confirmed significant differences between males (M = 111.04, *SD* = 9.22) and females (*M* = 104.73, *SD* = 9.06) in the Domain-Masculine Intelligence Type. The magnitude of the differences in the means (mean difference = -6.31, 95% CI:-9.89 to -2.73) was medium ($\eta^2 = .10$; Hedge's Adjustment d =.69). Hypothesis 1 was confirmed.

In order to confirm the incremental predictive power of Domain-Masculine Intelligence Type of gender differences in self-assessed abilities over mathematical/logical and spatial intelligences, two independent samples t-tests were also computed for mathematical/logical and spatial intelligences independently. For mathematical/logical intelligence the results were significant, t (113) = -2.29, p = .000, two-tailed (M_{male} = 111.84, SD_{male} = 13.12, vs. M_{female} = 106.10, SD_{female} = 12.38; Mean Difference =.-5.74, 95% CI: -10.70 to -.78; η^2 = .06, Hedge's Adjustment d =.45). For spatial intelligence, significant differences in scores were also found t (113) = -3.07, p = .000, two-tailed (M_{male} = 110.24, SD_{male} = 11.65, vs. M_{female} = 103.35, SD_{female} = 11.17; Mean Difference =.-6.89, 95% CI: -11.34 to -2.44; η^2 = .08, Hedge's Adjustment =.60). Here too, the most profound gender differences in provided self-estimates of ability were found on the Domain-Masculine Intelligence Type. For an overview of independent samples t-tests and effect sizes see Table 2.3.1. As such, the intelligence type was the best predictor of Hubris-Humility Effect. Hypothesis 2 was confirmed.

In order to investigate whether Hubris-Humility Effect is limited to DMIQ, mathematical/logical and spatial intelligences and not the remaining eight intelligences, a number of independent samples t-tests were conducted. Results are presented in Table 2.3.1. In agreement with the hypothesis, significant gender differences were observed only on DMIQ, mathematical/logical and spatial intelligences. However, significant gender differences were also observed on verbal intelligence. Hypothesis 3 was partially confirmed.

Table 2.3.1: Summary Statistics and Effect Sizes for DMIQ and 10 Multiple Self-Assessed Intelligences-Total Sample and Per Gender

Intelligence Type	Total M (SD) n	Males M (SD) n	Females M (SD) n	t(df)	Mean Difference	95% CI Lower to Upper²	η^2	*d*
DMIQ	106.81 (9.55) 115	111.04 (9.22) 38	104.73 (9.06) 77	-3.49(113)**	-6.31	-9.89 to -2.73	.10	.69
Math/Log	108.00 (12.86) 115	111.84 (13.12) 38	106.10 (12.38) 77	-2.29(113)*	-5.74	-10.70 to -.78	.06	.45
Spatial	105.63 (11.74) 115	110.24 (11.65) 38	103.35 (11.17) 77	-3.07(113)**	-6.89	-11.34 to -2.44	.08	.60
Verbal	110.87 (12.04) 115	114.87 (11.42) 38	108.90 (11.92) 77	-2.56(113)*	-5.97	-10.59 to -1.35	.05	.51
Musical	102.96 (15.88) 115	104.74 (15.15) 38	102.08 (16.25) 77	-.84(113)	-2.66	-8.90 to 3.59	.01	.17
Body-kinaesthetic	104.72 (11.26) 115	106.18 (11.48) 38	104.00 (11.16) 77	-.98(113)	-2.18	-6.61 to 2.24	.01	.19
Inter-personal	113.97 (10.95) 115	112.89 (11.60) 38	114.49 (10.65) 77	.74(113)	1.60	-2.71 to 5.91	.00	.14
Intra-personal	112.75 (12.48) 115	113.42 (13.96) 38	112.42 (11.76) 77	-.41(113)	-1.01	-5.92 to 3.91	.00	.04
Existential	111.03 (12.01) 115	109.42 (13.76) 38	111.82 (11.06) 77	1.01(113)	2.38	-2.32 to 7.11	.01	.19
Spiritual	105.83 (12.06) 115	108.92 (12.04) 38	104.31 (11.86) 77	-1.95(113)	-4.61	-9.29 to .07	.03	.38
Naturalistic	100.49 (11.39) 114	102.86 (10.52) 38	99.35 (11.68) 77	-1.55(112)	-3.51	-8.00 to .97	.02	.31

* p < .05. ** p < .01. *** p < .001. *d* = is Hedge's Adjustment , i.e. Cohen's *d* measure based on sample size.

2.3.3.2. Gender and Age as Predictors of the Domain-Masculine Intelligence Type

The relationship between DMIQ, gender and age was explored. Age was included because of its relationship with fluid and crystallised intelligence and SEI (e.g. Ackerman, 2006; Beier & Ackerman, 2003; Deary et al., 2001). The results of the correlational analysis are presented in Table 2.3.2. The only significant relationship between DMIQ was with gender (r = .31, *p*=.01), with males providing higher scores than females ($M_{Male} = 111.04$, $SD_{Male} = 9.22$; $M_{Female} = 104.73$, $SD_{Female} = 9.06$). No other significant relationships were noted.

Table 2.3.2: Correlations, Means and Standard Deviations between DMIQ, Gender and Age

		DMIQ	G	A
	X	106.81	1.33	19.46
	(SD)	(9.55)	(.47)	(4.06)
Domain-masculine IQ	(DMIQ)			
Gender	(G)	.31*		
Age	(A)	.13	-.10	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).   *N =115.*

2.3.3.3. Gender as the best predictor of Domain-Masculine Intelligence Type

Table 2.3.3. shows results of simultaneous multiple regression. Gender and age were predictor variables and the DMIQ was the criterion variable. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. The overall regression was significant, F(2.112) = 7.95, $p = .01$, Adj. $R^2 = .11$, $f^2 = .14$, with the overall model explaining 12% of total variance in the Domain-Masculine Intelligence Type. Gender ($\beta = .33$, $p = .00$, $r_{part} = .33$) was the best predictor of the Domain-Masculine Intelligence Type, accounting for 11% of variance. Hypothesis 4 was confirmed. Thus, hypotheses 1, 2, and 4 were confirmed and hypothesis 3 was partially confirmed.

Table 2.3.3: Simultaneous Multiple Regression of Gender and Age onto DMIQ

Regression Models	***Domain-Masculine IQ*** Standardised *β*	***t***
Gender	**.33**	3.70***
Age	.17	1.85
Regression Model²		F(2,112) = 7.95**
R^2		.12
R^2 Change		.12
Adj. R^2		.11
f^2		.14

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). *Note:* Significant values are in bold.

2.3.4. Discussion

As in Study 1, results of Study 2 authenticated the existence of HHE on DMIQ and confirmed the intelligence type as the most sensitive predictor of male overestimation and female underestimation of ability (η^2 =.10, *d* =.69). Just as in Study 1, gender differences were also observed on the mathematical/logical (η^2 =.06, *d* =.45) and spatial intelligences (η^2 =.08, *d* =.60). However, compared to Study 1, the observed effect sizes were smaller and more similar to the observed effect sizes in specific *'real'* intelligences. Equally, gender was affirmed as best predictor of DMIQ, over and above age, accounting for 11% of explained variance.

The only notable difference with Study 1 was the fact that HHE was also observed on verbal intelligence (η^2 =.05, *d* =.51), with males providing higher self-estimates. The observed effect size was medium. This finding is not unique as previous SEI studies reported male hubris in verbal abilities (e.g. Furnham, Callahan, & Akande, 2004; Furnham & Chamorro-Premuzic, 2005; Furnham, Hosoe, & Tang, 2003; Swami & Furnham, 2010). However, a meta-analytical study of gender differences in SEI revealed that males often provided lower verbal ability estimates than did females. The observed effect size for verbal abilities was also the smallest one (*d* =.07) among the investigated self-estimated intelligences (Szymanowicz, Chamorro-Premuzic, & Furnham, 2011, unpublished manuscript). Similarly, the observed medium effect size in verbal SEI in favour of men differs from the cross-culturally reported medium effect size in verbal abilities in favour of women (e.g. Halpern et al., 2007; Halpern & Wright, 1996; Ogle et al., 2003).

2.4 Summary

This chapter set out to corroborate the existence of the Hubris-Humility Effect in self-estimated intelligences, in particular on the numerical factor as well as to confirm the male-normativeness of intelligence stereotypical beliefs that were previously reported in the SEI research programme (e.g. Furnham, 2001; Furnham, Hosoe, & Tang, 2001; Furnham & Rawles, 1995; von Stumm et al., 2009). The results of Studies 1 and 2 validated these previous findings.

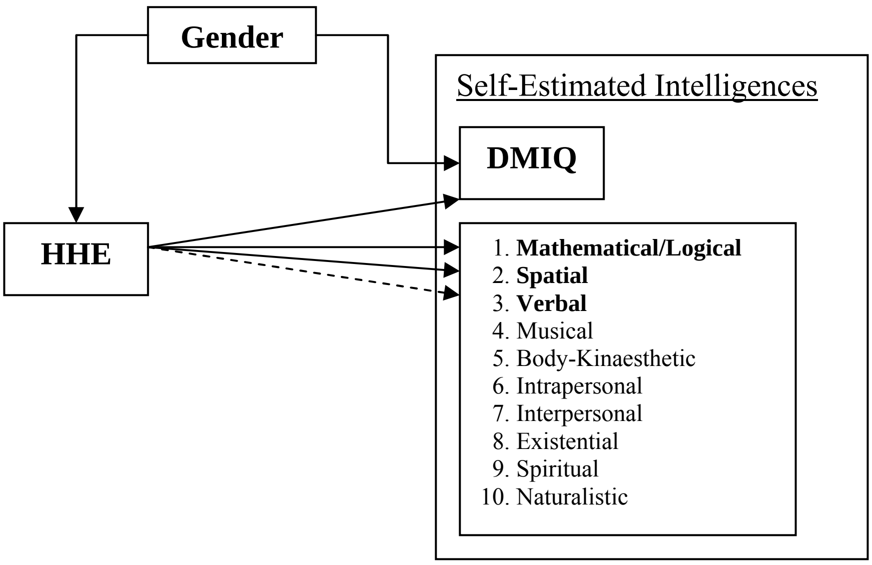
Study 1 introduced Domain-Masculine Intelligence Type in order to test whether the composite variable was a better predictor of gender differences in SEI than mathematical/logical and spatial intelligences individually. The results confirmed DMIQ as the most sensitive predictor of gender differences in SEI. Equally, gender was confirmed as the best predictor of the intelligence type.

Study 2 set out to replicate the findings of Study 1. At large, the findings validated results of Study 1 as well as the major findings in the existing SEI literature. Interestingly, the observed effect sizes in SEI were substantially larger than the reported effect sizes in *'real'* specific intelligences, providing further evidence for the assertion that gender differences in SEI or in HHE are caused by over-emphasis of gender attributes in peoples' stereotypes (cf. Swim, 1994).

Equally, the observed mean differences in DMIQ were -11.31 IQ points in Study 1 and -6.31 IQ points in Study 2. These values are higher than the reported sex differences in general and specific intelligences (cf. Feingold, 1988; Halpern et al, 2007; Hyde, 1996; Hyde, Fennema, & Lamon, 1990a; Linn & Petersen, 1985; Lynn & Irwing, 2002, 2004; Voyer et al, 1995) that are usually reported as male advantage of 4-5 IQ points.

The findings of Studies 1 and 2 are represented in Figure 2.4.1. The single-pointed arrows symbolize the direct relationship between two variables that were either predicted or found. The dashed arrow (between HHE and verbal SEI) represents a relationship that was not predicted. Variables that exhibited a relationship with DMIQ and HHE are in bold. The direction of the arrows implies causality that current results.

Figure 2.4.1: Pictorial representation of the findings of Studies 1 and 2



In summary, Chapter 2 provided further support for the existence of HHE in SEI, in particular on DMIQ, and affirmed the male-normativeness view of intelligence. In spite of the observed generational IQ gains or the so-called '*Flynn* *effect'* (Flynn, 1987), commonly-held gender-role stereotypes and performance biases in SEI are powerful and possibly damaging (e.g. Dar-Nimrod & Heine, 2006; Good et al., 2008). More research in suitable counter-strategies is necessary, such as training females to approach math differently (Kawakami et al., 2008) and removing all possible stereotype threats (Good et al., 2008). It must be reiterated that measuring actual cognitive sex differences in specific intelligences were beyond the scope of this thesis.

Chapter 3: Hubris-Humility Effect, Domain-Masculine Intelligence Type and ‘*g*’

3.1. General Introduction

Chapter 2 consisted of two correlational studies that confirmed the existence and limitation of the Hubris-Humility Effect on the numerical-spatial factor of self-estimated intelligences as well as established the predictive power of the Domain-Masculine Intelligence Type within the ten self-estimated intelligences. Equally, it also established that gender was the best predictor of DMIQ.

Chapter 3 continues to investigate the role gender plays in DMIQ and the relationship between HHE and DMIQ. It also introduces possible new determinants of DMIQ – psychometric intelligence ‘*g*’, implicit intelligence beliefs, gender identity concepts and self-constructs. Chapter 3 also seeks to confirm the major findings of previous literature on this topic.

The consent among social psychologists is that self-assessments are *‘subjective’* and as such unreliable and prone to overinflated or deflated ability beliefs and performance biases (Alicke et al., 1995; DeNisi & Shaw, 1977; Guenther & Alicke, 2010; Kruger & Dunning, 1999; Mabe & West, 1982; Moore & Small, 2007). The current consensus in the SEI research programme is that self-estimates of intelligence are not suitable as substitutes for psychometric intelligence, with the observed correlations usually in the region of r =.30 (Ackerman & Wolman, 2007; Borkenau & Liebler, 1993; Chamorro-Premuzic, Furnham, & Moutafi, 2004; Davidson & Downing, 2000; Furnham, 2005; Furnham & Fong, 2000; Furnham & Mottabu, 2004; Furnham & Rawles, 1999; Furnham, von Stumm, et al., 2009; Holling

& Preckel, 2005; Mabe & West, 1982; Paulhus et al., 1998; Rammstedt & Rammsayer, 2002b; Reilly & Mulhern, 1995; Visser, Ashton, & Vernon, 2008).

However, various researchers have demonstrated that individuals are capable of accurate self-estimates of ability (Ackerman et al., 2002; Allik et al., 2010; Borkenau & Liebler, 1993; Chamorro-Premuzic et al., 2010; Furnham & Rawles, 1999; Gati et al., 2006, Holling & Preckel, 2005; Kim et al., 2008; Swim, 1994), with SEI correlations with '*g*' ranging between $.26 < r < .55$, $.41 < r < .50$ with mathematical, and $.39 < r < .52$ with spatial abilities. Gender was also shown to moderate the relationship between self-estimated and psychometric intelligence, with males providing higher self-estimates of intelligence than females, even when psychometric intelligence scores were controlled for, suggesting that the gender differences in SEI are not a reflection of sex differences in psychometrically assessed intelligence (e.g. Furnham, Fong, & Martin,1999; Furnham & Rawles, 1999; Holling & Preckel, 2005; Rammstedt & Rammsayer, 2000b, 2001, 2002b; Reilly & Mulhern, 1995). To date, no studies in the SEI programme have investigated whether gender identity concepts and self-constructs influence the observed gender differences on the numerical-spatial factor of SEI (here assessed through DMIQ).

Study 3 introduces measures of fluid (Gf) and crystallised intelligence (Gc) in order to determine what role '*g*' plays as a predictor of DMIQ. It also aims to determine whether gender or '*g*' is the best predictor of DMIQ.

Study 4 seeks to replicate the results of Study 3 and ascertain whether gender or '*g*' is the best predictor of DMIQ. Implicit intelligence beliefs (Dweck, 1999) are included to determine whether they play a role in the prediction of DMIQ.

Study 5 aims to corroborate the findings of Study 3 and 4, in particular the role '*g*' plays in DMIQ. Gender identity variables, i.e. masculinity and femininity, and self-construct measures, i.e. self-esteem and self-control are introduced to determine their role as predictors of DMIQ as well as their relationship with DMIQ and gender.

3.2. Study 3

‘g’ and Gender as Predictors of the Domain-Masculine Intelligence Type: The role of fluid and crystallised intelligences

3.2.1. Introduction

The primary goal of this study was to investigate the role fluid (Gf) and crystallised (Gc) intelligences and gender play in the prediction of Domain-Masculine Intelligence Type. DMIQ was the most sensitive predictor of HHE in Studies 1 and 2. Hence, it was predicted that HHE’s occurrence on DMIQ would prevail in the current population (H1). Previous findings demonstrated that gender influenced the relationship between SEI and psychometrically assessed intelligence and that the gender differences remained even after ‘*g*’ was controlled for (e.g. Holling & Preckel, 2005; Rammstedt & Rammsayer, 2001, 2002b). Thus, gender was expected to influence the relationship between total ‘*g*’, i.e. the two fluid and one crystallised psychometric intelligence measures combined, and DMIQ (H2).

No uniform agreement exists about male advantage in general intelligence (‘*g*’) (e.g. Colom & Garcia-Lopez, 2002; Deary et al., 2003; Halpern et al., 2007; Jackson & Rushton, 2006; Lynn, 1999; Spelke, 2005), with historically no sex differences presumed, as evidenced by the development of standardised intelligence tests (e.g. Ackerman, 2006; Wechsler, 1944). Recent investigations have supplied contradicting data, with male advantage reported on various measures of Gf and Gc, such as General Knowledge Test (GKT), Raven’s Standard and Progressive Matrices (SPM) and Wechsler Adult Intelligence Scale (WAIS) (cf. Lynn, Allik, & Irwing, 2004; Lynn, Irwing, & Cammock, 2001; Lynn, Wilberg, & Margraf-Stiksrud, 2004).

Accordingly, male advantage was expected to be observed on the three psychometric measures - Baddeley Reasoning Test (BRT) (H3), Wonderlic Personnel Test (WPT) (H4), and General Knowledge Test (GKT) (H5). In Studies 1 and 2 gender was the best predictor of DMIQ. Consequently, it was hypothesised that gender would remain the best predictor of Domain-Masculine Intelligence Type over and above '*g*' as measured by BRT, WPT and GKT (H6).

3.2.2. Method

Participants

A total of eighty-five University College London undergraduate students took part in this study. There were 73 females (86%) and 12 males. Their age raged from 17 to 40 (M = 19.28, SD = 3.32) years.

Measures

Domain-Masculine Intelligence Type (DMIQ)

Based on the self-estimated measure (Furnham & Gasson, 1998) that was used in Chapter 2 (section 2.2.2), this is a shortened version with the exact same properties and layout, but containing only mathematical and spatial intelligences. The alpha for the Domain-Masculine Intelligence Type in this study was .76 and the inter-item correlation was r =.61.

Intelligence Measures

Fluid Intelligence (Gf):

Wonderlic Personnel Test (WPT: Wonderlic, 1992)

This 50-item test can be administered in 12 minutes and measures general intelligence. Scores can range from 0 to 50. Items include word and number comparisons, disarranged sentences, serial analysis of geometric figures and story problems that require mathematical and logical solutions, clearly measuring Gc and Gf. The test correlates very highly (r = .92) with the WAIS-R (Wechsler, 1981, see Wonderlic, 1992). The mean for the current study was 23.42 (SD = 6.38).

Baddeley Reasoning Test (BRT: Baddeley, 1968)

This 64-item test is administered in 3 minutes and measures Gf through logical reasoning. Scores can range from 0 to 64. Each item is presented in the form of grammatical transformation and answered with 'true/false', e.g. 'A precedes B-AB' (true), 'A does not follow B-BA' (false). It represents the quickest reliable measure of gf (Chamorro-Premuzic & Furnham, 2005). The mean score in this study was 29.79 (SD = 13.09).

Crystallized Intelligence (Gc):

General Knowledge Test (GKT: Irwing, Cammock, & Lynn, 2001)

This 72-item questionnaire is administered in 20 minutes and assess knowledge of the following areas: literature, general knowledge, science, medicine, games, fashion and finance. The mean score for the current population was 29.86 (SD = 10.24). The questionnaire has satisfactory psychometric properties (Furnham & Chamorro-Premuzic, 2006).

Procedure

All participants were first year psychology students, who took part in this study as part of an introductory psychology course. Participants had no background or in-depth knowledge of psychology and psychometric instruments. Tests were administered by three experimenters in a large and quiet lecture room. The ability measures were completed first, with a short break after each psychometric measure, followed by the DMIQ measure. Participants were fully debriefed about the purpose of the study at the end of the testing session. All participants were fluent in English and no problems were reported during the testing session.

3.2.3. Results

3.2.3.1. Hubris and Humility Effect and the Domain-Masculine Intelligence Type

An independent samples t-test, t(69) = 3.75, *p* =. 00, two-tailed, confirmed significant differences between males (M = 115.96, *SD* = 17.10) and females (*M* = 100.60, *SD* = 11.97) in the DMIQ. The magnitude of the differences in the means (Mean Difference = 15.36, 95% CI: 7.19 to 23.52) was large (η^2 = .17; Hedge's Adjustment *d* =1.05). Hypothesis 1 was confirmed.

3.2.3.2. Impact of Gender and Total 'g' on the Domain-Masculine Intelligence Type

The total scores of the three psychometric intelligence measures, the Baddeley Reasoning Test (BRT), the Wonderlic Personnel Test (WPT), and the General Knowledge Test (GKT) were combined, creating a new variable Total '*g*' (α = .69, inter-item *r* =.50). Total '*g*' was collapsed into a categorical variable with three groups, with Group 1 containing subjects that had the lowest Total '*g*' scores. Group 2 was made of subjects that had average Total '*g*' scores and Group 3 was made of subjects with highest Total '*g*' scores. Results are presented in Table 3.2.1.

Table 3.2.1: Overview of Total 'g' Banded

	Tot 'g' score	n
Group 1	<=70	30
Group 2	71-92	28
Group 3	93+	27

Note: Computed using Visual Bander technique (SPPS 13.0)

A two-way between-groups analysis of variance was conducted to explore whether gender influences the relationship between Total '*g*' and the DMIQ. Results are presented in Table 3.2.2.

Table 3.2.2: 2-way ANOVA (Tot 'g' and gender) on DMIQ

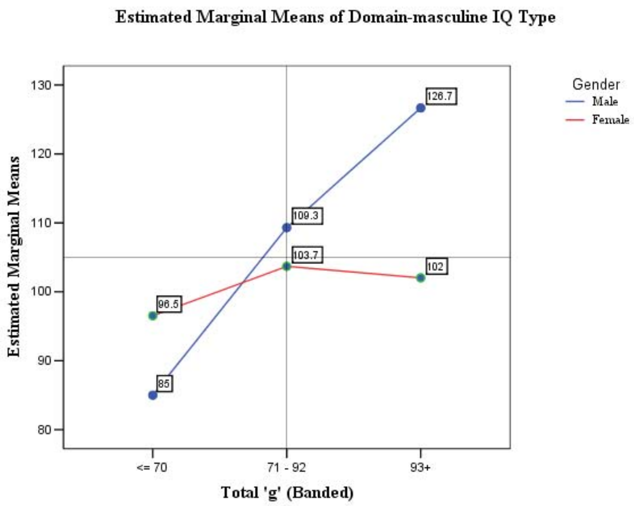
Variable	Tot '*g*' score	Mean Score (SD)			*F*-score		
		Total	Males	Females	Tot '*g*'	Gender	Tot '*g*' x Gender
DMIQ	G1 (L)	96.00 (10.19)	85.00 (10.11)	96.52 (10.13)	6.75**	1.65	5.04**
	G2 (M)	104.85 (14.27)	109.30 (13.84)	103.68 (14.51)			
	G3 (H)	103.20 (14.08)	115.96 (17.10)	102.03 (10.26)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: DMIQ1 = Domain-Masculine Intelligence Type.

The interaction effect between Total '*g*' and gender was significant, F (2,65) = 5.04, $p < .01$, $\eta_p^2 = .13$, with medium effect size. The main effect for Total '*g*', F (2,65) = 6.75, $p <.01$, $\eta_p^2 = .17$ was significant, with large effect size. The main effect for gender was not significant, F (1,65) = 1.65, $p =.20$, $\eta_p^2 = .03$. Planned contrasts revealed significant differences between Group 1 and Group 3 (Contrast Estimate – 19.66, $p <.01$). Post-hoc comparisons using Tukey HSD and Bonferroni tests indicated that mean scores for Group 1 (<=70) differed significantly from mean scores for Group 2 (71-92). Mean scores for Group 1 also significantly differed from mean scores for Group 3 (93+). Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets.

As the main interaction effect was significant, a further investigation of the relationship was warranted. Simple effects analysis was conducted. Data was split per gender and two one-way between-groups analysis of variance were conducted. For males, the one-way between-groups analysis of variance for Total '*g*' and DMIQ was significant, F (2,9) = 6.23, $p <.05$, $\eta^2 =.58$, with large effect size. As only one subject fell in the Group 1 (<=70) the robust tests of equality of means Welch and Brown-Forsyth as well as the Post Hoc tests were not computed. For females, there was no statistically significant difference in the three Total '*g*' groups on DMIQ, F(2,56) = 2.06, *p*=.14. The robust tests of equality of means, Welch (2,36) = 2.06, p =.13; Brown-Forsythe (2,48) = 2.03, p =.14 were not significant. The post-hoc comparisons using the Tukey HSD and Bonferroni tests revealed no significant differences between the three groups. Hypothesis 2 was confirmed.

Figure 3.2.1: 2-way ANOVA (Tot 'g' and gender) on DMIQ



3.2.3.3. Gender Differences in 'g'

In order to investigate whether gender differences occurred on the two Gf measures, BRT and WPT, independent samples t-tests were conducted. An independent samples t-test for BRT, t(83) = .32, *p* =.75, two-tailed, was not significant. Independent samples t-test for WPT, t(83) = 2.96, p <.01, two-tailed, confirmed significant differences between males and females. Thus, hypothesis 3 was not confirmed and hypothesis 4 was confirmed. To test whether gender differences occurred on the Gc measure, GKT, an independent samples t-test was computed. The results revealed significant differences, t(83) = 3.86, *p* =.00, two-tailed, between males and females. Hypothesis 5 was confirmed. An overview of the results is in Table 3.2.3.

Table 3.2.3: Independent Samples t-Tests and Effect Sizes for Three Psychometric Measures

	Males M (SD)	Females M (SD)	*t(df)*	Mean Diff.	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
BRT	30.92 (12.92)	29.60 (13.20)	.32(83)	1.31	6.84	9.47	.00	.10
WPT	28.25 (7.61)	22.63 (5.84)	2.96(83)**	5.62	1.84	9.40	.10	.82
GKT	39.67 (9.87)	28.25 (9.43)	3.86(83)***	11.42	5.54	17.30	.17	1.17

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Hedge's Adjustment or Cohen's d adjusted for sample size.

3.2.3.4. Gender and 'g' as Predictors of the Domain-Masculine Intelligence Type

The relationship between DMIQ, gender, '*g*' and age was explored. Age was included as it was shown to influence observed gender differences in SEI and sex differences in '*g*' (e.g. Beier & Ackerman, 2001, 2003; Deary et al., 2003; Lynn & Irwing, 2004; Rammstadt & Rammsayer, 2002b). Results of the correlational analysis are presented in Table 3.2.4. Gender correlated negatively (r = -.41, *p*=.00), with the Domain-Masculine Intelligence Type, with males providing higher scores than males (M_{Male} = 115.96, SD_{Male} = 17.10; M_{Female} = 100.60, SD_{Female} = 11.97).

From the three intelligence tests, only two correlated with the intelligence type. Positive correlations were observed between WPT (r=.48, *p* =.00) and the intelligence type as well as between GKT (r =.24, *p* <05) and the intelligence type. The three intelligence tests were strongly inter-correlated. A medium strength positive[7] relationship was observed between age and DMIQ (r =.38, *p* <.01). A negative strong relationship was observed between age and gender (r=-.50, *p* =.00).

[7] This thesis uses Cohen's (1988, pp. 79-81) guidelines for interpretation of correlation values, that is *r* =.10 to .29 is small, *r* = .30 to .49 is medium and *r* = .50 to 1.0 is large. These guidelines are identical for the positive and negative *r* values.

The age range of participants (17 years) was not disproportionate. To further investigate the impact of age on the intelligence type and reduce its impact on Gf, age was partialled out and the correlational matrix recomputed. The results are presented in Table 3.2.4.

Table 3.2.4: Correlations and Partial Correlations, Means and Standard Deviations between DMIQ, Gender, 'g' and Age

		DMIQ	G	BRT	WPT	GKT	A
	X	103.20	1.86	29.79	23.42	29.86	19.28
	(SD)	(14.08)	(.35)	(13.09)	(6.38)	(10.24)	(3.22)
Domain-masculine IQ	(DMIQ)						
Gender	(G)	-.41***					
Baddeley Reasoning Test	(BRT)	.23	-.04				
Wonderlic Personnel Test	(WPT)	.48***	-.31**	.56***			
General Knowledge Test	(GKT)	.24*	-.39***	.36**	.59***		
Age	(A)	.38**	-.50***	.03	.33**	.37**	
-Controlled for Age-							
Domain-masculine IQ	(DMIQ)						
Gender	(G)	-.28*					
Baddeley Reasoning Test	(B)	.24	-.03				
Wonderlic Personnel Test	(W)	.41***	-.18	.59***			
General Knowledge Test	(GK)	.12	-.26*	.38***	.53***		

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
N: between 71 and 85.

The inspection of the partial correlational matrix revealed no considerable differences in the pattern of the significant relationships, with the following exceptions. GKT and the DMIQ no longer correlated when age was partialled out nor did gender and WPT. The size of the observed partial correlations was smaller than the initial correlations. An independent samples t-test for age was not significant; t(11) 2.16, p =.54.

3.2.3.5. Gender as the best predictor of Domain-Masculine Intelligence Type

To further investigate whether the correlational patterns differed for males and females, the data was split per gender and the correlations recomputed. Results are presented in Table 3.2.5. For males, the only significant relationship that was observed was a strong positive correlation between the DMIQ and GKT (r = .63, p <.05). For females, the only significant relationship that was observed was a medium strength positive correlation between the intelligence type and the WPT (r =.38, p <.05).

Table 3.2.5: Correlations, Means and Standard Deviations between DMIQ, 'g' Measures, Intelligence Beliefs and Age – Per Gender

Variables	DMIQ Males	DMIQ Females
M	115.96	100.60
(SD)	(17.10)	(11.97)
n	12	59-73
BRT	.26	.24
WPT	.50	.38*
GKT	.63*	-.06
Age	.40	.09

* p < .05 ** p < .01 *** p < .001 (2-tailed).

Table 3.2.6. shows the results of a hierarchical regression analysis. Gender and the three psychometric measures were predictor variables and the DMIQ was the criterion variable. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. Stepwise method was used for each block. Gender (β = -.41, p =.00, r_{part} = -.41) was entered in Step 1, explaining 17% of the variance in DMIQ. The two fluid and the crystallised intelligence measures were added in Step 2, with gender (β = -.33, p <.01, r_{part} = -.30) remaining a significant predictor, explaining 9% of variance. From the three measures, only WPT (β = .48, p <.01, r_{part} = .34) significantly contributed to the

prediction of the intelligence type, explaining 12% of variance. The overall regression was significant, F(4,66) = 8.01, p= .00, f^2 =.49, with the overall model explaining 33% of total variance in DMIQ. Contrary to prediction, gender was not the best predictor of the DMIQ. WPT, a fluid intelligence measure was the best predictor of DMIQ, followed by gender. Hypothesis 6 was not confirmed.

Thus, hypotheses 1, 2, 4 and 5 were confirmed. Hypotheses 3 and 6 were not confirmed.

Table 3.2.6: Hierarchical Regression of Gender and Three Psychometric Measures onto DMIQ

Regression Models	***Domain-Masculine IQ***		
	Standardised β	***t***	r_{part}
Step 1:			
Gender	**-.41**	-3.75***	**-.41**
Regression Model[1]		F(1, 69) = 14.08***	
R^2		.17	
R^2 Change		.17	
Adj. R^2		.16	
f^2		.20	
Step 2:			
Gender	**-.33**	-2.94**	**-.30**
Baddeley Reasoning T.	.01	.07	.01
Wonderlic Personnel T.	**.48**	3.32**	**.34**
General Knowledge T	-.17	-1.31	-.13
Regression Model[2]		F(4, 66) = 8.01***	
R^2		.33	
R^2 Change		.16	
Adj. R^2		.29	
f^2		.49	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). *Note:* Significant values are in bold.

3.2.4. Discussion

The main focus of this study was to investigate the role '*g*' and gender play in prediction of DMIQ. As in Chapter 2, the existence of HHE on DMIQ was validated (η^2 = .17, d = 1.05), providing further evidence that large gender differences in SEI occur on the numerical-spatial factor.

The proposition that gender would influence the relationship between '*g*' and SEI was tested in hypothesis 2. A '*g*' x Gender ANOVA revealed a significant interaction and significant '*g*' effect, but no gender effects. Significant differences in DMIQ self-estimates were observed between the lowest and the highest and the lowest and medium general ability groups, with lowest DMIQ estimates provided by the lowest '*g*' group, highest DMIQ estimates by the medium '*g*', followed by slightly lower DMIQ estimates by the highest ability group. Males provided higher DMIQ self-estimates in the medium and highest general ability groups. Females provided higher DMIQ estimates than males in the lowest ability group. Further analyses revealed that men, but not women, provided significantly different DMIQ self-estimates in the three general ability groups.

These results seem to imply that individuals are reasonably rather accurate in estimating their abilities, except for the highest ability group. These results are in line with Swim's (1994) proposition that people are accurate or underestimate their scores. Equally, the hubris-humility effect seems to be replicated, with females providing higher self-estimates on the lowest ability group while males provided higher self-estimates on the higher ability groups. However, it is also possible that accurate self-assessment and self-knowledge played a role.

To test the male advantage on Gf and Gc measures, independent samples t-tests were computed. Results confirmed male advantage on WPT and GKT, but no significant gender differences were found for BRT. These findings are in line with the existing literature in the field (e.g. Ackerman, 2006; Lynn & Irwing, 2004; Lynn, Irwing, & Cammock, 2001). Correlational results affirmed female disadvantage on the three psychometric measures, but only two were significant.

This study also set out to validate the finding of Study 1 and 2 that gender is the best predictor of DMIQ. Correlational results confirmed that for males, DMIQ only correlated with GKT, and for females only with WPT. These results could mean that males and females vary in their '*definitions*' of intelligence, i.e. their understanding of what makes up intelligence. Based on the results, it seems that men hold more *'holistic'* or Cattell-ian definitions of DMIQ, while females adhere to more *'traditional or conservative'* views of DMIQ, reminiscent of Spearman. Hierarchical regression analysis revealed that WPT was the best predictor of DMIQ, accounting for 12% of explained variance. Contrary to prediction, gender was the second best predictor, explaining 9% of variance in DMIQ. Thus, a measure of general intelligence was a better predictor of the intelligence type than gender.

3.3. Study 4

<u>Gender, ‘g’ and Beliefs about Intelligence as Predictors of DMIQ</u>

3.3.1. Introduction

This study set to substantiate the findings of Study 3 and further examine the role gender and fluid (Gf) and crystallised (Gc) intelligence measures play in DMIQ. Hence, it was predicted that HHE would be reconfirmed on DMIQ (H1). Gender was expected to influence the relationship between total ‘*g*’, i.e. the fluid, general and crystallised psychometric intelligence measures combined and DMIQ (H2). Male advantage was expected to occur on the three ‘*g*’ measures, BRT (H3), WPT (H4), and GKT (H5).

In addition, a measure of implicit Intelligence Beliefs (Dweck, 1999) was introduced to determine whether such beliefs play role in DMIQ. These beliefs assert that individuals believe attributes, intelligence, abilities, and motives are either *‘malleable’* (incremental theory) or ‘*fixed’* (entity theory). Consequently, entity and incremental views of intelligence arouse different motivational attitudes and responses to success and failure. As such these beliefs are unrelated to ‘*g*’ but as with other self-theories they influence peoples’ beliefs systems and impact performance (Dweck, 1999). Consequently, no significant relationship was expected between ‘*g*’ and implicit Intelligence Beliefs but between DMIQ and Intelligence Beliefs (H6). Gender was expected as the best predictor of DMIQ over and above ‘*g*’ and the implicit Intelligence Beliefs (H7).

3.3.2. Method

Participants

A total of one hundred and twenty-one University College London undergraduate psychology students took part in this study. There were 82 females (68%) and 39 males. Their age raged from 17 to 24 (M = 19.13, SD = 1.32) years.

Measures

Domain-Masculine Intelligence Type (DMIQ)

See Study 3 (section 3.2.2). The alpha in this study was .80 and the inter-item correlation was .66.

Intelligence Measures

Fluid Intelligence (Gf):

The Baddeley Reasoning Test (BRT: Baddeley, 1968)

See Study 3 (section 3.2.2). The mean score in this study was 28.65 (SD = 12.51).

The Wonderlic Personnel Test (WPT: Wonderlic, 1992)

See Study 3 (section 3.2.2). The mean for the current study was 25.80 (SD = 6.62).

Crystallized Intelligence (Gc):

General Knowledge: General Knowledge (GKT: Irwing, Cammock, & Lynn, 2001)

See Study 3 (section 3.2.2). The mean score for the current population was 31.11 (SD = 10.34).

Intelligence Beliefs (Dweck, 1999)

Eight items from the Intelligence Beliefs measure (Dweck, 1999) were used, such as: *"I am not very confident about my intellectual ability"*, *"I believe I have a* *certain amount/level of intelligence and there is now not much I can do to change it"*, "*I could learn to do much better on intelligence tests if I wanted to",* and *"When I get new project/work, I am usually pretty sure I will be good at it"*. Items with reversed or negative wording about one's ability, indicated low belief in ability, low achievement drive or inclination to entity belief of intelligence (Dweck, 1999). This measure requires subjects to report their beliefs and judgements about malleability or inflexibility of intelligence and their ability. The alpha was .96 and the inter-item correlation was .73.

Procedure

The participants were first year students, who took part in this study as part of their coursework and who were unfamiliar with the concepts and measures. The tests were administered by three experimenters in a large and quiet lecture room. The ability measures were completed first, followed by a short break after each psychometric measure. The Domain-Masculine Intelligence Type measure and the Intelligence Beliefs measure were completed last. Participants were fully debriefed about the purpose of the study at the end of the session. All participants were fluent in English. No problems were reported during the testing session.

3.3.3. Results

3.3.3.1. Hubris and Humility Effect and the Domain-Masculine Intelligence Type

An independent samples t-test, t(119) = 7.46, *p* =. 00, two-tailed, confirmed significant differences between males (M = 120.64, *SD* = 14.34) and females (*M* = 102.59, *SD* = 11.45) in the DMIQ. The magnitude of the differences in the means

(Mean Difference = 18.05, 95% CI: 13.26 to 22.84) was very large ($\eta^2 = .32$; Hedge's Adjustment d =1.38). Hypothesis 1 was confirmed.

3.3.3.2. Impact of Gender and Total 'g' on the Domain-Masculine Intelligence Type

As in Study 3 the total scores of the three psychometric intelligence measures were combined, creating a new variable Total '*g*' ($\alpha = .55$, inter-item r =.34) that was then collapsed into a categorical variable with three groups. Group 1 contained subjects that had the lowest Total '*g*' scores; Group 2 was made of subjects that had average Total '*g*' scores and Group 3 was made of subjects with highest Total '*g*' scores. Results are presented in Table 3.3.1.

Table 3.3.1: Overview of Total 'g' Banded

	Tot '*g*' score	n
Group 1	<=73	41
Group 2	74-95	40
Group 3	96+	40

Note: Computed using Visual Bander technique (SPPS 13.0)

A two-way between-groups analysis of variance was conducted to explore whether gender moderates the relationship between Total '*g*' and DMIQ. Results are given in Table 3.3.2.

Table 3.3.2: 2-way ANOVA (Tot 'g' and gender) on DMIQ

Variable	Tot '*g*' score	Mean Score (SD)			*F*-score		
		Total	Males	Females	Tot '*g*'	Gender	Tot '*g*' x Gender
DMIQ	G1 (L)	101.40 (12.93)	107.00 (19.64)	101.40 (12.93)	5.46**	25.48***	7.59**
	G2 (M)	107.60 (12.72)	112.64 (13.36)	105.69 (12.16)			
	G3 (H)	116.40 (15.56)	127.43 (9.02)	101.47 (8.16)			

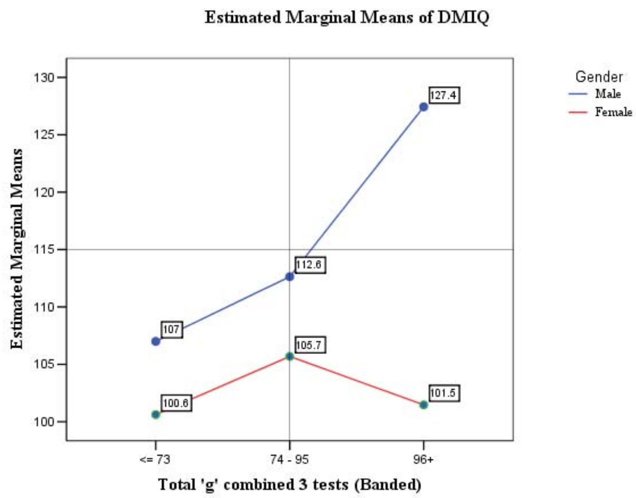
* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: DMIQ = Domain-Masculine Intelligence Type.

Levene's Test of Equality of Error Variance was significant (p <.05), indicating the groups variances were not equal. An alternative check for comparing variances was used. Firstly, the largest and the smallest standard deviations were squared. The largest squared SD was divided by the smallest squared SD, with resulting value of 1.72, which is smaller than the recommended value of 2, suggesting that the group variances, albeit not equal, were tolerable. Subsequently, the significance level was adjusted to p <.01.

The interaction effect between Total '*g*' and gender was significant, F (2,115) = 7.59, $p < .01$, $\eta_p^2 = .12$, with large effect size. The main effect for Total '*g*' was significant, F (2,115) = 5.46, $p < .01$, $\eta_p^2 = .09$, with medium effect size. The main effect for gender was also significant, F (1,115) = 25.48, $p = .00$, $\eta_p^2 = .18$, with large effect size. Planned contrasts revealed significant differences between Group 1 and Group 3 (Contrast Estimate –8.00, p <.05). Post-hoc comparisons using Games-Howell test indicated that mean scores for Group 1 (<=73) differed significantly from mean scores for Group 2(74-95). Mean scores for Group 2 also significantly differed from mean scores for Group 3 (96+). Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets.

As the main interaction effect was significant, a further investigation of the relationship was warranted. Simple effects analysis was conducted. Data was split per gender and two one-way between-groups analysis of variance were conducted. For males, the Levene's Test of Equality of Error Variance was significant (p <.05), indicating the groups variances were not equal. An alternative check for comparing variances was used. Firstly, the largest and the smallest standard deviations were squared. The largest squared SD was divided by the smallest squared SD, with resulting value of 4.72, which is bigger than the recommended value of 2, suggesting that the group variances were very equal. Subsequently, the significance level was adjusted to $p < .01$.

Figure 3.3.1: 2-way ANOVA (Tot 'g' and gender) on DMIQ



For males, the one-way between-groups analysis of variance for Total '*g*' and DMIQ was significant, $F(2,36) = 9.48$, $p = .00$, $\eta^2 = .35$, with large effect size. The robust tests of equality of means, Welch $(2,8) = 6.98$, $p < .05$; Brown-Forsyth $(2, 8) = 5.42$, $p < .05$ were significant. The Post Hoc tests using Games-Howell test revealed that mean scores for Group 2 (74-95) ($M_{Total} = 112.64$, $SD_{Total} = 13.36$) differed significantly from mean scores of Group 3 (96+) ($M_{Total} = 127.34$, $SD_{Total} = 9.02$).

For females, there was no statistically significant difference in the three Total '*g*' groups on DMIQ, $F(2,79) = 1.70$, $p = .19$. The robust tests of equality of means, Welch $(2,48) = 1.55$, $p = .22$; Brown-Forsythe $(2,76) = 1.93$, $p = .15$ were not significant. The post-hoc comparisons using the Tukey HSD and Bonferroni tests revealed no significant differences between the three groups. Hypothesis 2 was confirmed.

3.3.3.3. Gender Differences in 'g'

In order to investigate whether gender differences occurred on the two measures of fluid intelligence, the Baddeley Reasoning Test (BRT) and the Wonderlic Personnel Test (WPT), two independent samples t-tests were conducted. For BRT, t(94) = .31, p =.75, two-tailed, the test was not significant. For WPT, t(119) = 5.62, p =.00, two-tailed, the results confirmed significant differences between males and females. Thus, hypothesis 3 was not confirmed and hypothesis 4 was confirmed. An independent samples t-test was computed to investigate whether gender differences occurred on the crystallised intelligence measure, the General Knowledge Test (GKT). The results revealed significant differences, t(119) = 5.01, p =.00, two-tailed, between genders. Hypothesis 5 was not confirmed. Results are presented in Table 3.3.3.

Table 3.3.3: Independent Samples t-Tests and Effect Sizes for Three Psychometric Measures

	Males M (SD)	Females M (SD)	*t(df)*	Mean Diff.	95% CI		Effect Size	
					L	U	η^2	*d*
BRT	29.13 (10.45)	28.43 (13.44)	.31(94)	.70	3.74	5.14	.00	.13
WPT	30.18 (5.90)	23.72 (5.92)	5.62(119)***	6.46	4.18	8.74	.21	1.09
GKT	37.33 (9.22)	28.15 (9.54)	5.01(119)***	9.192	5.55	12.82	.17	.97

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Hedge's Adjustment or Cohen's d adjusted for sample size.

3.3.3.4. Gender, 'g', Intelligence Beliefs as Predictors of the Domain-Masculine Intelligence Type

First, the 8-item Intelligence Beliefs about Intelligence measure was analysed, using Principal Component Analysis, in order to identify the underlying structure of the measure. Prior to performing PCA, the suitability of data for factor analysis was assessed. Inspection of the correlation matrix revealed the presence of many coefficients of .3 and above. The Kaiser-Mayer-Olkin sampling measure value was .93 and exceeding the recommended value (Kaiser, 1970). The Bartlett's Test of Sphericity was significant, χ^2 (28) = 962, p =.00, supporting the factorability of the correlation matrix (Pallant, 2007).

The initial solution was rotated, using the Direct Oblimin procedure and factors with eigenvalues greater than 1.00. Absolute values less than .40 were suppressed. PCA revealed one component with eigenvalues over 1, accounting collectively for 76.37% of explained variance in the data. An inspection of the screeplot confirmed a clear break after the first component. Parallel Analysis (Monte Carlo PCA for Parallel Analysis, Watkins, 2000) confirmed the one component solution.

As only one component was extracted and the solution was not rotated, Pattern and Structure Matrices were not computed. Items 3, 5, 7 and 8 were recoded and the analysis recomputed, yielding same values and communalities. Table 4 shows the results of the PCA analysis, component matrix with component loadings and communalities, percentage of explained variance and alpha value. The Component Matrix revealed a simple structure with strong loadings. A single measure of Intelligence Beliefs was computed and used in further analyses.

Table 3.3.4: Principal Component Analysis (Direct Oblimin with Kaiser Normalisation) of Intelligence Beliefs (N=121) 1 Component Extracted

Component Matrix

Item	Component	Communalities
Q2	**.911**	.831
Q4	**.907**	.822
Q1	**.886**	.784
Q6	**.883**	.779
Q5	**.873**	.762
Q8	**.861**	.741
Q7	**.839**	.704
Q3	**.829**	.687
Eigenvalue	6.11	
% of Explained Variance	76.37%	
No. of Items	8	
Alpha (α)	.96	
Inter-item *r*	.72	

Note: Major loadings for each item are bolded.

The relationship between DMIQ, gender, '*g*', Intelligence Beliefs and age was explored. Results of the correlational analysis are presented in Table 3.3.5. Gender correlated negatively (r = -.56, *p*=.00), with DMIQ, with males providing higher scores than males (M_{Male} = 120.64, SD_{Male} = 14.34; M_{Female} = 102.59, SD_{Female} = 11.45).

From the three intelligence tests, only two correlated with the intelligence type. Positive correlations were observed between the WPT and DMIQ (r = .61, *p* =.00) and GKT and DMIQ (r =.29, *p* <01). The three intelligence tests were strongly inter-correlated. However, small negative correlations were observed between Intelligence Beliefs and WPT (r =-.21, *p* <.05) and GKT (r =-.21, *p* <.05) and no significant relationship was observed between Intelligence Beliefs and DMIQ. Thus, hypothesis 6 was not confirmed.

Table 3.3.5: Correlations and Partial Correlations, Means and Standard Deviations between DMIQ, Gender, 'g, Intelligence Beliefs and Age

		DMIQ	G	BRT	WPT	GKT	IQB	A
	X	108.41	1.68	28.65	25.80	31.11	3.48	19.13
	(SD)	(15.01)	(.47)	(12.51)	(6.62)	(10.34)	(1.15)	(1.32)
Domain-masculine IQ	(DMIQ)							
Gender	(G)	-.56***						
Baddeley Reasoning Test	(BRT)	.12	-.03					
Wonderlic Personnel Test	(WPT)	.61***	-.46**	.39***				
General Knowledge Test	(GKT)	.29**	-.42***	.23*	.40***			
Intelligence Beliefs	(IQB)	-.18	.17	-.14	-.21*	-.21*		
Age	(A)	.31**	-.31**	.03	.23*	.20*	-.09	
-Controlled for Age-								
Domain-masculine IQ	(DMIQ)							
Gender	(G)	-.52***						
Baddeley Reasoning Test	(B)	.12	-.02					
Wonderlic Personnel Test	(W)	.58***	-.42***	.39***				
General Knowledge Test	(GK)	.25**	-.38***	.22*	.37***			
Intelligence Beliefs	(O)	.-.16	.15	-.14	-.19*	-.20*		

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
N: between 71 and 85.

A medium strength positive relationship was observed between age and DMIQ (r =.31, p <.01). A negative strong relationship was observed between age and gender (r=--.31, p =.00). The age range of participants (7 years) was very small. To further investigate the impact of age on the intelligence type as well as to negate its influence on fluid intelligence (e.g. Beier & Ackerman, 2003), age was partialled out and the correlational matrix recomputed. The results are given in Table 3.3.5. The inspection of the partial correlational matrix revealed no differences in the pattern of significant relationships as well as in the observed values of the two correlational analyses. An independent samples t-test for age was significant; t(119) 3.52, p <.01, η^2 =.09; Hedge's Adjustment d = 1.19.

3.3.3.5. Gender as the best predictor of Domain-Masculine Intelligence Type

To further investigate whether the correlational patterns differed for males and females, the data was split per gender and the correlations recomputed. Results are presented in Table 3.3.6. For males, two significant relationships were observed; a strong positive correlation was noted between the WPT and DMIQ (r =.56, *p* =.00) as well as between GKT and DMIQ (r = .53, *p* =.00). For females, the only significant relationship that was observed was a medium strength positive correlation between WPT and DMIQ (r =.44, *p*=.00).

Table 3.3.6: Correlations, Means and Standard Deviations between DMIQ, 'g' Measures, Intelligence Beliefs and Age – Per Gender

Variables	DMIQ Males	DMIQ Females
M	120.64	102.59
(SD)	(14.34)	(11.45)
n	39	82
BRT	.15	.12
WPT	.56***	.44***
GKT	.53***	-.18
IQ Beliefs	-.17	-.04
Age	.17	.17

* *p* < .05   ** *p* < .01   *** *p* < .001 (2-tailed).

Table 3.3.7. shows the results of a hierarchical regression analysis. Gender, '*g*' as represented by the three intelligence tests and Intelligence Beliefs were the predictor variables and the DMIQ was the criterion varialbe. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. Stepwise method was used for each block. Gender (β = -.56, *p* =.00, r_{part} = -.56) was entered in Step 1, explaining 31% of the variance in domain-masculine intelligence. The three fluid and crystallised intelligence measures were added in Step 2, with gender (β = -.36, *p* =.00, r_{part} = -.30) a significant predictor, explaining 9% of variance. From the three measures, only WPT (β = .49, *p* =.00, r_{part} = .39) significantly contributed to the prediction of the

intelligence type, explaining 15% of variance. When Intelligence Beliefs were added at Step 3, only two variables significantly contributed in the prediction of the intelligence type. Gender (β = -.35, p =.00, r_{part} = -.29) explained 8% of variance and WPT (β = .49, p =.00, r_{part} = .39) accounted for 15% of the variance. None of the remaining predictor variables reached significance. The overall regression was significant, F (5,115) = 21.49, p= .00, f^2 =.92, with the overall model explaining 48% of total variance in DMIQ. As in Study 4, gender was the second best predictor of DMIQ and WPT, a fluid intelligence measure, was the best predictor of the intelligence type. Hypothesis 7 was not confirmed. Thus, hypotheses 1, 2, 4 and 6 were confirmed. Hypotheses 3, 5, and 7 were not confirmed.

Table 3.3.7: Hierarchical Regression of Gender, 'g' and Intelligence Beliefs onto DMIQ

Regression Models	***Domain-Masculine IQ*** Standardised β	t	r_{part}
Step 1:			
Gender	**-.56**	-7.46***	**-.56**
Regression Model[1]		F(1, 119) = 55.60***	
R^2		.32	
R^2 Change		.32	
Adj. R^2		.31	
f^2		.47	
Step 2:			
Gender	**-.36**	-4.43***	**-.30**
Baddeley Reasoning T.	-.07	-.97	-.07
Wonderlic Personnel T.	**.49**	5.81***	**.39**
General Knowledge T	-.04	.46	-.03
Regression Model[2]		F(4, 116) = 26.98***	
R^2		.48	
R^2 Change		.16	
Adj. R^2		.46	
f^2		.92	
Step 3:			
Gender	**-.35**	-4.37***	**-.29**
Baddeley Reasoning T.	-.07	-.99	-.07
Wonderlic Personnel T.	**.49**	5.74***	**.39**
General Knowledge T.	-.04	-.51	-.03
Intelligence Beliefs	-.03	-.48	-.03
Regression Model[3]		F(5, 115)= 21.49***	
R^2		.48	
R^2 Change		.00	
Adj. R^2		.46	
f^2		.92	

* p < .05 ** p < .01 *** p < .001 (2-tailed). *Note:* Significant values are in bold.

3.3.4. Discussion

The main focus of this study was, as in Study 3, to investigate the role ‘*g*’ and gender play in prediction of DMIQ. As in previous studies, the existence of HHE on DMIQ was validated (η^2 = .32, *d* = 1.38), providing further evidence that large gender differences in SEI occur on the numerical-spatial factor.

The proposition that gender would influence the relationship between ‘*g*’ and SEI was tested in hypothesis 2. A ‘*g*’ x Gender ANOVA revealed a significant interaction and significant ‘*g*’ and gender effects. Significant differences in DMIQ self-estimates were observed between the lowest and the medium groups and between the medium and the highest general ability groups, with lowest DMIQ estimates provided by the lowest ‘*g*’ group, medium DMIQ estimates by the medium ‘*g*’ and highest DMIQ estimates by the highest ability group.

The same pattern was observed for males, with lowest DMIQ estimates provided by the lowest ability group and the highest DMIQ estimates by the highest ability group. Males also provided higher DMIQ estimates than females in all three ability groups. Females’ DMIQ estimates in the lowest and highest ability groups were very similar, with highest DMIQ estimates provided by the medium ability group. Further analyses revealed that men in the medium and the highest ability groups provided significantly different DMIQ self-estimates. No differences were observed for females.

The results of Study 4 were similar to those in Study 3. In Study 4 individuals’ self-ratings of ability were even more in agreement with their psychometric scores than in Study 3. The findings have replicated HHE, with males providing higher self-estimates in all three ability groups, while females provided lower self-estimates.

To test the male advantage on Gf and Gc measures, independent samples t-tests were computed. Results replicated the findings of Study 3 and confirmed male advantage on WPT and GKT, with no significant gender differences on BRT. Correlational results confirmed these findings. Again, the results were in agreement with existing literature.

This study also set out to validate that gender is the best predictor of DMIQ. Correlational analysis revealed that in the male subsample, DMIQ correlated with WPT and GKT, while DMIQ only correlated with WPT in the female sub-sample. These results are slightly different from Study 3, but only for males. Here too, males and females seem to exhibit different insights of what constitutes intelligence; for males it is a combination of Gc and Gc and for females Gf. These perceptions could possibly influence the ability self-estimation process.

Contrary to prediction and previous findings, Intelligence Beliefs correlated with '*g*', challenging Dweck's assertions (1999) that '*g*' does not play role in implicit Intelligence Beliefs. Equally, no relationship was observed with DMIQ, further disproving assertions by attributional researchers (e.g. Blackwell, Trzesniewski, & Dweck, 2007; Dweck, 1999) that intelligence attributional theories impact peoples' ability belief systems.

As in Study 3, the hierarchical regression analysis revealed that WPT was the best predictor of DMIQ, accounting for 15% of explained variance. Gender was again second best predictor, explaining 8% of variance. Implicit Intelligence Beliefs did not contribute in the prediction of DMIQ and do not play a role in the prediction of the intelligence type. Hence, it appears that Gf is a more powerful determinant of DMIQ than gender.

3.4. Study 5

Gender, 'g', Gender Identity Concepts and Self-Constructs as Predictors of the DMIQ

3.4.1. Introduction

Studies 3 and 4 focused on the role '*g*' – fluid and crystallised – and gender play in the prediction of the Domain-Masculine Intelligence Type. This study aims to validate the previous findings and sets forth by introducing gender identity concepts, i.e. *'masculinity'* and *'femininity'* and self-constructs, i.e. *'self-esteem'* and *'self-control'* as possible determinants of DMIQ.

Thus, it was predicted that HHE would be observed on DMIQ (H1). Gender was expected to influence the relationship between total '*g*', i.e. fluid and crystallised psychometric intelligence measures combined and DMIQ (H2). Male advantage was expected to occur on WPT (H3) and GKT (H4).

Gender identity concepts or roles - masculinity (M) and femininity (F) - have been widely researched. The current agreement among researchers is that masculinity and femininity are culturally determined cognitive concepts used by individuals to classify their life (cf. Bem, 1974, 1981a; Lippa, 2001).

The layman's definitions of M and F are broader, incorporating personality traits, social roles, sexuality preferences and physical appearance (Lippa, 2001). In fact, M and F have been shown to correlate with gender role stereotypes (Biernat, 1991; Hirschy & Morris, 2002; Petrides, Furnham, & Martin, 2004; Rudman & Phelan, 2010) and personality traits (Marusic & Bratko, 1998). SEI studies have not included M/F in the investigation of gender differences but Furnham & Gasson (1998) proposed that national masculinity scores, as defined by Hofstede (1998), could play role in the SEI gender discrepancy. Likewise, Petrides, Furnham, and Martin (2004) reported that gender-role stereotypes play a role in the way people perceive intelligence, with psychometric intelligence perceived as masculine and emotional intelligence as feminine. While this study uses Bem's M/F measure, masculinity was expected to significantly contribute to the prediction of DMIQ (H5).

Research demonstrated a link between self-esteem and well-being, with small male advantage (e.g. Kling et al., 1999). Equally, females have been shown to have better self-control or self-discipline, which leads to superior academic performance, test results and achievement (Duckworth & Seligman, 2005, 2006). Thus, it seems plausible that high self-esteem and self-control lead to higher self-estimates of ability. Accordingly, self-esteem (H6) and self-control (H7) were expected to correlate with gender and DMIQ. Despite findings of Study 3 and 4, but in line with results of Study 1 and 2, gender was expected to be the best predictor of DMIQ, over and above '*g*', gender identity concepts and self-constructs (H8).

3.4.2. Method

Participants

A total of one hundred and two University College London undergraduate psychology students took part in this study. There were 79 females (78%) and 23 males. Their age raged from 17 to 46 (M = 19.46, SD = 4.31) years. 91% of participants stated to have accomplished A-levels as their highest educational qualification, 3% stated non-university higher education, 2% stated BA/BSc degree and 1% stated to have completed MA/MSc/MBA as their highest educational qualification. 49% of participants described themselves as Caucasian, 22% as Subcontinent Asian, 16% as Far East Asian, 2% as Caribbean and 1% as African.

92% of participants stated to be single, 4% to be married, and 2% living with partner. 37% were the youngest child, 33% the oldest, 15% were the middle child and 15% stated to be the only child in their family. 92% were right-handed and 8% were left-handed. 59% were native English speakers, 18% were native Chinese speakers, 3% were native Persian speakers, 2% were native Malay speakers, 2% were native Swedish speakers and 2% were native Punjabi speakers. All participants were fluent in English. No problems were reported during the testing session.

Measures

Domain-Masculine Intelligence Type (DMIQ)

See Study 3 (section 3.2.2). The alpha in this study was .51 and the inter-item correlation was r =.34.

Intelligence Measures

Fluid Intelligence (Gf):

The Wonderlic Personnel Test (WPT: Wonderlic, 1992)

See Study 3 (section 3.2.2.). The mean for the current study was 25.62 (SD = 5.63).

Crystallized Intelligence (Gc):

General Knowledge Test (GKT) (Irwing, Cammock, & Lynn, 2001)

See Study 3 (section 3.2.2). The mean score for the current population was 29.16 (SD = 10.24).

Gender Identity

Bem Sex Role Inventory (BSRI) (Bem, 1981a).

This non-timed 60-item measure is designed to measure the orthogonal constructs of masculinity and femininity. Each construct is made of 20 items, with the remaining 20-items measuring the gender-neutral or androgynous characteristics; the items are worded as adjectives. Items were scored using a 7-point scale, where 1 = *never or almost never true* and 7 = *almost always true,* e.g. athletic, sensitive to other's needs, solemn. The scale has been shown to have satisfactory internal reliability and homogeneity, with alphas for masculinity .86 and femininity .74 (Francis & Wilcox, 1998). The alphas for masculinity and femininity in this study were, .83 and .75, respectively.

Self-Constructs

Self-Esteem Scale (Rosenberg,1965).

This 10-item non-timed measure is designed for adults. Items were scored using a 4-point scale, where 1 = *strongly agree* and 4 = *strongly disagree* (e.g. "On the whole, I am satisfied with myself", "I certainly feel useless at times", and "I wish I could have more respect for myself"). Adequate internal reliability (alpha .85) and test-retest reliability (.87) has been reported (Pullman & Allick, 2000). The alpha in this study was .90 and the inter-item correlation r =.46.

Brief Self-Control Scale (BSCS) (Tangney et. al, 2004).

This 36-item measure is designed for adults, but it is face valid also for adolescents. Items are endorsed on a 5-point scale, where 1 = *not like me at all and* 5 = *very much like me* (e.g. "I have a hard time breaking bad habits", "I often act without thinking through all the alternatives", and "I am good at resisting temptation"). BSCS is a non-timed measure. Previous studies reported adequate internal reliability (alpha .85) and test-retest reliability (.87). The alpha in this study was .86.

Procedure

The students were second year students, who took part in this study as part of their coursework. All participants were debriefed at the end of the testing session.

3.4.3. Results

3.4.1. Hubris and Humility Effect and the Domain-Masculine Intelligence Type

An independent t-test, t(100) = -6.29, *p* =. 00, two-tailed, confirmed significant differences between males (M = 120.17, *SD* = 8.01) and females (*M* = 106.67, *SD* = 9.34) in DMIQ. The magnitude of the differences in the means (mean difference = -13.50, 95% CI:-17.77 to -9.24) was large (η^2 = .28; Hedge's Adjustment *d* =1.54). Hypothesis 1 was confirmed.

3.4.2. Impact of Gender and Total 'g' on the Domain-Masculine Intelligence Type

As in previous studies, the psychometric intelligence measures, the Wonderlic Personnel Test (WPT) and the General Knowledge Test (GKT) were combined, creating a new variable Total '*g'*. Group 1 contained subjects that had the lowest Total '*g*' scores; Group 2 was made of subject that had average Total '*g*' scores and Group 3 was made of subjects with highest Total '*g*' scores.

Table 3.4.1: Overview of Total 'g' Banded

	Tot 'g' score	n
Group 1	<=48.50	33
Group 2	48.51-62.00	34
Group 3	62.01+	31

Note: Computed using Visual Bander technique (SPPS 13.0)

A two-way between-groups analysis of variance was conducted to explore whether gender influences the relationship between Total '*g*' and DMIQ. The interaction effect between Total '*g*' and gender was not significant, $F(2,92) = .63$, $p = .53$, $\eta_p^2 = .01$. The main effect for Total '*g*', $F(2,92) = 4.97$, $p < .01$, $\eta_p^2 = .10$, was significant, with medium effect size. The main effect for gender was also significant, $F(1,92) = 32.47$, $p = .00$, $\eta_p^2 = .26$, with large effect size.

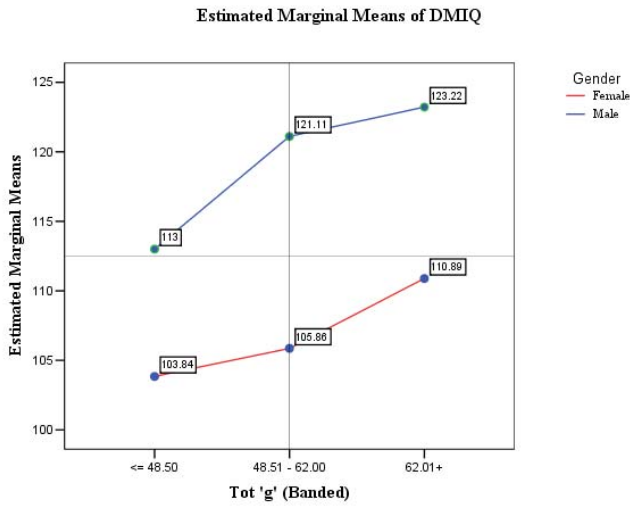
Table 3.4.2: 2-way ANOVA (Tot 'g' and gender) on DMIQ

Variable	Tot '*g*' score	Mean Score (SD)			*F*-score		
		Total	Males	Females	Tot '*g*'	Gender	Tot '*g*' x Gender
DMIQ1	G1 (L)	105.23 (11.58)	113.00 (9.08)	103.84 (11.56)	4.97**	32.47***	.63
	G2 (M)	109.90 (9.70)	121.11 (5.88)	105.86 (7.33)			
	G3 (H)	114.47 (9.15)	123.22 (7.56)	110.89 (7.18)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: DMIQ = Domain-Masculine Intelligence Type.

Planned contrasts revealed significant differences between Group 1 and Group 3 (Contrast Estimate –6.85, $p < .01$). Post-hoc comparisons using Tukey HSD and Bonferroni tests indicated that mean scores for Group 1 (<=48.50) differed significantly from mean scores for Group 3 (62.01+). Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 2 was confirmed.

Figure 3.4.1: 2-way ANOVA (Tot 'g' and gender) on DMIQ



3.4.3. Gender Differences in 'g'

Independent samples t-tests were computed in order to examine whether gender differences occurred on the WPT and GKT. Results are presented in Table 3.4.3. Hypotheses 3 and 4 were confirmed.

Table 3.4.3: Independent Samples t-Tests and Effect Sizes for Two Psychometric Measures

	Males M (SD)	Females M (SD)	*t(df)*	Mean Diff.	95% CI		Effect Size	
					L	U	η^2	*d*
WPT	28.09 (4.82)	24.90 (5.68)	-2.45(99)*	-3.19	-5.78	-.60	.06	.60
GKT	31.74 (9.18)	28.29 (10.46)	-1.42(96)*	-8.26	-8.26	1.37	.02	.35

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Hedge's Adjustment or Cohen's d adjusted for sample size.

3.4.4. Gender, 'g', Gender Identity Concepts, Self-Concept Constructs as Predictors of the Domain-Masculine Intelligence Type

The relationship between DMIQ, gender, 'g', gender identity and self-concept constructs as well as age was explored. The results of the correlational analysis are presented in Table 3.4.4.

Table 3.4.4: Correlations, Means and Standard Deviations between DMIQ, Gender, 'g', Gender Identity Concepts, Self-Constructs, and Age

		DMIQ	G	WPT	GKT	M	F	SE	SC	A
	X	109.72	1.23	25.62	29.10	4.50	4.89	1.96	3.17	19.46
	(SD)	(10.66)	(.42)	(5.63)	(10.24)	(.68)	(.57)	(.52)	(.45)	(4.31)
Domain-masculine IQ	(DMIQ)									
Gender	(G)	.53***								
WPT	(WPT)	.32**	.24*							
GKT	(GKT)	.32**	.14	.55***						
Masculinity	(M)	.26*	.32**	-02	.13					
Femininity	(F)	-.16	-.40**	-.21*	-.13	-.18				
Self-Esteem	(SE)	-.19	-.28**	.02	-.03	-.27**	.06			
Self-Control	(SC)	-.02	-.14	.11	.07	-.21*	.29**	-.13		
Age	(A)	.16	-.01	.04	.26*	.07	.13	.09	.21*	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).   *N*: between 97and 102.

Gender correlated positively (r = .53, *p*=.00), with the DMIQ, with males providing higher scores than males (M_{Male} = 120.17, SD_{Male} = 8.01; M_{Female} = 106.67, SD_{Female} = 9.34). Medium positive relationships were observed between the WPT (r=.32, *p* <.01) and DMIQ as well as between the GKT (r =.32, *p* <01) and DMIQ. Masculinity (r = .26, *p* <.05), but not femininity (r = -.16, *p* =.13), correlated positively with DMIQ. A negative relationship was observed between self-esteem and gender but no significant relationship was observed between self-esteem and DMIQ. Contrary to prediction, no significant relationships were observed between self-control and gender and DMIQ. Thus hypothesis 6 was partially confirmed and hypothesis 7 was not confirmed.

Despite the age range of participants (29 years) no significant relationship was observed between age and the intelligence type. An independent t-test for age was not significant; t(100) -.14, p =.89. No other significant relationships between DMIQ and the remaining variables were observed.

3.4.5. Gender as the best predictor of Domain-Masculine Intelligence Type

To further investigate whether the correlational patterns differed for males and females, the data was split per gender and the correlations recomputed. Results are presented in Table 3.4.5.

Table 3.4.5: Correlations, Means and Standard Deviations between DMIQ, 'g' Measures, Gender Identity, Self-Concept Constructs and Age – Per Gender

Variables	DMIQ Males	DMIQ Females
M	120.17	106.67
(SD)	(8.01)	(9.34)
n	23	79
WPT	.43*	.19
GKT	.49*	.25*
Masculinity	.02	.14
Femininity	.08	.06
Self-Esteem	.09	-.08
Self-Control	-.05	.10
Age	.10	.21

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

For males, a significant positive medium strength relationship was observed between the DMIQ and WPT, (r = .43, p <.05). A positive medium strength relationship was also observed between the intelligence type and GKT (r = -.49, p <.05). For females, the only positive medium strength relationship occurred between DMIQ and GKT (r =.25, p <.05).

Table 3.4.6. shows the results of a hierarchical regression analysis. Gender, 'g', gender identity and self-construct measures were regressed on DMIQ to ascertain which variable is the best predictor of the intelligence type. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity.

Stepwise method was used for each block. Gender (β = .53, p =.00, r_{part} = .53) was entered in Step 1, explaining 28% of the variance in DMIQ. WPT and GKT were added in Step 2, with gender (β = .48, p =.00, r_{part} = .47) being the only significant predictor of DMIQ, explaining 22% of variance. When the gender identity variables, i.e. masculinity and femininity, were added at Step 3, gender (β = .49, p =.00, r_{part} =

.42) remained the only significant predictor, explaining 18% of variance in DMIQ. None of the remaining predictor variables reached significance. When self-constructs, i.e. self-esteem and self-control, were added at Step 4, gender ($\beta = .48$, p =.00, r_{part} = .40) accounted for 16% of variance in DMIQ. As in Step 3, none of the remaining entered predictor variables reached significance. The overall regression was significant, $F(7,85) = 7.04$, p= .00, f^2 =.59, with the overall model explaining 37% of total variance in DMIQ. Hypothesis 5 was not confirmed and hypothesis 8 was confirmed. Thus, hypotheses 1, 2, 3, 4 and 8 were confirmed and hypotheses 5, 6 and 7 were not confirmed.

Table 3.4.6: Hierarchical Regression of Gender, 'g', Gender Identity Concepts and Self-Constructs onto DMIQ

Regression Models	***Standardised β***	***Domain-Masculine IQ*** ***t***	r_{part}
Step 1:			
Gender	**.53**	6.00***	**.53**
Regression Model[1]		F(1, 91) = 35.96***	
R^2		.28	
R^2 Change		.28	
Adj. R^2		.28	
f^2		.39	
Step 2:			
Gender	**.48**	5.47***	**.47**
WPT	.09	.86	.07
GKT	.20	1.98	.17
Regression Model[2]		F(3, 89) = 15.96***	
R^2		.35	
R^2 Change		.07	
Adj. R^2		.33	
f^2		.54	
Step 3:			
Gender	**.49**		
WPT	.12	4.92***	**.42**
GKT	.19	1.15	.10
Masculinity	.10	1.79	.15
Femininity	.11	1.06	.09
Regression Model[3]		1.13	.10
R^2		F(5,87) = 10.05***	
R^2 Change		.37	
Adj. R^2		.02	
f^2		.33	
		.59	
Step 4:			
Gender	**.48**	4.65***	**.40**
WPT	.12	1.19	.10
GKT	.18	1.76	.15
Masculinity	.09	.96	.08
Femininity	.10	1.03	.09
Self-Esteem	-.03	-.30	-.03
Self-Control	.01	.09	.01
Regression Model[4]		F(7,85) = 7.04***	
R^2		.37	
R^2 Change		.00	
Adj. R^2		.32	
f^2		.59	

* $p < .05$ ** $p < .01$ *** $p < .001$. *Note:* Significant values are in bold.

3.4.4. Discussion

The main focus of this study was to confirm previous findings about the role '*g*' and gender play in prediction of DMIQ. As in previous studies, the existence of HHE on DMIQ was validated ($\eta^2 = .28$, $d = 1.54$), further affirming that large gender differences in SEI occur on the numerical-spatial factor.

Next, the role gender plays in the relationship between '*g*' and SEI was investigated. Results revealed significant '*g*' and gender effects, with significant differences in DMIQ self-estimates provided by the lowest and the highest '*g*' groups. Consistent with previous findings, males provided higher DMIQ estimates on all three ability groups and supplied further support for HHE. Compared to Study 3 and 4, males' and females' DMIQ estimates were even more *'accurate'*, i.e. lowest DMIQ estimates were provided by the lowest ability group, medium estimates by the medium group and highest estimates by the highest ability group.

Subsequently, results replicated the findings of Study 3 and 4 and confirmed male advantage on WPT and GKT. Correlational results confirmed these findings. The results were in line with existing literature.

This study also set out to validate that gender is the best predictor of DMIQ. Correlational analysis revealed that in the male subsample, DMIQ correlated with WPT and GKT, further validating results of Study 4. However, the female results differed from Study 3 and 4, with GKT and not WPT significantly correlated with DMIQ.

In disagreement with the literature in the field (e.g. Duckworth & Seligman, 2005, 2006; Kling et al., 1999), no relationship was observed between gender and self-control and females had higher self-esteem than males in this sample. Equally, neither self-esteem nor self-control correlated with DMIQ.

Finally, unlike Study 3 and 4, but in line with results of Study 1 and 2, the hierarchical regression analysis revealed that gender was the best and only predictor of DMIQ, accounting for 16% of explained variance. Contrary to expectations masculinity did not significantly contribute to DMIQ prediction, nor did self-esteem or self-control. Moreover, with the introduction of gender identity and self-construct variables, the psychometric measures ceased to be the best predictor of DMIQ and gender regained its standing as the best determinant of DMIQ.

3.5. Summary

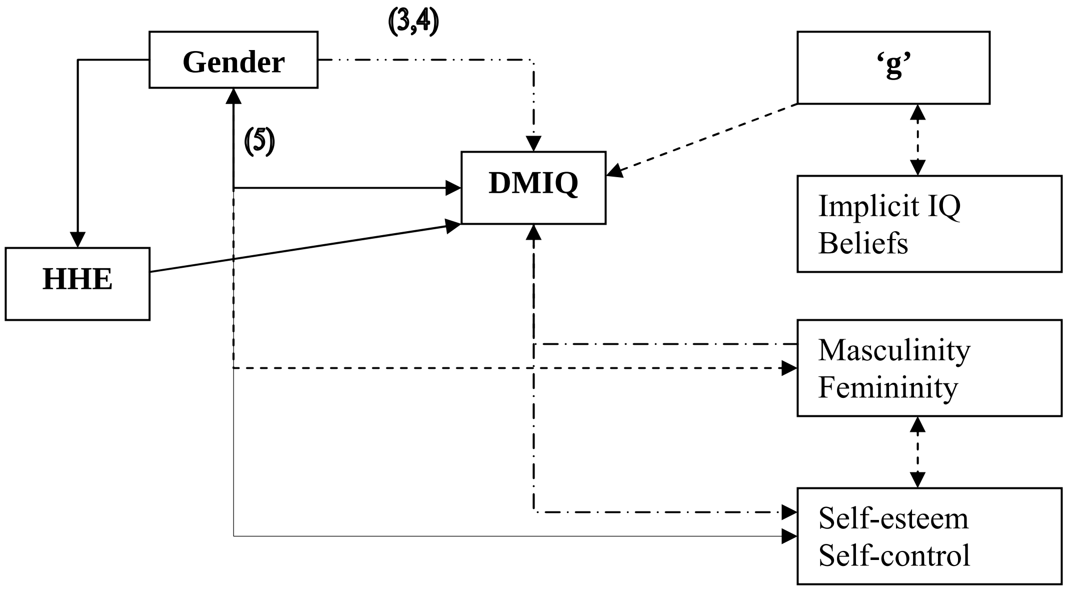
The three experiments reported in this chapter examined the role fluid (Gf) and crystallised (Gc) intelligence and gender play in the Domain-Masculine Intelligence Type. In addition, all three studies confirmed the existence of HHE on DMIQ, providing further evidence that gender differences on the mathematical/logical and spatial SEI are large indeed.

Study 3 revealed that '*g*', i.e. a measure of general intelligence, and not gender was the best determinant of DMIQ. Study 4 also introduced implicit Intelligence Beliefs to ascertain whether they play a role in the prediction of DMIQ. The results revealed that '*g*', followed by gender was the best predictor of DMIQ. Intelligence Beliefs played no role. These results differ from the findings of Studies 1 and 2. Several factors seemed to have played a role. Firstly, both studies have shown that individuals are capable of accurate self-estimates of intelligence when compared to their 'actual' intelligence scores. Secondly, for females, only fluid intelligence seems to constitute intelligence. Male understanding of intelligence, i.e. combination of crystallised and fluid intelligence, is less '*restrictive*' and resembles the academic viewpoint. Thirdly, if females view intelligence as fluid, it seems likely that exposure to situations that focus on (fluid) intelligence will activate stereotypical beliefs and '*prime'* female behaviour accordingly, as evidenced through the stereotype threat bias (Aronson & Steele, 2005; Dar-Nimrod, 2007; Steele, 1997). Consequently, such mechanisms could account for the finding that '*g*' was the best predictor of DMIQ, followed by gender.

Study 5 aimed to establish whether gender identity concepts and self-constructs play a role in prediction of the intelligence type. Neither gender identity concepts nor self-construct contributed significantly to the prediction of DMIQ. Contrary to results of Study 3 and 4, but in accordance with the prediction, gender was the best and the only predictor of DMIQ. The results resembled the findings of Study 1 and 2.

The findings of Studies 3, 4, and 5 are represented in Figure 3.5.1. The single-pointed arrows symbolize a direct relationship between two variables. The dashed arrows (e.g. between '*g*' and DMIQ) represent relationships that were not predicted. The patterned arrows (e.g. between gender and DMIQ) represent relationships that were predicted but not observed. Brackets contain studies the results are referring to if non-uniform results were observed. Variables that exhibited a relationship with DMIQ and HHE are in bold. The direction of the arrows implies causality that is based on this chapter's results.

Figure 3.5.1: Pictorial representation of the findings of Studies 3, 4 and 5



Chapter 4: Gender, Gender Identity, Affect Measures and Self-Constructs as Predictors of Domain-Masculine Intelligence Type

4.1. General Introduction

Chapter 2 confirmed the existence of the Hubris-Humility Effect on the Domain-Masculine Intelligence Type (DMIQ) and demonstrated that gender was the best predictor of the intelligence type. Chapter 3 investigated the role general intelligence (‘*g*’) and gender play in the prediction of DMIQ. A number of possible DMIQ predictors, i.e. implicit Intelligence Beliefs, gender identity and self-constructs were introduced. None of the newly introduced variables contributed significantly to the prediction of the intelligence type. Equally, nearly all hypotheses concerning the novel variables were refuted. In the same way, the overall results were ambiguous in that ‘*g*’ was the best predictor of the intelligence type in the first two studies, but in the last study, that also included gender identity and self-constructs, gender was the only significant predictor of DMIQ.

Hence, the aim of Chapter 4 is to re-examine the role gender, gender identity variables and self-constructs play in the prediction of DMIQ. Equally, it aims to ascertain whether gender is the best predictor of the intelligence type. As in previous studies the occurrence of Hubris-Humility Effect on DMIQ will be investigated.

Study 6 introduces a two-dimensional model of affect or the experience of emotion, i.e. positive and negative affect (PA/NA) (e.g. Watson & Tellegen, 1985). The affect model has not been included in the SEI research programme previously despite a possible interface. Researchers demonstrated that PA/NA correlate with personality traits (Costa & McCrae, 1980; Watson et al., 1988). Similarly, SEI researchers (cf. Chamorro-Premuzic & Furnham, 2005; Chamorro-Premuzic, Furnham, & Moutafi, 2004; Furnham ,2005; Furnham & Chamorro-Premuzic, 2006) established a relationship between self-estimated intelligence model and personality traits. Thus, Study 6 sets to examine whether PA/NA constructs play a role in the prediction of DMIQ. Equally, as in Studies 4 and 5, gender identity variables, masculinity and femininity, will be included to determine whether they play a role in DMIQ. Based on the previous findings, gender is expected to remain the best predictor of DMIQ.

Study 7 will conclude by re-introducing self-esteem and self-control and together with masculinity, femininity and PA/NA, and gender seek to establish whether they are valid predictors of the intelligence type as well as to review the previous findings and the existing literature.

4.2. Study 6

Gender, Gender Identity Variables and Affect Measures as Predictors of Domain-Masculine Intelligence Type

4.2.1. Introduction

The aim of this study is to investigate the role gender, gender identity concepts, i.e. masculinity and femininity, and affect measures, i.e. positive and negative affect, play in the prediction of the Domain-Masculine Intelligence Type. Positive affect (PA) is the degree to which individuals feel active and enthusiastic, with high PA associated with energy and Extraversion and low PA with unhappiness and lethargy. Similarly, negative affect (NA) is the degree of general upset, with high NA associated with anger, fear and guilt and low NA with composure and calmness (Watson et al., 1988). In addition, low PA correlates with depression and NA with anxiety and Neuroticism (Costa & McCrae, 1980; Watson & Tellegen, 1980). Since, personality trait literature repeatedly documented that women score higher on Neuroticism (Costa & McCrae, 1980), it is unsurprising that women score higher on negative affect than men (Smith & Reise, 1998). To date, the self-estimated intelligence research programme have not included affect measures, despite a possible interface, as affect has been shown to correlate with personality traits (e.g. Watson et al., 1988) and impact (female) performance (Smith & Reise, 1998), whilst SEI has been shown to be strongly related to personality traits (e.g. Chamorro-Premuzic & Furnham, 2005). Thus, PA/NA are introduced to establish whether they play a role in the prediction of DMIQ as well as to examine the relationship between the construct and DMIQ.

In addition, masculinity and femininity are also included in this study to examine whether they play a role in the prediction of DMIQ. As discussed, gender stereotypes are thought to play a role in HHE (e.g. Furnham, 2001; Petrides, Furnham, & Martin, 2004) and were shown to be strongest in the areas that are stereotypically associated with masculinity and femininity, such as mathematics and arts (Brown & Josephs, 1999). Indeed, math ability stereotypes were shown to evoke underperformance in women as well as induce different performance anxiety in men and women (cf. Bonnot & Croizet, 2007; Rudman & Phelan, 2010). The fact that women are more susceptible to gender role stereotypical beliefs seems a plausible explanation for the existence of the Hubris-Humility Effect. Thus, the relationship between masculinity and femininity and DMIQ will be examined. Based on the existing literature, it seems likely that masculinity but not femininity will be a predictor of DMIQ.

Hypotheses are based on results of previous studies or literature findings. Thus, HHE is expected to occur on DMIQ (H1). Masculinity (H2) and positive affect (H3) are expected to be significantly correlated with DMIQ. This would support the finding that math-spatial abilities are perceived as male-normative or masculine and in turn, that PA is perceived as more ‘masculine’. Gender is expected to be the best predictor of DMIQ over and above gender identity concepts and affect measures (H4). Masculinity (H5) and PA (H6) are expected to be significant predictors of the intelligence type.

4.2.2. Method

Participants

A total of one hundred and forty-three participants took part in this study. There were 79 females (55%) and 64 males. Their age raged from 17 to 46 (M = 23.02, SD = 6.87) years. 52% of the participants reported their ethnicity as Caucasian, 33% as Asian and 2% as African. 96% reported BA/BSc level as their highest level of education. 81% of participants were single and 10% were married or were living with partner. 33% disclosed to be first born child, 43% the youngest, 11% were the middle child, and 14% the only child. 94% were right-handed and 6% were left-handed.

Measures

Domain-Masculine Intelligence Type (DMIQ)

See Study 3 (section 3.2.2). Alpha for Domain-Masculine Intelligence Type was .57 and the inter-item correlation r =.41.

Gender Identity

Bem Sex Role Inventory (BSRI) (Bem, 1981a).

See Study 5 (section 3.4.2.). The alphas for masculinity and femininity in this study were, .84 and .81, respectively.

Affect Measures

Positive Affect and Negative Affect Schedule (PANAS) (Watson, Clark & Tellegen, 1988).

This non-timed 20-item scale measures the independent constructs of positive and negative affect. Each construct is measured through 10 items that are alternated. Items were scored using a 5-point scale, where 1 = *very slightly or not at all* and 5 = *extremely*, e.g. afraid, nervous, irritable (NA) and proud, enthusiastic, inspired (PA).

The scale has shown a significant level of stability and no consistent sex differences have been reported (Watson, Clark & Tellegen, 1988). The scale's intercorrelations and internal consistency reliabilities are very high, ranging from 0.86 to 0.90 for positive affect and 0.84 - 0.87 for negative affect (Watson, Clark & Tellegen, 1988). Alphas for the positive and negative affect in this study were, .81 and .85 respectively. The inter-item correlations were .30 and .36 respectively.

Procedure

The participants were recruited through word of mouth and via a snow-balling effect, i.e. circular emails were sent to various contacts with request to share the survey link with their own contacts. Those who responded to the emails were sent a URL link and took the survey online (www.zoomerang.com). Debrief feedback was given at the end of the online survey, together with an option for personalised feedback report by the main researcher. All participants were fluent in English and no problems were reported via the feedback/comments box.

4.2.3. Results

4.2.3.1. Hubris and Humility Effect and the Domain-Masculine Intelligence Type

A samples independent t-test, t(141) = -5.81, p =. 00, two-tailed, confirmed significant differences between males (M = 116.82, SD = 10.68) and females (M = 106.92, SD = 9.66) in the Domain-Masculine Intelligence Type (DMIQ). The magnitude of the differences in the means (mean difference = -9.90, 95% CI:-13.27 to -6.53) was large ($\eta^2 = .19$; Hedge's Adjustment d =.97). Hypothesis 1 was confirmed.

4.2.3.2. Gender, Gender Identity Concepts and Affect Measures as Predictors of DMIQ

The relationship between DMIQ, gender, gender identity and affect measures as well as age was explored. The results of the correlational analysis are presented in Table 4.2.1. Gender correlated positively (r = .44, *p*=.00), with DMIQ, with males providing higher scores than males (M_{Male} = 116.82, SD_{Male} = 10.68; M_{Female} = 106.92, SD_{Female} = 9.66). Medium strength positive correlations were observed between the intelligence type and masculinity (r =.20, *p* <.05) and positive affect (r =.25, *p*<.01) respectively. Thus, hypotheses 2 and 3 were confirmed.

Table 4.2.1: Correlations, Means and Standard Deviations between DMIQ, Gender, Gender Identity, Affect Measures and Age

		DMIQ	G	M	F	PA	NA	A
	X	111.35	1.45	4.70	4.76	3.68	2.12	23.02
	(SD)	(11.24)	(.50)	(.70)	(.65)	(.58)	(.70)	(6.87)
Domain-masculine IQ	(DMIQ)							
Gender	(G)	.44***						
Masculinity	(M)	.20*	.49***					
Femininity	(F)	-.08	-.42***	-.34***				
Positive Affect	(PA)	.25**	.35***	.39***	-.09			
Negative Affect	(NA)	-.00	-.13	-.17*	.08	-.10		
Age	(A)	.09	.10	.27**	.04	.25**	-.04	

* *p* < .05 ** *p* < .01 *** *p* < .001 (2-tailed). *N between 136 and 143.*

In line with literature findings, men had significantly higher PA scores than females and masculinity was higher in males and femininity in females. Also, PA correlated positively with masculinity but not femininity. Despite the age range of participants (29 years) no significant relationship was observed between age and the intelligence type. An independent t-test for age was not significant; t(141) -1.14, *p*

=.25. No other significant relationships between DMIQ and the remaining variables were observed.

4.2.3.3. Gender as the best predictor of Domain-Masculine Intelligence Type

To further investigate whether the correlational patterns differed for males and females, the data was split per gender and correlations recomputed. Results are presented in Table 4.2.2. In the male sub-sample, no significant relationships were observed between DMIQ, masculinity, femininity and PA/NA. For females, a positive relationship occurred between DMIQ and positive affect (r =.23, p <.05). These results are interesting as females seem to believe that positive affect is necessary for high DMIQ estimates, but males still score higher on positive affect.

Table 4.2.2: Correlations, Means and Standard Deviations between DMIQ, Gender Identity, Affect Measures and Age – Per Gender

Variables	DMIQ Males	DMIQ Females
M	116.82	106.92
(SD)	(10.68)	(9.66)
n	64	79
Masculinity	-.22	.17
Femininity	.15	.09
Pos. Affect	-.09	.23*
Neg. Affect	.14	-.02
Age	-.03	.12

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

Table 4.2.3. shows the hierarchical regression results. Gender, gender identity and affect measures were the predictor variables and the DMIQ was the criterion variable. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. Stepwise method was used for each block. Gender (β = .43, p =.00, r_{part} = .43) was entered in Step 1, explaining 18% of the variance in DMIQ. When the gender identity variables, i.e. masculinity and femininity, were added at Step 2, gender explained 15% of

variance (β = .48, p =.00, r_{part} = .39) and was the only significant predictor. When positive and negative affect were added at Step 3, gender (β = .46, p =.00, r_{part} = .37) continued to be the only significant predictor of the intelligence type, explaining 14% of variance. The overall regression was significant, F(7,126) = 5.02, p = .00, f^2 =.28, with the overall model explaining 22% of total variance in DMIQ. In accordance with prediction, gender was the best predictor of DMIQ. Hypothesis 4 was confirmed and hypotheses 5 and 6 were refuted.

Thus, hypotheses 1, 2, 3 and 4 were confirmed and hypotheses 5 and 6 were not confirmed.

Table 4.2.3: Hierarchical Regression of Gender, Gender Identity and Affect Measures onto DMIQ

Regression Models	***Standardised β***	***Domain-Masculine IQ*** ***t***	r_{part}
Step 1:			
Gender	**.43**	5.48***	**.43**
Regression Model[1]		F(1, 132) = 30.03***	
R^2		.19	
R^2 Change		.19	
Adj. R^2		.18	
f^2		.23	
Step 2:			
Gender	**.48**	5.00***	**.39**
Masculinity	.00	.04	.00
Femininity	.12	1.30	.10
Regression Model[2]		F(3, 130) = 10.56***	
R^2		.20	
R^2 Change		.01	
Adj. R^2		.18	
f^2		.25	
Step 3:			
Gender	**.46**	4.66***	**.37**
Masculinity	-.04	-.42	-.03
Femininity	.09	1.01	.08
Positive Affect	.14	1.52	.12
Negative Affect	.08	.96	.08
Regression Model[3]		F(5,128) = 9.05***	
R^2		.22	
R^2 Change		.02	
Adj. R^2		.19	
f^2		.28	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). *Note:* Significant values are in bold.

4.2.4. Discussion

The aim of this study was to investigate the role affect measures and gender identity variables and gender play in the prediction of DMIQ. Additionally, gender was expected to be the best predictor of the intelligence type but masculinity and positive affect were also expected to significantly contribute in the prediction.

The first hypothesis set out to confirm the existence of HHE on DMIQ as was observed in all previous studies. The data supported the hypothesis ($\eta^2 = .19$, $d = .97$), providing further evidence that large gender differences occur on the numerical-spatial factor. As uniform findings were obtained in the six studies it could be argued that female underestimation of ability or '*humility*' is indeed a result of self-handicapping beliefs, negative self-perception biases, low self-confidence, or reaction to gender role stereotypes such as stereotype-threat (e.g. Betsworth, 1999; Beyer, 1990, 1998; Carr et al., 2008; Ehrlinger & Dunning, 2003; Steele & Aronson, 1995).

Measure of affect was included to test the premise that affect is a likely predictor for DMIQ given its relationship with personality traits, that in turn were shown to predict SEI (e.g. Chamorro-Premuzic & Furnham, 2005). Equally, gender identity variables were included since they correlate strongly with gender stereotypes, which are thought to play a role in HHE (e.g. Brown & Josephs, 1999).

Thus, masculinity and positive affect were expected to be significantly correlated with DMIQ. Results supported the hypothesis, providing support for the proposition that math-spatial abilities are perceived as male-normative or masculine. Likewise, positive affect correlated with DMIQ and with masculinity, confirming the hypothesis as well as previous claims that PA is associated with masculinity. To confirm previous results, gender was expected to be the best predictor of DMIQ over and above gender identity concepts and affect measures. The results confirmed that gender was the best predictor of the intelligence type, explaining 14% of variance. As in Study 5, gender was the only predictor of DMIQ, refuting the prediction that masculinity and PA play role in the prediction of DMIQ.

4.3. Study 7

Gender, Gender Identity, Affect and Self-Concept Measures as Predictors of Domain-Masculine Intelligence Type

4.3.1. Introduction

This study sets out to validate the findings of Study 6 as well as extend the previous findings and existing literature by re-introducing self-esteem and self-control. As discussed in Study 5, high self-esteem and self-control (Duckworth & Seligman, 2005, 2006; Kling et al., 1999) are likely to play a role in DMIQ.

Thus, it was hypothesised that the occurrence of Hubris-Humility Effect on the Domain-Masculine Intelligence Type would be validated in this sample (H1). Masculinity (H2), positive affect (H3), self-esteem (H4) and self-control (H5) are expected to correlate with DMIQ. The focus of this study is to examine what roles do gender, affect, gender role and self-constructs play in the prediction of DMIQ. In Study 5 and 6 gender was the best and only predictor of the intelligence type. Consequently, gender is expected to be the best predictor of DMIQ over and above gender identity concepts, affect measures and self-constructs (H6).

4.3.2. Method

Participants

A total of one hundred and thirty-nine participants from the general public took part in this study. There were 78 females (56%) and 61 males. Their age raged from 17 to 71 (M = 25.77, SD = 10.54) years. 89% were native English speakers and 87% reported their ethnicity as Caucasian. 67% of participants quoted A-levels as their highest educational qualification, 7% stated non-university education 9% stated

BA/BSc, 10% MA/MSc/MBA and 5% had completed education to the PhD level. 31% were married or with living with a partner, 21% were single, 25% divorced and were 14% widowed. 38% reported to be very religious, 10% reported to be only religious when surrounded by their family, 17% reported to be religious but not to actively practice, 11% stated to be hardly observant and 25% disclosed not to be religious. 33% were the first born child, 32% the youngest, 27% the middle child, 7% the only child and 1.4% had a twin. 82% were right-handed and 16% were left-handed; 2% were ambidextrous.

Measures

Domain-Masculine Intelligence Type (DMIQ)

See Study 3 (section 3.2.2). Alpha for Domain-Masculine Intelligence Type was .35 and the inter-item correlation r =.21.

Gender Identity

Bem Sex Role Inventory (BSRI) (Bem, 1981a).

See Study 5 (section 3.4.2). The alphas for masculinity and femininity in this study were, .81 and .77, respectively.

Affect Measures

Positive Affect and Negative Affect Schedule (PANAS) (Watson, Clark & Tellegen, 1988).

See Study 6 (section 4.2.2). Alphas for the positive and negative affect in this study were, .85 and .85 respectively. The inter-item correlations were .36 and .35.

Self-Constructs

Self-Esteem Scale (Rosenberg,1965).

See Study 5 (section 3.4.2) The alpha in this study was .75.

Brief Self-Control Scale (BSCS) (Tangney et. al, 2004).

See Study 5 (section 3.4.2). The alpha in this study was .81.

Procedure

Participants were recruited through an advertisement in The Jewish Chronicle, a weekly newspaper for the British Jewish community and a posting on the social networking site Facebook. Those who responded to the advertisement were sent either a paper version of the survey, together with a prepaid return envelope or an email with instructions and the survey in attachment. Detailed scoring instructions were given for each measure and debrief feedback was given at the end of the survey document.

As a gesture of appreciation for taking part in the study, a free personalised feedback report option was offered to the participants. The production of a detailed personalised feedback took the researcher about 45 minutes. Approximately 75% of participants have taken up the offer for free personalised feedback report. About 50% of those who responded to the initial advertisement subsequently referred their friends, family members and colleagues, substantially enlarging the total sample. All participants were fluent in English.

4.3.3. Results

4.3.3.1. Hubris and Humility Effect and the Domain-Masculine Intelligence Type

An samples independent t-test, t(137) = -5.51, p =. 00, two-tailed, confirmed significant differences between males (M = 114.20, SD = 11.73) and females (M = 103.24, SD = 11.57) in DMIQ. The magnitude of the differences in the means (mean difference = -10.96, 95% CI:-14.90 to -7.03) was large (η^2 = .18; Hedge's Adjustment d =.94). Hypothesis 1 was confirmed.

4.3.3.2. Gender, Gender Identity Concepts, Affect Measures and Self-Concept Constructs as Predictors of DMIQ

The relationship between DMIQ, gender, gender identity, affect measures and self-concept constructs as well as age was explored. Age was included in order to explore the reported assertions that it influences the observed gender differences in self-estimated intelligence (e.g. Beier & Ackerman, 2001, 2003; Deary et al., 2003; Lynn & Irwing, 2004; Rammstadt & Rammsayer, 2002b). In particular, age was partialled out to inspect its impact on DMIQ. The results of the correlational analysis are presented in Table 4.3.1.

Gender correlated positively (r = .43, p=.00), with DMIQ, with males providing higher scores than females (M_{Male} = 114.20, SD_{Male} = 11.73; M_{Female} = 103.24, SD_{Female} = 11.57). Medium strength positive correlations were observed between DMIQ and masculinity (r =.48, p =.00) and DMIQ and PA (r =.27, p<.01), confirming hypotheses 2 and 3. A negative relationship was observed between DMIQ and age (r = -.24, p <.01), with younger participants providing higher self-estimates. No other significant relationships between DMIQ and the remaining variables were observed. Thus, hypotheses 4 and 5 were not confirmed.

Given the age range of the participants (54 years) and the significant relationship between the intelligence type and age, partial correlational analysis was run with age partialled out. The results are presented in Table 4.3.1. Inspection of the two analyses revealed no significant differences in the significant relationship pattern as well as the observed values, with the following exceptions. When age was partialled out, the relationship between self-esteem and negative affect ceased to be significant. In addition, a small negative relationship was observed between self-control and negative affect (r =-.20, p<.05). An independent t-test for age was not significant; t(137) .66, p =.51.

Table 4.3.1: Correlations and Partial Correlations, Means and Standard Deviations between DMIQ, Gender, Gender Identity, Affect Measures, Self-Constructs and Age

		DMIQ	G	M	F	PA	NA	SE	SC	A
	X	108.05	1.44	4.54	4.80	3.49	2.13	2.29	3.04	25.77
	(SD)	(12.82)	(.50)	(.73)	(.65)	(.73)	(.73)	(.25)	(.46)	(10.54)
Domain-masculine IQ	(DMIQ)									
Gender	(G)	.43***								
Masculinity	(M)	.48***	.27**							
Femininity	(F)	-.11	-.33***	-.12						
Positive Affect	(PA)	.27**	.08	.37***	.28**					
Negative Affect	(NA)	.08	-.07	.05	.03	-.08				
Self-Esteem	(SE)	-.01	.05	-.17*	-.14	-.28**	.19*			
Self-Control	(SC)	.04	.05	-.09	.15	.05	-.14	.12		
Age	(A)	-.24**	.06	-.23**	.04	-.13	-.23**	-.15	-.22**	
Controlled For Age										
Domain-masculine IQ	(DMIQ)									
Gender	(G)	.43***								
Masculinity	(M)	.45***	.26**							
Femininity	(F)	-.10	-.33***	-.11						
Positive Affect	(PA)	.25**	.08	.35***	.29**					
Negative Affect	(NA)	.02	-.08	-.01	.04	-.12				
Self-Esteem	(SE)	-.05	.04	-.21*	-.13	-.31***	.16			
Self-Control	(SC)	-.01	.04	-.15	.16	.02	-.20*	.09		

* p < .05 ** p < .01 *** p < .001 (2-tailed). *N between 136 and 137.*

4.3.3.3. Gender as the best predictor of Domain-Masculine Intelligence Type

To further investigate whether the correlational patterns differed for males and females, the data was split per gender and the correlations recomputed. Results are presented in Table 4.3.2. In the male subsample, medium strength positive relationship was observed between DMIQ and masculinity (r =.44, p =.00) and between DMIQ and positive affect (r =.34, *p* <.01). These results corroborate previous findings in the field (e.g. Lippa, 2001; Smith & Reise, 1998).

Table 4.3.2: Correlations, Means and Standard Deviations between DMIQ, Gender Identity, Affect Measures, Self-Concept Constructs and Age – Per Gender

Variables	DMIQ Males	DMIQ Females
M	114.20	103.24
(SD)	(11.73)	(11.57)
n	61	78
Masculinity	.44***	.40***
Femininity	.09	-.01
Pos. Affect	.34**	.20
Neg. Affect	.16	.08
Self-Esteem	-.06	-.02
Self-Control	.15	-.11
Age	-.21	-.26*

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

For females, DMIQ was positively correlated with masculinity (r =.40, *p* =.00), which was an unexpected result. However, it validates the assertions that women are more susceptible to gender role stereotypes, especially about maths, spatial abilities and sciences (Eccles, 1987; Massa et al., 2005; Rudman & Phelan, 2010; Vispoel et al., 2000). In addition, it demonstrates that women perceive DMIQ as 'masculine' and the skills required for successful performance on DMIQ are also associated with masculinity. In addition, a negative relationship was observed between age and DMIQ (r =-.26, *p* <.05), revealing that younger women gave higher estimates.

Table 4.3.3. shows the results of the hierarchical regression. Gender, gender identity, affect measures, and self-constructs were the predictor variables and the DMIQ was the criterion variable. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. Stepwise method was used for each block. Gender (β = .43, *p* =.00, r_{part} = .43) was entered in Step 1, explaining 18% of the variance in DMIQ.

When the gender identity variables, i.e. masculinity and femininity, were added at Step 2, gender explained 10% of variance (β = .33, p =.00, r_{part} = .31). Masculinity was the other significant predictor (β = .40, p =.00, r_{part} = .38) accounting for 14% of explained variance.

Table 4.3.3: Hierarchical Regression of Gender, Gender Identity, Affect Measures and Self-Constructs onto DMIQ

Regression Models	**Domain-Masculine IQ**		
	Standardised β	***t***	r_{part}
Step 1:			
Gender	**.43**	5.51***	**.43**
Regression Model[1]		F(1, 137) = 30.37***	
R^2		.18	
R^2 Change		.18	
Adj. R^2		.18	
f^2		.22	
Step 2:			
Gender	**.33**	4.38***	**.31**
Masculinity	**.40**	5.39***	**.38**
Femininity	.05	.71	.05
Regression Model[2]		F(3, 135) = 21.92***	
R^2		.33	
R^2 Change		.15	
Adj. R^2		.31	
f^2		.49	
Step 3:			
Gender	**.33**	4.33***	**.30**
Masculinity	**.34**	4.30***	**.30**
Femininity	.01	.14	.01
Positive Affect	.12	1.44	.10
Negative Affect	.09	1.28	.09
Regression Model[3]		F(5,133) = 13.94***	
R^2		.34	
R^2 Change		.02	
Adj. R^2		.32	
f^2		.52	
Step 4:			
Gender	**.32**	4.12***	**.29**
Masculinity	**.36**	4.39***	**.31**
Femininity	.00	.05	.00
Positive Affect	.13	1.51	.11
Negative Affect	.09	1.22	.09
Self-Esteem	.04	.58	.04
Self-Control	.06	.80	.06
Regression Model[4]		F(7,131) = 10.06***	
R^2		.35	
R^2 Change		.01	
Adj. R^2		.32	
f^2		.52	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). *Note:* Significant values are in bold.

When positive and negative affect were added at Step 3, gender (β = .33, p =.00, r_{part} = .30) and masculinity (β = .34, p =.00, r_{part} = .30) were the only significant predictors of the intelligence type, each accounting for 9% of explained variance. When self-esteem and self-control were added at Step 4, the situation remained unchanged, with gender (β = .32, p =.00, r_{part} = .29) and masculinity (β = .36, p =.00, r_{part} = .31) the only significant predictors, explaining 8% and 10% of the variance in the intelligence type respectively.

The overall regression was significant, $F(7,131) = 10.06$, p = .00, f^2 =.52, with the overall model explaining 35% of total variance in DMIQ. Contrary to prediction, masculinity was the best predictor of DMIQ. Gender was the second best predictor. Hypothesis 6 was not confirmed. Thus, hypotheses 1, 2, and 3 were confirmed and hypotheses 4, 5 and 6 were not confirmed.

4.3.4. Discussion

Study 7 set out to validate the findings of Study 6 as well as previous studies and existing literature. Two self-concept measures, i.e. self-esteem and self-control were re-introduced and were expected to play a role in the prediction of DMIQ, together with affect and gender role measures.

The first hypothesis set out to confirm the existence of HHE on DMIQ as was observed in Studies 1 to 6. The data supported the hypothesis (η^2 = .18, d = .94), providing further evidence that large gender differences occur on the numerical-spatial factor. Thus, consistent findings about gender differences on the numerical-spatial factor of the multiple SEI model were obtained in all seven studies. As previously proposed, evidence seems to support the notion that the female underestimation of ability indeed results from self-handicapping beliefs, negative self-perception biases, low self-confidence, or reaction to gender role stereotypes such as stereotype-threat (e.g. Betsworth, 1999; Beyer, 1990, 1998; Carr et al., 2008; Ehrlinger & Dunning, 2003; Steele & Aronson, 1995).

Based on existing literature about affect, gender identity and self-constructs, positive affect, masculinity and self-esteem and self-control were expected to correlate with DMIQ. The results partially supported these claims, with only masculinity and PA significantly related to the intelligence type. These findings confirm previous findings and existing assertions in the field. Interesting results were obtained when the data was split per gender. Corroborating existing literature (e.g. Lippa, 2001; Smith & Reise, 1998), observed positive relationships between DMIQ and masculinity and PA but only for males. For females, masculinity also correlated positively with DMIQ, validating the assertion that women are more susceptible to gender role and domain-stereotypes, such as maths, spatial abilities and sciences (Eccles, 1987; Massa et al., 2005; Rudman & Phelan, 2010; Vispoel et al., 2000). Moreover, this data demonstrated that women perceive DMIQ as 'masculine'.

To confirm previous results, gender was expected to be the best predictor of DMIQ over and above gender identity and affect measures and self-concepts. The results failed to confirm the hypothesis. Masculinity was the best predictor of DMIQ, accounting for 10% of explained variance. Gender was the second best predictor with 8% of explained variance.

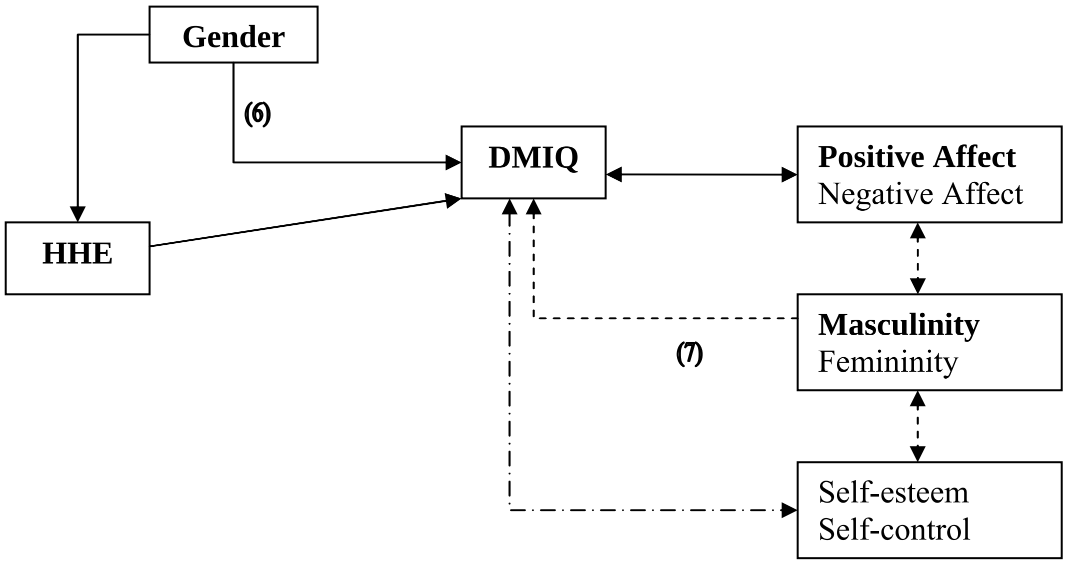
4.4. Summary

The two correlational studies reported in this chapter examined the role affect and gender identity measures, self-constructs and gender play in the Domain-Masculine Intelligence Type. In addition, both studies confirmed the existence of HHE on DMIQ, providing further evidence that gender differences on the mathematical/logical and spatial SEI are large indeed.

Study 6 revealed that gender was the best and only determinant of DMIQ, while in Study 7 masculinity, followed by gender was the best predictor of the intelligence type. Femininity, positive and negative affect or self-esteem and self-control did not significantly contribute to the prediction of DMIQ. These results are in agreement with study 5 that also revealed that self-concept and gender identity measures are not determinants of the intelligence type.

The findings of Studies 6 and 7 are represented in Figure 4.4.1. The single-pointed arrows symbolize the direct relationship between two variables. The dashed arrows (e.g. between masculinity and DMIQ) represent relationships that were not predicted. The patterned arrows (e.g. between self-esteem/self-control and DMIQ) represent relationships that were predicted but not observed. Brackets contain studies with results referring to when non-uniform results were observed. Variables that exhibited a relationship with DMIQ and HHE are in bold. The direction of the arrows implies causality that is based on this chapter's results.

Figure 4.4.1: Pictorial representation of the findings of Studies 6 and 7



Chapter 5: Domain-Masculine Intelligence Type and Culture

5.1. General Introduction

Chapter 2 validated the existence of the Hubris-Humility Effect on Domain-Masculine Intelligence Type, while Chapter 3 revealed that '*g*' as measured by Gf and Gc measures was the best predictor of DMIQ in two out of three studies. Despite the fact that gender was the best predictor of DMIQ in Study 5 and second best predictor in Studies 3 and 4. Chapter 4 focused on the role gender, gender identity variables, affect and self-concept measures plays in the prediction of DMIQ. Results revealed gender was the best predictor of the intelligence type in Study 6 while in Study 7 masculinity was the best predictor, followed by gender. None of the *'novel'* variables contributed significantly to the prediction of the type.

A plethora of studies in the SEI research programme (e.g. Furnham, Fong, & Martin, 1999; Furnham, Hosoe, & Tang, 2001; Furnham & Mottabu, 2004; Furnham, Rakow, et al., 1999; Furnham, Shahidi, & Baluch, 2002; Furnham, von Stumm, et al., 2009) were conducted about the role culture plays in the self-estimation process (see also Section 1.2.7 and Table 1.2.4. for more detailed overview), alongside studies about the cross-cultural understanding of intelligence (e.g. Favia & Fontane, 1997; Segall et al., 1999; Yang & Sternberg, 1997. The results of the cross-cultural studies of intelligence confirmed that cultures differ in some elements of their definitions of intelligence, although overall, peoples' understanding of intelligence is pan-cultural and resembles that of academics (e.g. Furnham & Baguma, 1999; Yang & Sternberg, 1997; Wober, 1973). Thus, intelligence should not be attempted to be understood outside a particular cultural context (Sternberg & Grigorenko, 2006).

Within the SEI cross-cultural programme that was conducted in more than 20 countries, clear and consistent cultural differences were found (e.g. Furnham & Akande, 2004; Furnham & Fong, 2000; Furnham & Fukumoto, 2008; Nasser & Singhal, 2006; Yuen & Furnham, 2005). In particular, Asians were most prone to humility, while Americans were most prone to display hubris during the self-estimation process, with Europeans in between the two. However, the previously observed SEI, generational and HHE estimation patterns were replicated across cultures, with males in particular providing significantly higher mathematical/logical and spatial intelligence estimates than did females (e.g. Furnham & Baguma, 1999; Furnham, Hosoe, & Tang, 2001; Yuen & Furnham, 2005; see also Table 1.2.4.).

These results are in stark contrast with the existing general intelligence (‘*g*’) literature that has shown an Asian ‘*g*’ advantage over Europeans and Americans (e.g.; Jackson & Rushton, 2006; Lynn et al., 2004). Furthermore, few African studies were exceptional in that HHE was either not observed or the gender differences were small. These results were attributed to diverse cultural norms and values (cf. Furnham, Rakow, & Mak, 2002; Hofstede, 1998, 2003; McSweeney, 2002).

Equally, one of the most used models in cross-cultural research is Hofstede’s model of cultural dimensions (1998, 2003, n.d.). It asserts the existence of four universal cultural dimensions: Power Distance, Uncertainty Avoidance, Individualism-Collectivism, and Masculinity-Femininity.

Power Distance (PDI) refers to the degree and acceptance of equality/inequality between people in a particular society. Thus, in high Power Distance society one’s social status and rightful place must be clear (De Mooij & Hofstede, 2010). The second dimension, Uncertainty Avoidance (UAI) refers to the degree to which people in a society feel threatened by uncertainty and ambiguity.

Countries high in Uncertainty Avoidance will rely on rules, regulations, controls, laws and expert opinions (Hofstede, n.d.). Individualism-Collectivism (IND) concerns the level to which a society reinforces individual or collective achievement and interpersonal relationships. Hence in highly individualistic cultures people are *"I-conscious"* and self-actualisation is important. In highly collectivistic cultures people are "*We-conscientious"* and their identity is determined by the social group they belong to (De Mooij & Hofstede, 2010; Hofstede, n.d.).

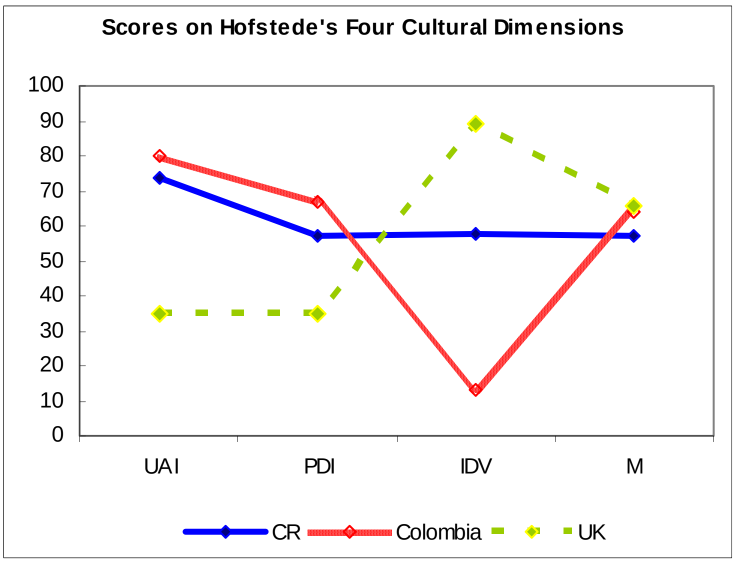
The fourth dimension, Masculinity-Femininity refers to the degree societies reinforce the traditional masculine work role model of male achievement, control, and power (Hofstede, 2003; n.d.). Societies high in Masculinity experience a higher degree of gender differentiation. In such cultures, males tend to dominate a significant portion of the society and power structure. An example of a high Masculinity country is Japan and of low Masculinity is Sweden.

The aim of this chapter is to investigate whether culture plays a role in the Domain-Masculine Intelligence Type and whether it affects the occurrence of HHE on the intelligence type in three distinct cultures, i.e. Czech Republic in Study 8, Colombia and the United Kingdom in Study 9. That is, will the gender differences, as observed in Studies 1 to 7 and previous SEI literature, be observed in these three countries? Although ample studies were conducted with British participants, no study was to date conducted in the Czech Republic or Colombia. Yet, Furnham & Chamorro-Premuzic (2005) confirmed the existence of HHE in an Argentinean sample, while Furnham, Rakow, et al. (1999) reported non-significant gender differences in a Slovakian sample and attributed this rare finding to years of Communist gender-equality norms and rearing practices. Thus, the Czech Republic and Colombia were chosen because both countries were presumed culturally similar to these studies, yet with differing masculinity scores.

Equally, Chapter 5 will examine the role gender identity variables, and in particular masculinity (Bem, 1974, 1981 a,b; Lippa, 2001), play in the prediction of the intelligence type in each culture. Existing literature classifies Colombia and the United Kingdom as high Masculinity countries, while the Czech Republic is categorized as average on this dimension (e.g. Hofstede, 1998, 2003, n.d.)

Figure 5.1.1. shows the scores on Hofstede's cultural dimensions for the three countries. Czech Republic data shows a very strong Uncertainty Avoidance (UAI score 74), an average Power Distance (PDI score 57), average Masculinity (M score 57), and average Individualism (IDV score 58). (Hofstede, n.d.).

Figure 5.1.1. Hofstede's Four Cultural Dimension Scores for the Czech Republic, Colombia and the United Kingdom



Colombia is typified by very high Uncertainty Avoidance (UAI score 80), high Power Distance (PDI score 67), high Masculinity (M score 64), and very low Individualism (IDV score 13) (Colombia, 2010). Whilst the scores are similar to other

South American countries, Colombia has the second highest masculinity rankings in South America (Hofstede, 2003, n.d.).

Hofstede's data for the United Kingdom show a very different cultural pattern. UK scores low on Uncertainty Avoidance (UAI score 35) and on Power Distance (PDI score 35) and very high on Individualism (IDV score 89). Surprisingly, it also scores highly on Masculinity (M score 66). Thus, each country has a distinct culture.

Consequently, Study 8 focuses on confirming the existence of HHE on DMIQ as well as ascertaining the best predictor of DMIQ. Furthermore, the Czech results are compared with the results of Studies 1 to 7 (all UK samples).

Study 9 sets to confirm the findings of Study 8 with two small sub-samples, i.e. Colombian and British.

5.2. Study 8

Hubris-Humility Effect and Domain-Masculine Intelligence Type in Czech Republic

5.2.1. Introduction

This study aims to confirm the existence of the Hubris-Humility Effect on the Domain-Masculine Intelligence Type with participants from the Czech Republic. Equally, the effect sizes for the Czech sample are expected to be smaller than those found in Studies 1 to 7 with various British populations. This expectation is based on existing literature and previous findings (e.g. Furnham, Rakow, et al., 1999; Hofstede, 2003, n.d.) that demonstrated distinct cultural differences and socio-political backgrounds between the Czech Republic and the United Kingdom as well as lack of HHE in a similar, i.e. Slovakian, culture. Thus, it was hypothesized that HHE will be observed on DMIQ (H1) but that the observed effect size will be smaller than in Studies 1-7 (H2).

Gender stereotypes are thought to a play role in HHE (e.g. Petrides et al., 2004) and were shown to be most pronounced in areas that are associated with *'masculine'* and *'feminine'* characteristics, such as math/sciences and arts (Brown & Josephs, 1999). These stereotypes were also exposed to negatively impact performance and ability perception in women on tasks that are perceived as masculine, such as math (cf. Dar-Nimrod, 2007; Kiefer & Sekaqueptewa, 2007; Rudman & Phelan, 2010; Rydell et al., 2010; Steele & Aronson, 1995). Accordingly, gender is expected to influence the relationship between masculinity and DMIQ (H3) and between femininity and the intelligence type (H4).

Studies 5 and 6 that also used gender identity variables established gender as the best predictor of DMIQ, while in Study 7 masculinity was the best predictor. Given that the Czech Republic is not a high masculinity culture (see Figure 5.1.1.) it seems probable that gender will account for most variance in the intelligence type. Consequently, gender is expected to be the best predictor of DMIQ over and above masculinity and femininity (H5).

5.2.2. Method

Participants

A total of one hundred and sixteen participants took part in this study. There were 85 females (73%) and 31 males. Their age raged from 17 to 50 (M = 30.83, SD = 8.19) years. 78% of the participants had achieved A-level or similar level of education, 4% reached non-university level of education and 17% had earned BA/BSc level of education. 41% of the participants were single, 41% were married, 11% were living with a partner, and 7% were divorced. 41% were the oldest, 32% the youngest, 9% were the middle child, 17% the only child and one participant had a twin sibling. 85% were right-handed, 10% were left-handed and 5% were ambidextrous. 61% indicated not to be religious and 18% said they were religious and 14% were undecided. A wide range of professions was observed from teachers (30%), to nursing students (27%), to civil servants (10%) to liberal professions (16%), police officers (6%), managers (3%), managing directors (2%), entrepreneurs (2%), secretaries (2%), and chefs (2%).

Measures

Domain-Masculine Intelligence Type (DMIQ)

See Study 3 (section 3.2.2). The alpha for Domain-Masculine Intelligence Type was .65 and the inter-item correlation for the two-item measure was = .48.

Gender Identity

Bem Sex Role Inventory (BSRI) (Bem, 1981a).

See Study 5 (section 3.4.2). The alphas for masculinity and femininity in this study were, .81 and .76, respectively.

Procedure

All participants were living in Prague, Czech Republic and Czech was their mother tongue. They were recruited by a local research assistant. The data was collected face-to-face by the research assistant who handed out hard copies of the survey questionnaire, together with translated Data Protection documents and consent forms. The participants were also given a brief description of all measures, with short feedback and background of the study. As a reward for participation, a detailed individual report (in English only) was offered. The questionnaires were then posted back to UCL, attention of the main researcher, where they were scored and entered into SPSS.

All documents were translated into Czech and back-translated into English by the main researcher and the local research assistant. The main researcher is fluent in Czech and the local research assistant was a native Czech speaker, fluent in English. No discrepancies were found. Prior to the main survey, the Czech questionnaire was given to five ‘control’ subjects, with no difficulties or discrepancies reported. No issues were found and hence the questionnaire was deemed ready for administration.

5.2.3. Results

5.2.3.1. Hubris-Humility Effect and the Domain-Masculine Intelligence Type

An independent t-test, t(114) = -3.05, p <. 01, two-tailed, confirmed significant differences between males (M = 107.66, SD = 10.61) and females (M = 100.75, SD = 10.89). The magnitude of the differences in the means (mean difference = -6.91, 95% CI:-11.41 to -2.42) was medium (η^2 = .08; Hodge's Adjustment d =.64). Hypothesis 1 was confirmed.

Table 5.2.1: Overview of Independent t-Tests and Effect Sizes for DMIQ – Studies 1 to 8

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI		Effect Size	
					L	U	*η²*	*d*
Study 1 UK	117.72 (13.72) 52	106.41 (11.01) 77	-5.18(127)***	-11.31	-15.64	-6.99	**.17**	**.90**
Study 2 UK	111.04 (9.22) 38	104.73 (9.06) 77	-3.49(113)**	-6.31	-9.89	-2.73	.10	.69
Study 3 UK	115.96 (17.10) 12	100.60 (11.97) 59	3.75(69)***	.15.36	7.19	23.52	**.17**	**1.05**
Study 4 UK	120.64 (14.34) 39	102.59 (11.45) 82	7.46(119)***	18.05	13.26	22.84	**.32**	**1.38**
Study 5 UK	120.17 (8.01) 23	106.67 (9.34) 79	-6.29(100)***	-13.50	-17.77	-9.24	**.28**	**1.54**
Study 6 UK	116.82 (10.68) 64	106.92 (9.66) 79	-5.81(141)***	-9.90	-13.27	-6.53	**.19**	**.97**
Study 7 UK	114.20 (11.73) 61	103.24 (11.57) 78	-5.51(137)***	-10.96	-14.90	-7.03	**.18**	**.94**
Study 8 CZ	107.66 (10.61) 31	100.75 (10.89) 85	-3.05(114)**	-6.91	-11.41	-2.42	.08	.64

* p < .05 ** p < .01 *** p < .001 (2-tailed). Note: d = Hedge's Adjustment d is adjusted for sample size and used in all studies. Studies 1 to 7 = UK population; Study 8 = Czech Republic. Large effect sizes are in bold.

Table 5.2.1. shows the results of the independent samples t-tests for Studies 1 to 8 for the Hubris-Humility Effect. Results confirm the existence of HHE on DMIQ in all studies, with medium to very large effect sizes. Studies 1 to 7 were conducted with various UK populations. Six out of seven studies reported large to very large effect sizes and only one study reported a medium effect size. The smallest observed effect size was found in the Czech Republic sample (η^2 =.08; Hedge's Adjustment *d* =.64), confirming hypothesis 2.

5.2.3.2. Impact of Gender, Masculinity and Femininity on the DMIQ

Masculinity was collapsed into categorical variable, with Group 1 containing subjects with lowest masculinity scores, Group 2 subjects with average masculinity and Group 3 subjects with highest masculinity scores. Identical analysis was carried out with Femininity. Results are presented in Table 5.2.2.

Table 5.2.2: Overview of Masculinity and Femininity Banded

	Masculinity	n
Group 1	<=4	39
Group 2	5	39
Group 3	6+	38
	Femininity	n
Group 1	<=5	39
Group 2	5-6	38
Group 3	6+	38

Note: Computed using Visual Bander technique (SPPS 13.0)

Two two-way between-groups analyses of variance were conducted to explore whether gender moderates the relationship between masculinity and the intelligence type as well as femininity. Results are presented in Table 5.2.3.

Table 5.2.3. Two 2-way ANOVAs (Masculinity and Gender and Femininity and Gender) on DMIQ

Variable	Tot '*g*' score	Mean Score (SD)			*F*-score		
		Total	Males	Females	M/F	Gender	M x G
Masculinity	G1 (L)	98.87 (9.52)	105.00 (10.61)	98.17 (9.29)	.51	5.09*	.61
	G2 (M)	104.14 (12.28)	105.91 (10.20)	103.45 (13.11)			
	G3 (H)	104.83 (10.94)	109.53 (11.19)	101.41 (9.62)			
Femininity	G1 (L)	103.03 (10.41)	106.43 (9.18)	101.12 (10.74)	.38	9.08**	.31
	G2 (M)	100.97 (12.91)	106.50 (14.63)	99.00 (11.91)			
	G3 (H)	103.58 (10.32)	112.08 (6.21)	101.98 (10.21)			

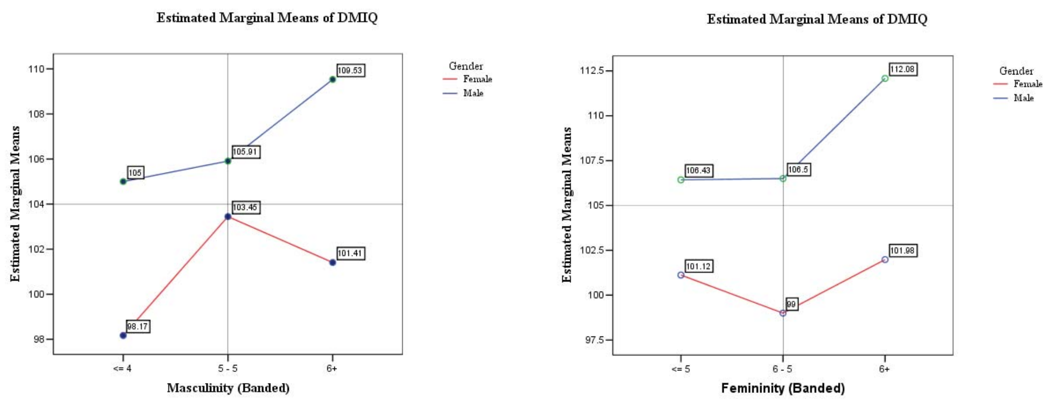
* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: DMIQ1 = Domain-Masculine Intelligence Type.

For masculinity, the interaction effect was not significant, F (2,110) = .61, p = .55, $\eta_p^2 = .01$. The main effect for masculinity, F (2,110) = .68, p =.51, $\eta_p^2 = .01$ was also not significant. The main effect for gender was significant, F (1,110) = 5.09, p <.05, $\eta_p^2 = .04$, with small effect size. Planned contrasts revealed no significant differences between the three groups. Post-hoc comparisons using the Tukey HSD test indicated that mean scores for Group 1 (<=4) significantly differed from mean scores for Group 3 (6+). However, the more stringent Bonferroni test revealed no significant differences in mean scores between three groups. Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 3 was partially confirmed.

For femininity, the interaction effect was not significant, F (2,109) = .31, p = .73, $\eta_p^2 = .01$. The main effect for femininity, F (2,109) = .97, p =.38, $\eta_p^2 = .02$ was also not significant. The main effect for gender was significant, F(1,109) = 9.86, p

<.01, η_p^2 = .08, with medium effect size. Planned contrasts revealed no significant differences between the three groups. Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated the mean scores of the three groups did not significantly differ. Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 4 was partially confirmed.

Figure 5.2.1: Two 2-way ANOVAs (M and Gender and F and Gender) on DMIQ



5.2.3.3. Gender and Gender Identity as Predictors of the Domain-Masculine Intelligence Type

The relationship between gender, gender identity variables and DMIQ was explored. As in Studies 6 and 7 age was also included to further examine its role in the SEI process (cf. Beier & Ackerman, 2001, 2003; Rammstedt & Rammsayer, 2002b). The results of the correlational and partial correlational analyses are presented in Table 5.2.4.

Gender correlated positively (r = .27, *p* <.01), with DMIQ, with males providing higher scores than females (M_{Male} = 107.66, SD_{Male} = 10.61; M_{Female} = 100.75, SD_{Female} = 10.89). Masculinity (r = .28, *p* <.01), but not femininity (r = -.05, *p*

=.61) correlated with DMIQ, which supports the assertion that DMIQ is perceived as male-normative and as such evokes masculinity gender-role stereotypical associations (e.g. Brown & Josephs, 1999; Furnham, 2001). A small positive correlation was observed between the intelligence type and age (r =.21, p <.05), revealing that older participants provided higher DMIQ estimates. Given the significant relationship between age and the intelligence type, correlations were re-computed, with aged being partialled out and are presented in Table 5.2.4. Examination of the partial correlational analysis revealed no significant differences in observed values from the previous analysis. An independent t-test for age was not significant; t(113) -1.02, p =.31.

Table 5.2.4: Correlations and Partial Correlations, Means and Standard Deviations between DMIQ, Gender and Gender Identity and Age

		DMIQ	G	M	F	A
	X	102.59	1.27	4.45	4.99	30.83
	(SD)	(11.20)	(.44)	(.76)	(.60)	(8.19)
Domain-masculine IQ T1	(DMIQ)					
Gender	(G)	.27**				
Masculinity	(M)	.28**	.34***			
Femininity	(F)	-.05	-.20*	.15		
Age	(A)	.21*	.10	.13	.03	
Controlled for Age						
Domain-masculine IQ T1	(DMIQ1)					
Gender	(G)	.26**				
Masculinity	(M)	.26**	.33***			
Femininity	(F)	-.06	-.20*	.14		

* p < .05 ** p < .01 *** p < .001 (2-tailed). *N between 115 and 116.*

To further investigate whether the correlational patterns differed for males and females, the data was split per gender and the correlations recomputed. Results are presented in Table 5.2.5. The only significant relationship was observed in the male subsample, between DMIQ and masculinity, (r = .38, p <.05), with highly masculine males providing higher DMIQ estimates. For females no significant relationships were observed. Again, the results corroborate the assertion that individuals, and especially females, perceive DMIQ as male-normative or domain-masculine.

Table 5.2.5: Correlations, Means and Standard Deviations between DMIQ, Gender Identity and Age – Per Gender

Variables	DMIQ Males	DMIQ Females
M	107.66	100.75
(SD)	(10.61)	(10.89)
Masculinity	.38*	.14
Femininity	.05	-.02
Age	.22	.18

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
N =between 30 and 85.

In order to test hypothesis 5 hierarchical regression was computed. Gender and gender identity were regressed on DMIQ to ascertain which variable is the best predictor of the intelligence type. Results are presented in Table 5.2.6. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. The stepwise method was used for each block.

Table 5.2.6: Hierarchical Regression of Gender and Gender Identity Variables onto DMIQ

Regression Models	***Standardised β***	***Domain-Masculine IQ*** ***t***	r_{part}
Step 1:			
Gender	**.27**	3.03**	**.27**
Regression Model[1]		F(1, 113) = 9.20**	
R^2		.08	
R^2 Change		.08	
Adj. R^2		.07	
f^2		.09	
Step 2:			
Gender	.19	1.94	.17
Masculinity	**.22**	2.26*	**.20**
Femininity	-.04	-.45	-.04
Regression Model[2]		F(3, 111) = 4.86**	
R^2		.12	
R^2 Change		.04	
Adj. R^2		.09	
f^2		.14	

* $p < .05$ ** $p < .01$ *** $p < .001$.   *Note:* Significant values are in bold.

Gender (β = .27, p <.01, r_{part} = .27) was entered in Step 1, explaining 7% of the variance in DMIQ. When the gender identity variables, i.e. masculinity and femininity, were added at Step 2, gender failed to reach significance. Masculinity (β = .22, p <.05, r_{part} = .20) was the only significant predictor, explaining 4% of variance. Femininity did not reach significance. The overall regression was significant, F (3, 111) = 4.86, p< .01, f^2 =.14, with the overall model explaining 12% of total variance in the intelligence type. Contrary to the hypothesis gender was not a significant predictor but masculinity was the best predictor of DMIQ.

Thus, hypotheses 1 and 2 were confirmed, hypotheses 3 and 4 were partially confirmed and hypothesis 5 was not confirmed.

5.2.4. Discussion

This study sat out to confirm the existence of HHE on DMIQ in a Czech sample. Based on results of a SEI study with a Slovakian sample (Furnham, Rakow, et al., 1999), a culture historically and socio-politically similar to the Czech Republic that did not replicate the existence of HHE, the effect sizes were expected to be smaller than those found in Studies 1 to 7. The results confirmed the existence of HHE on DMIQ (η^2 = .08, d = .64). However, the observed medium effect size was the smallest among the eight studies. The first seven studies were done with various British populations and the observed effect sizes were large and very large, with only one medium effect size. Hence, the results provided support for the existence of cultural disparity in gender differences in HHE and DMIQ between the Czech Republic and the United Kingdom as well as affirmed the uniqueness of Czech culture from the Slovakian culture.

Hypotheses 3 and 4 were concerned with gender's influence on the relationship between masculinity, femininity and DMIQ in order to provide further evidence for the finding that gender stereotypes are most pronounced in areas that are associated with *'masculinity'* and *'femininity'*. Moreover, gender stereotypes impact negatively on performance and ability perception in women, especially on *'masculine'* tasks (cf. Dar-Nimrod, 2007; Kiefer & Sekaqueptewa, 2007; Rudman & Phelan, 2010; Rydell et al., 2010; Steele & Aronson, 1995) and are assumed to play a role in HHE (Petrides et al., 2004). The results revealed significant gender effects for both masculinity and femininity but no significant interaction effect. Males provided higher DMIQ estimates on all three masculinity groups and surprisingly also on all three femininity groups. This finding suggests that Czech men either do not differentiate between the gender identity variables or view both as equally important. Equally, the results confirm Hofstede's claim (2003, n.d.) that the Czech Republic is an average masculine society as well as affirming the existence of *'hubris'* among Czech males.

Finally, as in all previous studies, this study aimed to confirm gender as the best predictor of the intelligence type. The correlational analysis validated the assertion that SEI, and in particular mathematical/logical and spatial intelligences, are perceived as male-normative (Furnham, 2001) as demonstrated through DMIQ's relationship with masculinity but not femininity. These results were confirmed when the sample was split by gender and the correlations recalculated, revealing a medium positive relationship between masculinity and DMIQ, but only for males. This implies that Czech men associate DMIQ with masculinity.

Since previous studies that used gender identity variables found conflicting results, i.e. Studies 5 and 6 established gender but Study 7 revealed masculinity as the best predictor, and given that the Czech Republic is an average masculinity country, gender was forecasted to be the best predictor of DMIQ. The results refuted this claim and confirmed masculinity as the best and only predictor, accounting for 4% of explained variance in DMIQ. Overall, the findings confirm the existing literature and previous studies in the area within a Czech Republic sample.

5.3. Study 9

Hubris and Humility Effect and the Domain-Masculine Intelligence Type in Dual Culture Study: Colombia and the UK.

5.3.1. Introduction

The aim of this study is to replicate the findings of Study 8 in a small dual-cultural sample. Although the existence of Hubris-Humility Effect was confirmed in another South American culture, i.e. Argentina (Furnham & Chamorro-Premuzic, 2005) and in nearly all studies with various British populations (cf. Furnham, 2001; Furnham, Clark, & Bailey, 1999; von Stumm et al., 2009), no other study investigated the existence of HHE on DMIQ in a Colombian and British sample. Thus, HHE is expected to occur in both cultures (H1).

According to Hofstede's cultural model (1998, 2003, n.d.) Colombia and the United Kingdom are divergent cultures (see Figure 5.1.1.). However, both countries score highly on Masculinity, with Colombia having the second highest national score among South American nations (e.g. Hofstede, 2003). Accordingly, gender is expected to influence the relationship between masculinity (Bem, 1981, a, b) and DMIQ in both cultures (H2). Given the results of Study 7 and 8 as well as the fact that both countries are highly 'masculine', it is expected that masculinity will be the best predictor of DMIQ in both cultures (H3).

5.3.2. Method

Participants

A total of one hundred and two participants took part in this study. There were 54 males (53%) and 48 females. Their age raged from 18 to 33 (M = 23.30, SD = 3.60) years. 52 participants (51%) were native English speakers and 50 were native Spanish speakers from Colombia. In the Colombian population (n =50), there were 28 males (56%) and 22 males, with their age ranging from 18 to 33 (M = 23.86, SD = 3.93) years. In the UK population (n =52), there were 26 males (50%) and 26 females, with their age ranging from 18 to 32 (M =22.77, SD = 3.20) years.

Measures

Domain-Masculine Intelligence Type (DMIQ)

See Study 3 (section 3.2.2). The alpha for Domain-Masculine Intelligence Type was .40 and the inter-item correlation for the two-item measure was = .25.

Gender Identity

Bem Sex Role Inventory (BSRI) (Bem, 1981a).

See Study 5 (section 3.4.2). The alphas for masculinity and femininity in this study were, .83 and .80, respectively.

Procedure

Participants in both countries were recruited through word of mouth among student populations and the general public. In the UK, the majority of the participants were associated with Goldsmiths College, University of London, and were recruited through a flyer posted in the Psychology Department. Colombian participants were recruited through a local research co-ordinator, who was a native Spanish speaker.

The data was collected face-to-face by the UK and Colombian research administrators, who handed out hard copies of the survey questionnaire, together with Data Protection documents. Participants were also given a brief description of all measures, with short feedback and background of the study. The questionnaires were then posted back to UCL, where they were scored, entered into SPSS and analysed. For the Colombian population, all documents were translated into Spanish and back-translated to English by the local Colombian research co-ordinator.

Prior to the main survey, the Spanish questionnaire was tested on a number of control subjects, with no difficulties or discrepancies reported. The pilot study indicated that it took approximately 30 minutes to complete the questionnaire. No issues were found, hence the questionnaire was deemed ready for administration. Contact details of the lead researcher were given in order to answer any questions or provide more information on the survey.

Participants were aware that they were free to withdraw their participation at any point or leave questions unanswered. No problems were reported during the testing sessions. The study has met the Ethics requirements of the Psychology Department and followed BPS ethical procedures, including seeking informed consent from all participants before undertaking part in the survey.

5.3.3. Results

5.3.3.1. Hubris-Humility Effect and the Domain-Masculine Intelligence Type

Independent samples t-tests were computed for each population. Results are presented in Table 5.3.1. Significant gender differences, with males providing higher self-estimates on DMIQ than females were observed in the Colombian and the UK samples. The observed effect sizes were large, with a larger ES for Colombia. Hypothesis 1 was confirmed.

Table 5.3.1: Independent Samples t-Tests and Effect Sizes for DMIQ – Colombia and the United Kingdom

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
Colombia	110.36 (10.93) 28	100.75 (9.43) 22	-3.27(48)**	-9.61	-15.51	-3.71	**.18**	**.94**
UK	114.37 (9.21) 26	105.50 (11.38) 26	-3.09(50)**	-8.87	-14.63	-3.10	**.16**	**.86**

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: *d* = Cohen's d. Large effect sizes are in bold.

5.3.3.2. Impact of Gender and Masculinity on the Domain-Masculine Intelligence Type

At the outset the dataset was split per nationality. In both samples, masculinity was collapsed into categorical variable, with Group 1 containing subjects with lowest masculinity scores, Group 2 subjects with average masculinity scores and Group 3 subjects with highest masculinity scores. Results are presented in Table 5.3.2.

Table 5.3.2: Overview of Masculinity Banded

	Masculinity	n
Colombia		
Group 1	<=4	17
Group 2	5	15
Group 3	6+	18
UK		
Group 1	<=4	19
Group 2	5	17
Group 3	6+	16

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender influences the relationship between masculinity and DMIQ in Colombia and the UK. Results are presented in Table 5.3.3.

In the Colombian sample, the homogeneity of variance assumption was violated (Levene Statistic $p < .05$), indicating the groups variances were not equal. An alternative check for comparing variances was used. Firstly, the largest and the smallest standard deviations were squared. The largest squared SD was divided by the smallest squared SD, with a resulting value of 1.43, which is smaller than the recommended value of 2, suggesting that the group variances, albeit not equal, were tolerable. Subsequently, the significance level was adjusted to p <.01.

The interaction effect between gender and masculinity was not significant, F (2,44) = .29, $p = .75$, $\eta_p^2 = .01$. The main effect for masculinity, F (2,44) = .18, $p > .50$, $\eta_p^2 = .10$ was not significant. The main effect for gender was also not significant, F(1,44) = 1.30, $p = .26$, $\eta_p^2 = .03$. Planned contrasts revealed no significant differences between the three groups. Post-hoc comparisons using the Games-Howell and Bonferroni tests revealed no significant differences in mean scores between the three groups. Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets.

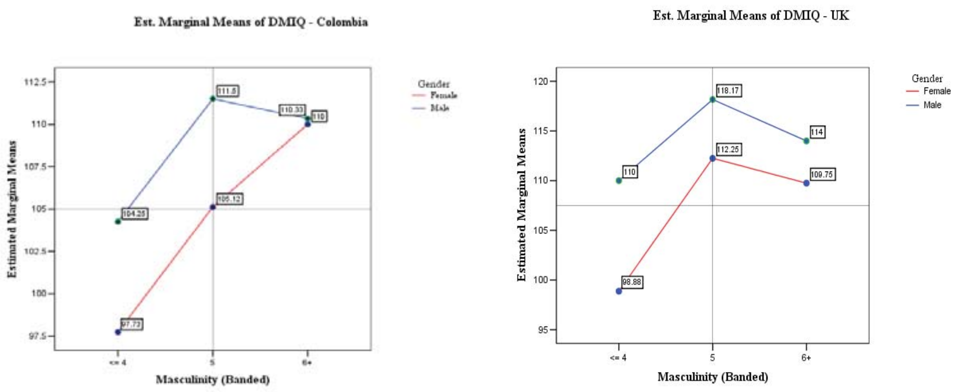
Table 5.3.3: 2-way ANOVA (Masculinity and Gender) on DMIQ – Colombia and the United Kingdom

Variable	Tot '*g*' score	Mean Score (SD)			*F*-score		
		Total	Males	Females	Masculinity	Gender	M x G
Colombia							
Masculinity	G1 (L)	98.50 (8.44)	104.25 (15.20)	97.73 (7.72)	.18	1.30	.29
	G2 (M)	109.80 (13.87)	111.50 (14.31)	105.13 (13.23)			
	G3 (H)	110.28 (7.27)	110.33 (7.84)	110.00 (4.33)			
UK							
Masculinity	G1 (L)	102.97 (10.42)	110.00 (9.13)	98.88 (9.09)	5.92**	6.99*	.61
	G2 (M)	115.38 (9.78)	118.17 (10.40)	112.25 (8.59)			
	G3 (H)	112.41 (9.56)	114.00 (7.38)	109.75 (12.76)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

For the United Kingdom sample, the interaction effect between gender and masculinity was not significant, F (2,46) = .61, $p = .55$, $\eta_p^2 = .03$. The main effect for masculinity, F (2,46) = 5.92, $p < .01$, $\eta_p^2 = .21$ was significant, with large effect size. The main effect for gender was also significant, F(1,44) = 6.99, $p < .05$, $\eta_p^2 = .13$, with medium effect size. Planned contrasts revealed significant differences between Group 1 and Group 3 (Contrast Estimate -9.10, $p < .01$). Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that mean scores for Group 1 (<=4) differed significantly from mean scores for Group 2 (5) as well as Group 3 (6+). Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Thus, hypothesis 2 was partially confirmed.

Figure 5.3.1: Two 2-way ANOVAS (Masculinity and Gender) on DMIQ – Colombia & UK



5.3.3.3. Gender and Gender Identity Variables as Predictors of DMIQ in Colombia and the UK

The dataset was split per nationality before all analyses were computed in order to test the hypotheses. The relationship between DMIQ, gender and gender identity variables was explored. Given that age was shown to impact the SEI estimations (e.g. Rammstedt & Rammsayer, 2002b) and correlated with DMIQ (see Studies 6, 7 and 8) the variable was included in the analysis to consider whether it plays a role in this dual-culture sample. The results of the correlational and partial correlational analyses are presented in Table 5.3.4.

For the Colombian population, a medium positive correlation was observed between DMIQ and gender (r =.43, *p* <.01), with males providing higher scores than females (M_{Male} = 110.36, SD_{Male} = 10.93; M_{Female} = 100.75, SD_{Female} = 9.43). Medium positive relationships were observed between DMIQ and masculinity (r = .39, *p* <.01) and between DMIQ and age (r = .29, *p* <.05), with older Colombian participants providing higher DMIQ estimates. This finding validates the findings of Study 8. A medium negative relationship was observed between the intelligence type and femininity (r = -.29, *p* <.05).

Given the significant relationship between age and DMIQ, the correlational analysis was recomputed, with age partialled out. An inspection of the partial correlational matrix revealed no significant differences in the correlational pattern from the initial analysis. However, an independent samples t-test for age was significant; t(48) -2.26, p <.05; M_{Male} = 24.93, SD_{Male} = 3.90; M_{Female} = 22.50, SD_{Female} = 3.62, with older Colombian participants being male. The magnitude of the differences in the means (mean difference = -2.43, 95% CI:-4.59 to -.26) was medium (η^2 = .10; Cohen's *d* =.65). It should be noted that the very small sample size (N=50) is likely to have influenced the results.

Table 5.3.4: Correlations and Partial Correlations, Means and Standard Deviations between DMIQ, Gender, Gender Identity, and Age – Colombia (n =50) and the UK (n =52)

UK / Colombia		DMIQ	G	M	F	A
UK	*X*	109.93	1.50	4.67	4.59	22.77
	(SD)	(11.19)	(.51)	(.76)	(.68)	(3.20)
Colombia	*X*	106.13	1.50	4.82	4.78	23.86
	(SD)	(11.28)	(.51)	(.73)	(.68)	(3.93)
Domain-masculine IQ	(DMIQ)		.40**	.45**	.05	.34*
Gender	(G)	.43**		.30*	-.32*	.54***
Masculinity	(M)	.39**	.63***		-.21	.22
Femininity	(F)	-.29*	-.43**	-.18		-.23
Age	(A)	.29*	.31*	.37**	.07	
Controlled for Age UK / Colombia						
Domain-masculine IQ	(DMIQ1)		.27	.41**	.14	
Gender	(G)	.37**		.22	-.24	
Masculinity	(M)	.32*	.58***		-.16	
Femininity	(F)	-.32*	-.48**	-.22		

* *p* < .05 ** *p* < .01 *** *p* < .001 (2-tailed).

For the United Kingdom population, a medium positive correlation was observed between DMIQ and gender (r =.40, *p* <.01), with males providing higher scores than females (M_{Male} = 114.37, SD_{Male} = 9.21; M_{Female} = 105.50, SD_{Female} = 11.38). Medium positive relationships were observed between DMIQ and masculinity (r = .45, p <.01) and between DMIQ and age (r = .34, p <.05), with older British participants providing higher DMIQ estimates. Again, these results replicate the findings of Study 8. No other significant relationships were observed.

Given the significant relationship between age and DMIQ, the correlational analysis was recomputed, with aged partialled out. An inspection of the partial correlational matrix revealed three significant differences in the correlational pattern. When age was controlled for, gender no longer correlated with DMIQ. Likewise, the previously significant relationships between masculinity, femininity and gender lost significance. An independent t-test for age was significant; t(50) -4.47, p =.00; M_{Male} = 24.46, SD_{Male} = 2.87; M_{Female} = 21.08, SD_{Female} = 2.58, with older British participants being male. The magnitude of the differences in the means (mean difference = -3.39, 95% CI:-4.91 to -1.86) was large (η^2 = .29; Cohen's d =1.24). As in the Colombian sample, the size of the UK sample (N=52) is likely to have influenced the results. Overall the results imply that age influenced DMIQ estimates in both cultures.

5.3.3.4. Gender as the best predictor of DMIQ

To further investigate whether the correlational patterns differed for males and females, the data was split per gender and the correlations recomputed. Results are presented in Table 5.3.5. For Colombia, no significant relationships were observed. In the British sample, the only significant relationship was observed between DMIQ and masculinity (r =.47, p <.05) but only for females. Although an unexpected finding, it confirms female susceptibility to gender role stereotypes that appear to be the strongest in areas perceived as *'masculine'*, such as maths, spatial abilities and sciences (Eccles, 1987; Massa et al., 2005; Rudman & Phelan, 2010; Vispoel et al., 2000). At the same time, the results confirm that females associate DMIQ with 'masculine' qualities. Interestingly, the same results were obtained in Study 7, also with a British population.

Table 5.3.5: Correlations, Means and Standard Deviations between DMIQ, Gender Identity and Age – Per Gender and Nationality

	Colombia		United Kingdom	
Variables	DMIQ Males	DMIQ Females	DMIQ Males	DMIQ Females
M	110.36	100.75	114.37	105.50
(SD)	(10.93)	(9.43)	(9.21)	(11.38)
n	28	22	26	26
Masculinity	.03	.34	.22	.47*
Femininity	-.19	-.00	.33	.06
Age	.16	.21	.08	.25

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

In order to test hypothesis 3, hierarchical regression was computed with the Colombian population. Results are presented in Table 5.3.6. Gender and gender identity were regressed on DMIQ to ascertain whether masculinity was the best predictor. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. Stepwise method was used for each block.

Gender ($\beta = .43$, $p < .01$, $r_{part} = .43$) was entered in Step 1, explaining 19% of variance in domain-masculine intelligence. When gender identity variables were added at Step 2, gender failed to reach significance but neither masculinity nor femininity did reach significance. The overall regression was significant, $F(3,45) = 4.13$, $p < .01$, $f^2 = .28$, with the overall model explaining 22% of total variance in DMIQ. Thus, hypothesis 3 was not confirmed in the Colombian sample.

Table 5.3.6: Hierarchical Regression of Gender and Gender Identity Constructs onto DMIQ – Colombian Sample (n=50)

Regression Models	***Standardised β***	***Domain-Masculine IQ*** *t*	r_{part}
Step 1:			
Gender	**.43**	3.24**	**.43**
Regression Model[1]		F(1, 47) = 10.49**	
R^2		.18	
R^2 Change		.18	
Adj. R^2		.17	
f^2		.22	
Step 2:			
Gender	.23	1.21	.16
Masculinity	.22	1.28	.17
Femininity	-.15	-1.02	-.13
Regression Model²		F(3, 45) = 4.13**	
R^2		.22	
R^2 Change		.04	
Adj. R^2		.17	
f^2		.28	

* $p < .05$ ** $p < .01$ *** $p < .001$.   *Note:* Significant values are in bold.

Table 5.3.7. shows the hierarchical regression results for the British population. Gender and gender identity were regressed on DMIQ to ascertain whether masculinity was the best predictor. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. Stepwise method was used for each block.

Gender ($\beta = .40$, $p <.01$, $r_{part} = .40$) was entered in Step 1, explaining 16% of variance in DMIQ. When masculinity and femininity were added at Step 2, gender ($\beta = .36$, $p <.01$, $r_{part} = .33$) explained 11% of variance. As predicted, Masculinity ($\beta = .39$, $p <.01$, $r_{part} = .37$) was also a significant predictor of the intelligence type. Masculinity explained 14% of variance in DMIQ and as such was its best predictor. Femininity did not significantly contribute to the prediction. The overall regression was significant, $F(3,48) = 7.98$, $p< .001$, $f^2 =.49$, with the overall model explaining 33% of total variance in DMIQ. Hence, hypothesis 3 was confirmed in the British sample.

Table 5.3.7: Hierarchical Regression of Gender and Gender Identity Constructs onto DMIQ – United Kingdom Sample (n =52)

Regression Models	***Standardised β***	***Domain-Masculine IQ*** *t*	r_{part}
Step 1:			
Gender	**.40**	3.09**	**.40**
Regression Model[1]		F(1, 50) = 9.53**	
R^2		.16	
R^2 Change		.16	
Adj. R^2		.14	
f^2		.19	
Step 2:			
Gender	**.36**	2.82**	**.33**
Masculinity	**.39**	3.16**	**.37**
Femininity	.24	1.93	-.23
Regression Model²		F(3, 48) = 7.98***	
R^2		.33	
R^2 Change		.17	
Adj. R^2		.29	
f^2		.49	

* $p < .05$ ** $p < .01$ *** $p < .001$. *Note:* Significant values are in bold.

Thus, hypothesis 1 was confirmed and hypotheses 2 and 3 were partially confirmed.

5.3.4. Discussion

Study 9 intended to validate the findings of Study 8 as well as confirm the previous literature findings. In addition, this study was unique in that it compared two distinctive cultures, Colombia and the United Kingdom. To date no SEI study was conducted with a Colombian sample.

The first hypothesis aimed to confirm the existence of HHE on DMIQ as was observed in Studies 1 to 8. The data supported the hypothesis for both cultures, with Colombia having a slightly larger effect size ($\eta^2 = .18$, $d = .94$) than the British sample ($\eta^2 = .16$, $d = .86$). The results confirm the claim that gender differences in SEI, and in particular on DMIQ, are universal and pan-cultural (cf. Furnham, 2001; von Stumm et al., 2009).

The second hypothesis, which expected gender to influence the relationship between masculinity and DMIQ in both cultures, was partially confirmed. No significant effects were observed in the Colombian sample. Nonetheless, the small sample size is likely to have influenced the results. For the British sample, a large significant masculinity effect and a medium gender effect were observed. The main interaction was not significant. The results have shown that individuals with the lowest masculinity provided lowest DMIQ estimates that differed significantly from the estimates of average and highest masculinity individuals. Unexpectedly, individuals with average masculinity provided the highest DMIQ estimates. The very same estimation pattern was observed for both genders, with average masculine males and females providing the highest DMIQ estimates. Furthermore, males had higher DMIQ estimates than females in all three masculinity groups, providing further support for male hubris in estimation.

Equally, correlational analyses revealed that masculinity correlated positively with DMIQ in both cultures, while femininity correlated negatively with DMIQ, but only in the Colombian sample. Moreover, age influenced DMIQ estimates in both samples, further confirming existing literature (e.g. Beier & Ackerman, 2001, 2002; Deary et al., 2003; Rammstedt & Rammsayer, 2002b). The results also revealed that British females, but not males, perceived DMIQ as masculine, replicating results of Study 7 and confirming the assertion of male-normativeness of intelligence (cf. Furnham, 2001).

Given that both cultures are highly Masculine (Hofstede, 1998, 2003, n.d.) masculinity was expected to be the best predictor of DMIQ, over and above gender and femininity. The results partially confirmed hypothesis 3, with masculinity as the best predictor of the intelligence type, but only in the British sample. Although the overall hierarchal regression was significant in the Colombian sample, no variable significantly contributed in the prediction of DMIQ. This finding is startling, given that Colombia is the second highest masculine culture in South America (Hofstede, 2003). Yet, the small sample sizes are likely to have influenced the results in both cultures.

5.4. Summary

The two correlational studies reported in this chapter set out to validate the existence of Hubris-Humility Effect on the Domain-Masculine Intelligence Type in three distinct cultures, the Czech Republic, Colombia and the United Kingdom and to provide further evidence for the existence of universal gender differences on the numerical-spatial factor of SEI (e.g. Furnham & Baguma, 1999; Furnham, Hosoe, & Tang, 2001; Yuen & Furnham, 2005; see also Table 1.2.4.). Both studies attempted to confirm whether gender influences the relationship between HHE and DMIQ. Finally, as in all previous studies, the best predictor of DMIQ was sought, with gender expected in the Czech Republic sample, and masculinity in the dual-cultural sample.

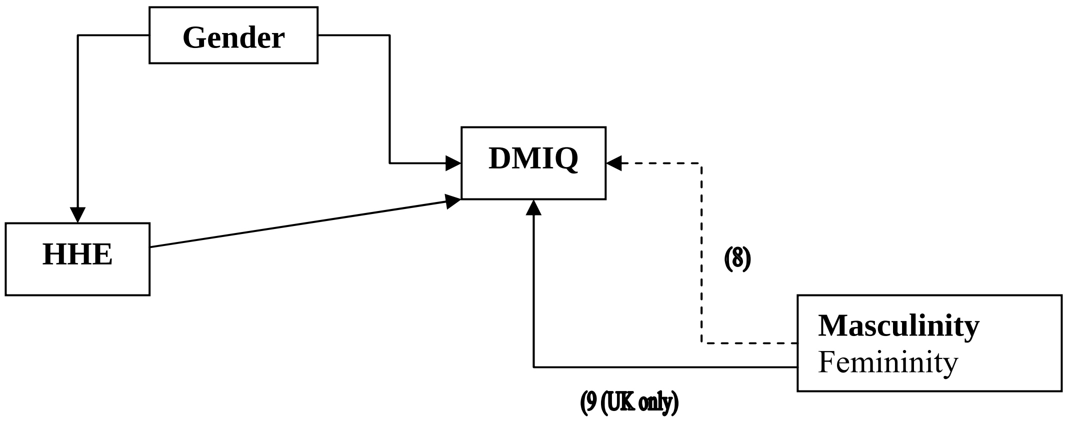
Study 8 confirmed the existence of HHE on DMIQ in the Czech sample. However, the medium observed effect size was the smallest among the eight reported studies. Gender significantly influenced the relationship between masculinity and DMIQ as well as between femininity and DMIQ. Czech males provided higher DMIQ estimates across all masculinity and femininity groups, which suggests that to them gender identity variables were evenly important. Contrary to prediction and literature in the field, masculinity was the best and only significant predictor of DMIQ.

Study 9 aimed to replicate findings of Study 8. However, given that both Colombia and United Kingdom are highly masculine countries, masculinity was expected to be the best predictor of DMIQ. As in Study 8, HHE existence on DMIQ was confirmed in both countries, providing additional evidence that gender differences in the numerical-spatial SEI factor are pan-cultural. Gender did not influence the relationship between masculinity and DMIQ in the Colombian sample but it did in the British sample. British males provided higher DMIQ estimates in all three masculinity groups, affirming the male '*hubris'* estimation pattern. Interestingly, British males and females with average masculinity scores provided the highest DMIQ estimates. These results provide support for the assertion that peoples' ability assessments are inaccurate and subject to cognitive biases and stereotypical beliefs (cf. Ehrlinger & Dunning, 2003; Kruger & Dunning, 1999). Equally, contrary to previous reports, highly masculine subjects do not automatically perceive DMIQ as 'masculine' (e.g. Chatard et al., 2007; Dar-Nimrod & Heine, 2006; Furnham, 2001; Kiefer & Sekaquaptewa, 2007). In accordance with prediction, masculinity was the best predictor of DMIQ, but only in the British sample.

Thus, Study 8 and 9 results confirm the assertion that intelligence can only be understood within a particular cultural context (Sternberg & Grigorenko, 2006).

The findings of Studies 8 and 9 are represented in Figure 5.4.1. The single-pointed arrows symbolize the direct relationship between the two variables. The dashed arrows (e.g. between masculinity and DMIQ in Study 8) represent relationships that were not predicted. Brackets contain studies the results are referring to in case non-uniform results were observed. Variables that exhibited a relationship with DMIQ and HHE are in bold. The direction of the arrows implies causality that is based on this chapter's findings.

Figure 5.4.1: Pictorial representation of the findings of Studies 8 and 9



Chapter 6: Domain-Masculine Intelligence Type and Precocity: Study with Mensa UK Members

6.1. Study 10

6.1.1. Introduction

While the previous chapters focused on confirming the existence of Hubris-Humility Effect in the Domain-Masculine Intelligence Type as well as establishing the best determinant of the intelligence type, Chapter 5 focused on corroborating these results across three distinct cultures.

This chapter focuses on validating the previous findings and SEI literature in a precocious population with the members of British Mensa. This is the first study of its kind within the SEI research programme. Mensa is an international non-political organisation for highly intelligent, gifted and talented individuals. Founded in Britain in 1946, it has more than 100,000 members in more than forty countries (Mensa UK, 2010). The society's objectives are to provide a stimulating intellectual and social environment for its members, to identify and foster human intelligence for the benefit of humanity, and to encourage research into the nature, characteristics, and uses of intelligence (Mensa UK, 2010). Membership is open to anyone who can demonstrate an IQ in the top two per cent of the population, measured by a recognised or approved IQ testing process, usually through Cattell's Culture Fair IQ Test (Mensa UK, 2010).

Studies comparing gifted/precocious and normal populations found little differences between the two groups. Similarly, gender differences in high ability groups mirror those found in normal populations (Roznowski, Reith, & Hong, 2000; Shea et al., 2001). However, the two groups differ in that precocious students display less gender stereotype beliefs and gifted girls' interests resemble those of normal males (Lubinski & Humphreys, 1990). Equally, gifted adolescents exhibit more desirable behaviours, such as work harder at school, participate in more preparatory courses, take more math/science courses, do more homework and get higher grades.

However, gender was shown to play a role in precocious groups. The most remarkable is the finding that gifted females tend to have lower educational and career expectations than gifted men, despite regularly outperforming gifted males (Lubinski & Humphreys, 1990; Roznowski et al., 2000). Parenthood, family-childcare needs, cyclical career path choices, and achievement conflicts influence life and career choices of profoundly gifted females (Benbow et al., 2000; Ferriman et al., 2009; Roznowski et al., 2000; Xie & Shauman, 2003).

Study 10 sets out to investigate whether gender differences in the numerical-spatial factor of SEI will be confirmed among Mensa UK members. Given the similarities between gifted and normal populations and the demonstrated *'humility'* among gifted females (e.g. Roznowski et al., 2000), it is predicted that HHE will prevail on DMIQ (H1). Mensa UK keeps its members abreast about diverse findings and developments in the intelligence research. Equally, it seems natural for highly gifted individuals to be more aware of their abilities and have a thorough understanding of expert and laymen views of intelligence. Likewise, previous research has shown that cultures do not differ in their understanding and beliefs about intelligence (e.g. Swami et al., 2008). This claim will be tested with the highly intelligent sample, using a questionnaire based on experts' opinions about intelligence, but in regards to gender differences. Based on the above, it is predicted that no significant gender differences will be observed in Beliefs about Intelligence among British Mensa members (H2). Moreover, gender identity variables are reintroduced to ascertain whether the previous findings about the observed relationship with DMIQ with normal populations will be replicated in the precocious sample. Thus, as in Studies 5, 6, 7, 8 and 9 a positive relationship between masculinity and DMIQ is expected to be observed (H3).

The relationship between gender, gender identity variables, Beliefs about Intelligence and DMIQ will be explored next. Based on literature about the role of age in SEI (e.g. Rammstedt & Rammsayer, 2002b) and on results of Study 3, 4, 7, 8 and 9, age is also included in the analysis to establish whether the previously observed age-DMIQ relationship will be replicated in this sample. Thus it is predicted that gender, age and Beliefs about Intelligence will be correlated with DMIQ (H4). In accordance with reported findings (e.g. Roznowski, Reith, & Hong, 2000; Shea et al., 2001) and results of Studies 1, 2, 5, and 6 gender is expected to be the best predictor of DMIQ over and above gender identity variables and Beliefs about Intelligence (H5).

6.1.2. Method

Participants

A total of two hundred and seventy-eight British Mensa members took part in this study. There were 143 males (51%) and 135 females. Their age ranged from 17 to 75 (M = 47.39, SD = 15.02) years. All participants were fluent in English, with 95% native English speakers. 95% of the participants reported to be White/Caucasian, 1.5% of Asian and 1.1% of African descent. 57% were married or living with partner, 27% were single, 11.5% divorced or separated, and 4% widowed. In all, 36.2% had completed non-university, higher-level education, 33.8% achieved BA/BSc level, 21.2% MA/MSc level and 5% achieved PhD/Doctorate or equivalent level of education.

48.9% reported to be in full-time employment, 8.6% in part-time, 12.2% were self-employed, 17.3% were retired, 6.5% were students, 2.2% homemakers, and 3.2% were unemployed. 13.7% disclosed to be working in education, 9.4% in computer/hardware/software/internet industry, 7.9% in finance/banking/insurance, 6.8% in healthcare/medical industries, 4.3% in consulting, 4.3% in military/for the UK government, 3.6% in accounting, 3.2% in non-profit organisations, 2.5% in engineering/architecture, 2.5% in manufacturing, 2.2% in legal industry, 1.8% aerospace/aviation/automotive industry, 1.8% in media/publishing/printing and 1.8% in retail/wholesale/trading industries. 38.5% were the oldest child in the family, 27.3% the youngest, 16.9% the only child and 9.4% the middle child. 41.8% reported to be religious, 41% reported not to be religious, whilst 17.2% were undecided. 69.8% were right-handed and 16.9% were left-handed; 12.2% were ambidextrous. 57.9% disclosed not to be religious, 24.5% reported to be religious/observant and 15.1% were undecided.

Measures

Domain-Masculine Intelligence Type (DMIQ)

See Study 3 (section 3.2.2). The alpha in this study was .51 and the inter-item correlation r = .35.

Intelligence Beliefs

Meaning and Measurement of Intelligence Questionnaire (Furnham, 2003).

This non-timed 30 item measure is designed to measure general public beliefs about intelligence. The questionnaire items were gathered from a summary of 50 (Western) psychologists and experts on intelligence research (reprinted in Gottfredson, 1997a). The summary was a response to an uproar over the publication of *The Bell Curve* (Herrnstein & Murray, 1994) in *Wall Street Journal* (15 December 1994) (Swami et al, 2008, p.238). The items concern, among other statements, what intelligence is, e.g. *Intelligence is a very general mental capability that involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly from experience; IQ scores predict equally accurately for all groups regardless of race and social class; Members of the same family also tend to differ substantially in intelligence for both genetic and environmental reasons.* The items are scored using an 8-point Liker scale, where 1 is *Strongly Disagree* and 8 is *Strongly Agree*. Previous research has shown good internal consistency, i.e. Cronbach's α = .81 (Swami et al, 2008). The alpha in this study was .81.

Gender Identity

Bem Sex Role Inventory (BSRI) (Bem, 1981a).

See Study 5 (section 3.4.2). The alphas for masculinity and femininity in this study were, .86 and .77, respectively.

Procedure

All participants were members of the British Mensa, who completed the survey either online or in a paper version that was sent to them with a pre-paid return envelope. Two hundred and seventy participants (97%) took the survey online. Eight Mensans – those without internet access, the most elderly and a handful from the Isle of Man, returned the paper questionnaires by post. Detailed scoring instructions were given at the beginning of each measure and the participants were aware that the study was approved by UCL Ethics Committee, meeting confidentiality and Data Protection requirements. Debrief feedback was available at the end of the survey questionnaire.

The main researcher contacted the co-ordinator of British Mensa, with a request for its members to participate in the study. After several telephone discussions, British Mensa agreed to participate in the study. An article on the hubris-humility effect and self-assessed intelligence, based on Prof. Dr. A. Furnham's research as well as detailed background information on the study was drafted and approved by Mensa's co-ordinator and press officer. The article, including participation encouragement by Mensa's co-ordinator, with the URL link (www.zoomerang.com) and contact details for the main researcher (for the paper version), was published in Mensa's monthly central newsletter (April/May 2007); shorter versions of the article were also published in the regional Mensa newsletters (May 2007).

The main researcher has received a number of letters from the participants, in direct response to their participation in the study (often in a reaction to the items/measures they have just completed), with personal anecdotes about what life is like for the highly intelligent as well as their thoughts on intelligence in general, and the study's elements. No problems were reported either during the testing session or received through the feedback/comments box in the online survey.

6.1.3. Results

6.1.3.1. Domain-masculine intelligence and the Hubris and Humility Effect

An independent t-test, $t(243) = 5.56$, $p =. 00$, two-tailed, confirmed significant differences between highly intelligent males (M = 143.92, *SD* = 12.53) and highly intelligent females (*M* = 134.43, *SD* = 14.58) in the Domain-Masculine Intelligence Type. The magnitude of the differences in the means (mean difference = 9.49, 95%

CI:6.13 to 12.85) was medium ($\eta^2 = .11$; Cohen's $d = .70$). Hypothesis 1 was confirmed.

6.1.3.2. Gender Differences in Beliefs about Intelligence

First, the 30-item Beliefs about Intelligence measure was analysed, using Principal Component Analysis, in order to identify the underlying structure of the measure. Prior to performing PCA, the suitability of data for factor analysis was assessed. Inspection of the correlation matrix revealed the presence of many coefficients of .3 and above. The Kaiser-Mayer-Olkin sampling measure value was .76 and exceeding the recommended value (Kaiser, 1970). The Bartlett's Test of Sphericity was significant, χ^2 (435) = 2643, $p = .00$, supporting the factorability of the correlation matrix (Pallant, 2005).

The initial solution was rotated, using the Direct Oblimin procedure and factors with eigenvalues greater than 1.00. Absolute values less than .40 were suppressed. PCA revealed eight components with eigenvalues over 1, accounting collectively for 60.51% of explained variance in the data. An inspection of the screeplot revealed a clear break after the seventh component. Results of Parallel Analysis (Monte Carlo PCA for Parallel Analysis, Watkins, 2000) further supported this finding, showing only seven components with eigenvalues exceeding the corresponding criterion values for a randomly generated data matrix of the same size (30 variables x 278 respondents).

Direct Oblimin procedure was repeated with seven extracted components, explaining a total of 57.05% of the variance. Table 3 shows the results of the PCA analysis, pattern and structure matrices with component loadings, percentages of explained variance, number of items and alpha levels per component. The Pattern

Matrix revealed a simple structure with strong loadings. The seven components were named to capture the common meaning among the items:

Table 6.1.1: Seven Components of Beliefs about Intelligence

Component	Component Name
C1	High IQ Advantage
C2	IQ Tests
C3	Intelligence As General Mental Ability
C4	IQ & Environment
C5	IQ Unchangeable
C6	IQ & Racial/Ethnic Differences
C7	IQ & Individual Differences

Interestingly, the components resemble the expert view of intelligence (e.g. Gottfredson, 1997a, b; 2000; Rushton & Jensen, 2005b) and affirm that precocious individuals have better understanding of intelligence than the normal population. Small to medium sized correlations were observed between the seven components. The seven components were used in the further analyses.

Table 6.1.2: Principal Component Analysis (Direct Oblimin with Kaiser Normalisation) of Intelligence Beliefs (N=278) 7 Components Extracted

Item	**Pattern Matrix Coefficients**							**Communalities**
	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6	Component 7	
Q15	**.777**	.022	.019	.046	-.065	-.075	.099	.630
Q18	**.759**	.164	.209	-.071	.036	.030	-.039	.638
Q14	**.748**	-.127	-.063	-.065	-.086	-.034	-.058	.585
Q16	**.543**	-.075	-.069	.032	.072	.192	.048	.422
Q20	**.517**	-.025	-.116	-.060	.006	.054	.065	.297
Q19	**.491**	-.027	.157	.182	.026	.039	-.050	.360
Q7	-.111	**-.801**	.127	.070	.159	-.085	-.206	.732
Q8	-.046	**-.797**	.125	.071	.135	-.069	-.243	.754
Q4	.176	**-.701**	-.003	-.125	-.110	.072	.268	.649
Q6	-.053	**-.659**	-.115	.070	-.091	-.033	-.044	.484
Q3	.192	**-.633**	.069	-.216	-.361	.007	.256	.724
Q17	-.281	.063	**.655**	-.059	-.055	.045	.131	.476
Q1	.048	-.076	**.635**	.017	-.222	.154	.186	.576
Q10	.081	.011	**.634**	.096	-.131	-.259	.086	.548
Q2	.235	.007	**.596**	-.083	.036	-.106	-.001	.448
Q5	.026	-.102	**.556**	.072	.039	.104	-.013	.364
Q9	.051	-.115	**.547**	.168	.409	.146	-.103	.567
Q25	-.089	-.012	-.075	**.767**	-.058	-.129	.289	.654
Q28	.062	-.048	.196	**.665**	.059	.072	-.189	.580
Q27	.462	.071	.129	**.479**	-.078	-.015	.033	.523
Q24	.188	-.138	.064	.162	**-.690**	.075	-.033	.639

Q30	-.011	.043	-.081	.069	**.658**	-.024	.394	.635
Q26	.001	-.140	-.045	.400	**-.440**	.241	-.013	.496
Q13	.309	-.119	-.065	.036	**.418**	.099	.113	.343
Q11	.016	.116	.110	-.157	-.056	**.877**	-.009	.775
Q12	.007	.062	.023	-.009	-.027	**.874**	.022	.764
Q29	.048	-.121	-.224	.386	.054	**.533**	.040	.577
Q21	.045	.058	.162	.009	.125	.104	**.715**	.642
Q22	-.091	.045	.038	.226	.059	-.066	**.712**	.569
Q23	.202	.085	.128	-.152	.119	.054	**.685**	.666

Item	**Structure Matrix Coefficients**							**Communalities**
	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6	Component 7	
Q15	**.781**	-.142	.158	.113	-.059	.122	.244	.630
Q18	**.751**	.005	.311	.000	.055	.187	.149	.638
Q14	**.729**	-.286	.049	.016	-.109	.142	.054	.585
Q16	**.606**	-.192	.031	.117	.064	.336	.156	.422
Q20	**.539**	-.175	.252	.256	.017	.181	.066	.297
Q19	**.523**	-.114	-.032	-.009	.006	.176	.152	.360
Q20	.085	**-.809**	.137	.203	.003	-.019	-.331	.732
Q8	.025	**-.790**	.132	.196	.028	-.047	-.306	.754
Q4	.380	**-.706**	.072	.031	-.198	.178	.212	.649
Q6	.058	**-.671**	-.088	.161	-.190	.015	-.150	.484
Q3	.367	**-.662**	.134	-.071	-.441	.098	.197	.724
Q1	.240	-.142	**.670**	.131	-.227	.194	.249	.476
Q10	.134	-.036	**.664**	.137	-.132	-.253	.133	.576
Q2	.295	-.042	**.623**	-.008	.031	-.055	.095	.548
Q17	-.159	.105	**.615**	-.014	-.043	-.020	.149	.448
Q5	.166	-.151	**.572**	.168	.019	.133	.047	.364
Q9	.193	-.145	**.567**	.271	.382	.189	-.013	.567
Q25	.001	-.087	.025	**.744**	-.043	-.027	.273	.654
Q28	.149	-.207	.264	**.705**	.040	.166	-.134	.580
Q27	.517	-.122	.257	**.525**	-.067	.156	.152	.523
Q24	.256	-.321	.122	.220	**-.712**	.148	-.035	.639
Q30	.043	.183	-.043	.061	**.687**	.012	.426	.635
Q26	.118	-.296	.011	.450	**-.459**	.302	-.028	.496
Q13	.372	-.120	.002	.096	**.407**	.198	.184	.343
Q12	.211	-.016	.031	.099	-.012	**.871**	.099	.775
Q11	.203	.050	.097	-.049	-.036	**.851**	.080	.764
Q29	.210	-.221	-.159	.457	.043	**.607**	.069	.577
Q21	.222	.135	.239	.057	.173	.169	**.763**	.642
Q23	.334	.158	.207	-.105	.169	.130	**.753**	.569
Q22	.048	.123	.117	.227	.106	-.006	**.709**	.666
Eigenvalue	5.13	3.40	2.35	1.86	1.64	1.42	1.04	
% of Explained Variance	17.11%	11.35%	7.83%	6.20%	5.48%	4.74%	3.45%	
No. of Items	6	5	6	3	4	3	3	
Alpha (α)	.75	.79	.70	.58	.59	.73	.71	
Inter-item *r*	.34	.43	.29	.31	.27	.46	.47	

Note: Major loadings for each item are bolded.

In order to test hypothesis 2 seven independent samples t-tests were computed. Results are presented in Table 6.1.3. Contrary to the prediction, significant gender differences were observed on two out of seven IQ components, i.e. on High IQ

Advantage (IQ1) and Unchangeable IQ (IQ5), with males believing significantly more than females that high IQ is an advantage and that one's intelligence cannot be changed. Yet, the observed effect sizes were small. These findings are at odds with Dweck's (1999) assertion that *'entity'* theorists, i.e. those who believe that intelligence can not be changed, perform worse than those who believe in malleability of intelligence. Equally, Dweck asserted that females are more likely to adhere to the entity attributional theory. It seems that precocious individuals' beliefs about intelligence resemble those of experts rather than laymen. Hypothesis 2 was not supported.

Table 6.1.3: Independent Samples t-tests and Effect Sizes for 7 Beliefs about Intelligence

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI		Effect Size	
					L	U	η^2	*d*
IQ 1	34.11 (7.19) 138	32.23 (6.74) 132	2.21(268)*	1.87	.20	3.55	.02	.27
IQ 2	22.92 (7.41) 142	22.98 (6.84) 133	-.08(273)	-.07	-1.77	1.63	.00	.01
IQ 3	39.63 (6.99) 142	40.62 (4.24) 133	-1.43(235)	-.99	-2.35	.37	.01	.17
IQ 4	17.18 (4.50) 140	17.75 (3.78) 133	-1.14(271)	-.57	-1.57	.42	.01	.14
IQ 5	13.83 (4.37) 138	12.82 (3.89) 132	2.01(268)*	1.02	.02	2.01	.02	.24
IQ 6	16.62 (4.00) 143	16.39 (3.91) 132	.49(273)	.24	-.70	1.18	.00	.06
IQ 7	15.35 (3.00) 142	15.12 (2.75) 134	.67(274)	.23	-.45	.92	.00	.08

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Hedge's Adjustment *d* is adjusted for sample size. Large effect sizes are in bold. IQ1 = High IQ advantage; IQ2 = IQ Tests; IQ3 = IQ as general mental ability; IQ4 = IQ & Environment; IQ5 = IQ Unchangeable; IQ6 = IQ & Racial/ethnic differences; IQ7 = IQ & Individual differences.

6.1.3.3. Gender, Gender Identity and Beliefs about Intelligence and the Relationship with DMIQ

The relationship between gender, gender identity variables, Beliefs about Intelligence and DMIQ was explored. Table 6.1.4. shows the results of the correlational analysis. Gender correlated (r = .33, p =.00), with DMIQ, with males providing higher scores than females (M_{Male} = 143.92, SD_{Male} = 12.53; M_{Female} = 134.43, SD_{Female} = 14.58). Masculinity (r = .26, p =.00), but not femininity (r = -.07, p =.29) correlated with DMIQ, confirming hypothesis 3. This is in line with the results of Study 5, 6, 7, 8, and 9 and confirms the finding that precocious individuals do not differ from the normal population in their gender role attitudes.

Out of the seven Beliefs about Intelligence components, only High IQ Advantage (IQ1) correlated (r = .19, p <.01) positively with DMIQ. High IQ factor (IQ1) resembles the Practical importance of intelligence factor (factor 2) reported by Swami et al. (2008).

Table 6.1.4: Correlations, Means and Standard Deviations between DMIQ, Gender, Gender Identity Variables, Beliefs about Intelligence and Age

		DMIQ	G	M	F	IQ1	IQ2	IQ3	IQ4	IQ5
	X	139.31	1.49	4.86	4.50	33.19	22.95	40.11	16.51	15.77
	(SD)	(14.35)	(.50)	(.80)	(.65)	(7.02)	(7.13)	(5.83)	(3.95)	(3.91)
Domain-masculine IQ	(DMIQ)									
Gender	(G)	-.33***								
Masculinity	(M)	.26***	-.22***							
Femininity	(F)	-.07	.21**	-.12						
IQB1	(IQ1)	.19**	-.13*	.22***	-.13*					
IQB2	(IQ2)	.04	.01	.10	-.01	.22***				
IQB3	(IQ3)	-.01	.09	.09	.14*	.25***	.14*			
IQB4	(IQ4)	.03	-.03	.00	.03	.28***	.21**	.30***		
IQB5	(IQ5)	-.08	-.09	.16	-.01	.08	.20	.17	.28*	
IQB6	(IQ6)	-.01	-.13*	-.02	-.12	.34***	.04	.03	.18**	.04
IQB7	(IQ7)	.02	.07	-.01	.08	.28***	-.18**	.26***	.16**	-.42***
Age	(A)	-.10	-.06	.01	.00	.22***	.07	.16**	.12	.13

* p < .05 ** p < .01 *** p < .001 (2-tailed).
N between255 and 278 for all variables, except IQ5 (n = 70-79).

Given the wide age range in the participants (58 years) and the previous results that confirmed a relationship between age and DMIQ (Study 3,4, 7, 8 and 9) as well as literature findings about its impact on SEI estimates (e.g. Rammstedt & Rammsayer, 2002b), age was included in the analysis. However, the relationship was not significant nor was an independent t-test for age. Thus, hypothesis 4 was only partially supported.

To further investigate whether the correlational patterns differed for males and females, the data was split per gender and the correlations recomputed. Results are presented in Table 6.1.5.

Table 6.1.5: Correlations, Means and Standard Deviations between DMIQ, Gender, Beliefs about Intelligence and Age – Split Per Gender

Variables	DMIQ Males	DMIQ Females
M	143.92	134.43
(SD)	(12.53)	(14.58)
Masculinity	.15	.25**
Femininity	-.07	.07
IQB1	.18	.13
IQB2	-.03	.12
IQB3	.09	-.12
IQB4	.02	.04
IQB5	-.27	-.01
IQB6	-.06	-.04
IQB7	.14	-.04
Age	-.22*	-.04

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
N =between 38 and 143.

For precious males, a negative relationship was observed between age and DMIQ (r = -.22, p <.05), with younger males providing higher DMIQ estimates. For precocious females, DMIQ correlated positively with masculinity (r =.25, p <.01), replicating results of Study 7 and 9 and supporting the view that gifted females beliefs and choices resemble those of males in normal populations (Lubinski & Humphreys, 1990).

6.1.3.4. Gender as the best predictor of DMIQ

In order to test hypothesis 5 and ascertain whether gender is the best predictor of DMIQ over and above gender identity variables and the seven Beliefs about Intelligence in the highly gifted population, hierarchical regression was computed. Results are presented in Table 6.1.6. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. Stepwise method was used for each block.

Gender (β = -.33, p <=.01, r_{part} = -.33) was entered in Step 1, explaining 11% of the variance in DMIQ. When the gender identity variables were added at Step 2, gender explained 8% of variance (β = -.29, p <.01, r_{part} = -.28). Neither masculinity nor femininity were significant predictors of the intelligence type. When the seven Beliefs about Intelligence factors were added in Step 3, gender (β = -.29, p <.01, r_{part} = -.28) remained the best and only significant predictor of the intelligence type, explaining 8% of variance. However, the overall regression was not significant, F(10,59) = 1.41, p = .20, f^2 =.23. As such hypothesis 5 was partially confirmed.

Thus, hypotheses 1 and 3 were supported and hypotheses 4 and 5 were partially supported, while hypothesis 2 was refuted.

Table 6.1.6: Hierarchical Regression of Gender, Gender Identity Variables and Beliefs about Intelligence onto DMIQ

Regression Models	***Standardised β***	***Domain-Masculine IQ*** **t**	r_{part}
Step 1:			
Gender	**-.33**	-2.89**	**-.33**
Regression Model[1]		F(1, 68) = 8.38**	
R^2		.11	
R^2 Change		.11	
Adj. R^2		.10	
f^2		.12	
Step 2:			
Gender	**-.29**	-2.46*	**-.28**
Masculinity	.19	1.69	.19
Femininity	.03	.13	.01
Regression Model[2]		F(3, 66) = 3.78*	
R^2		.15	
R^2 Change		.04	
Adj. R^2		.11	
f^2		.18	
Step 3:			
Gender	**-.29**	-2.37**	**-.28**
Masculinity	.19	1.53	.18
Femininity	.03	.24	.03
IQB1	.16	1.15	.14
IQB2	.00	-.00	.00
IQB3	-.01	-.06	-.01
IQB4	.06	.45	.05
IQB5	-.20	-1.33	-.16
IQB6	-.08	-.66	-.08
IQB7	-.08	-.52	-.06
Regression Model[3]		F(10, 59) = 1.41	
R^2		.19	
R^2 Change		.05	
Adj. R^2		.06	
f^2		.23	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). *Note:* Significant values are in bold.

6.1.4. Discussion

Study 10 set out to validate the existence of gender differences on the numerical-spatial factor of SEI in a sample of precocious individuals. This was the first time highly gifted participants were used in the SEI research programme. Hypothesis 1 was concerned, as in all previous studies, with validating the existence of HHE on DMIQ. The results confirmed that even precocious individuals, who are fully aware of their superior cognitive abilities, fall prey to gender stereotypical beliefs of hubris-humility ($\eta^2 = .11$, $d = .70$). These results affirm the existing suppositions about similarities between gifted and normal populations and the demonstrated *'humility'* among gifted females (e.g. Lubinski & Humphreys, 1990; Roznowski et al., 2000).

Hypothesis 2 set out to confirm that no gender differences will be observed among Mensa UK members in their understanding of meaning and measurement of intelligence, given the increased awareness of their abilities and appreciation of intelligence. Using a questionnaire that reflects experts' view of intelligence (Furnham, 2003; Swami et al., 2008), Beliefs about Intelligence, the results refuted Hypothesis 2. Gender differences were observed on two out of seven Beliefs about Intelligence factors, with precocious males, but not females, believing that having a high IQ is an advantage and that intelligence is not malleable. These results contradict the assertions of attributional theorists (cf. Dweck, 1999) and indicate that precocious males' views of intelligence resemble those of experts (e.g. Gottfredson, 1997a, b, 2000).

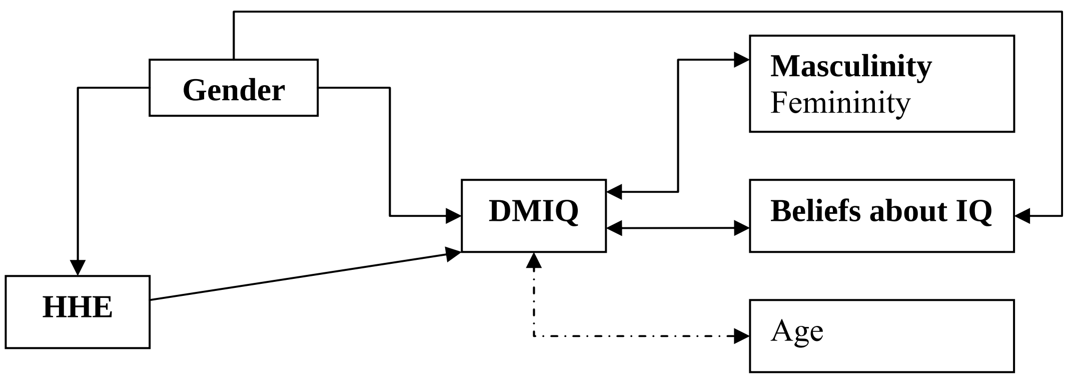
Hypotheses 3 and 4 were concerned with replicating previous results about the relationship between masculinity and DMIQ and the relationship between gender, age and Beliefs about Intelligence and DMIQ. The results confirmed that precocious individuals, alike normal populations, suppose a relationship between masculinity and DMIQ. Interestingly, but in agreement with Studies 7 and 9, gifted females but not males associate DMIQ with masculinity. Gender and Beliefs about Intelligence were related to DMIQ, but no relationship was observed between age and DMIQ. Among the Beliefs about Intelligence, only the first component, i.e. High IQ Advantage correlated with DMIQ.

In accordance with results of Studies 1, 2, 5, and 6 and existing literature (e.g. Roznowski, Reith, & Hong, 2000; Shea et al., 2001) gender was expected to be the best predictor of DMIQ over and above gender identity variables and Beliefs about

Intelligence. The results confirmed gender as the best and only predictor of the intelligence type, accounting for 8% of explained variance. However, the overall analysis was not significant.

The findings of Study 10 are represented in Figure 6.1.1. The single-pointed arrows symbolize direct relationship between two variables. The patterned arrows (e.g. between age and DMIQ) represent a relationship that was predicted but not observed. Brackets contain studies the results are referring to in case non-uniform results were observed. Variables that exhibited a relationship with DMIQ and HHE are in bold. The direction of the arrows implies causality that is based on this chapter's results.

Figure 6.1.1: Pictorial representation of the findings of Study 10



Chapter 7: Hubris-Humility Effect and Domain-Masculine Intelligence Type Measured through Experimental Design

7. 1. General Introduction

The first part of this thesis, i.e. Studies 1 to 10, were correlational studies that were concerned with validating the existence of the Hubris-Humility Effect on the numerical-spatial factor of SEI or the Domain-Masculine Intelligence Type. Secondly, each chapter introduced potential new determinants of the intelligence type, such as '*g*', gender-identity and self-construct variables, and affect measures. Gender was expected to be the best predictor of DMIQ. In addition, Chapter 5 set to corroborate the above in three distinct cultures and Chapter 6 to ascertain these hypotheses in a precocious population. The existence of HHE on DMIQ was validated in all ten studies, with medium to very large effect sizes. Yet, contrary to predictions, gender was not uniformly confirmed as DMIQ's best predictor, with '*g*' and masculinity alternating being the best predictor in Studies 3, 4, 7, 8, and 9. Gender was the best predictor in Studies 1, 2, 5, 6, and 10.

The second part of this thesis, confined in Chapter 7, sets out to investigate whether repeated measurement of the intelligence type as well as inclusion of real-life psychometric problems (TCAP) and task-success probability probes (TSP) does: 1) impact the occurrence of HHE on DMIQ, 2) facilitate size reduction in HHE from the initial (T1) to the post-task (T2) estimation condition, 3) assist explanation of DMIQ's best predictor, 4) enable an in-depth investigation of gender's role in the relationships between DMIQ and TCAP and DMIQ and TSP, and 5) facilitate understanding of the role gender plays in TCAP and TSP. To that end, five experimental studies were designed and run between 2007 and 2010, using a specialised online survey engine.

Albeit similar in design and execution, each study was unique in the number and sort of the psychometric problems (TCAP) and the number of task-success probes (TSP) asked. All experiments adhered to the following design. At the onset of each experiment participants were asked to estimate their DMIQ. This was followed by a number of psychometric problems that were alternated with probes about subjects' confidence about their ability to successfully solve a similar but more difficult problem. The psychometric problems were of medium difficulty level and were based on real-life IQ test questions that are used in the graduate recruitment process by corporations. Participants were then asked to re-estimate their DMIQ.

Repeated measurement of DMIQ estimation is implemented in order to test whether HHE can be manipulated, reduced or possibly made more accurate following a gender-stereotype inducing task. The limited data that exists on the effect of repeated measures on behaviour and performance suggest that mood and confidence can be altered when subjects are required to undergo multiple measurements of (e.g. Bartsch & Nesselroade, 1973; Ryckman et al., 1971). Similarly, literature suggests that the ability to accurately estimate one's abilities is correlated with one's intelligence, with higher IQ leading to more accurate estimation (Davidson & Downing, 2000).

The gender-stereotype inducing task, i.e. numerical, reasoning and crystallised psychometric problems (TCAP) is included to examine the claims that individuals are likely to overestimate or inflate their ability or performance beliefs on easy tasks and underestimate their abilities on difficult tasks (cf. Alicke et al., 1995; Burson et al., 2006; Guenther & Alicke, 2010; Moore & Small, 2007). These cognitive biases thus appear to influence self-perceptions and the accuracy of judgements and performance (e.g. Ariely, 2008; Ehrlinger & Dunning, 2003; Kahneman & Tversky, 2000). Equally, self-estimates were shown to be moderated by self-enhancement motives (Guenther & Alicke, 2010), which could explain the hubris occurrence. As such these biases are very similar to the Hubris-Humility Effect.

Yet, other researchers (e.g. Ackerman et al., 2002; Chamorro-Premuzic et al., 2010; Hall & Carter, 1999; Swim, 1994) have demonstrated that individuals are capable of accurate self-estimates of ability and tend to be accurate about gender-stereotypes. In fact, the ability to accurately assess gender differences was demonstrated to be an individual difference (Hall & Carter, 1999, p.350).

Thus, given the conflicting evidence and claims in the literature, the psychometric task (TCAP) is included to examine whether it will have impact on the DMIQ estimation process from T1 to T2, and whether it will affect the HHE. Likewise, the inclusion of TCAP should satisfy critics of the SEI research that claim it is subjective and of limited validity due to the exclusion of objective measures (Johnson & Bouchard, 2007).

Moreover, gender differences in mathematics achievement, attitudes and affect have been widely researched and documented (cf. Halpern et al., 2007). In general, females are shown to hold more negative or self-handicapping attitudes towards math, have lower math ability self-confidence, tend to stereotype math as domain-masculine, perform worse on standardised math tests, and opt out of STEM careers (Crombie et al., 2005; Beyer, 1990, 1998; Hyde et al., 1990a,b; Linn & Hyde, 1989; Meelissen & Luyten, 2008; Sax & Harper, 2007; The College Board, 1998). Males also associate math with domain-masculinity but are also more self-confident and display more positive math attitudes (Meece et al., 2006; Meelissen & Luyten, 2008;

van der Sluis et al., 2010). In fact, self-confidence is shown as one of the key predictors of gender differences in math achievement, with females reporting lower self-confidence than males, despite no differences in the actual performance (e.g. Carr et al., 2008). Yet, despite the lack of self-confidence, female math skills estimations are more accurate than males' (Carr et al., 2008; Pallier, 2003). Thus, self-confidence about one's math/spatial or domain-masculine abilities is likely to also play a role in HHE. In order test this claim, the experiment procedure also includes task-success probability estimations or success confidence probes. The probes are asked after each block of the psychometric problems and thus varied per experiment.

It should be reiterated that this thesis does not aim to contribute to the discussion on sex differences in intelligence. However, sex differences in TCAP and gender differences in TSP will be assessed to explore whether such differences contribute to the over- and underestimation behaviour of males and females (HHE).

Although the SEI research programme has generated a notable number of studies, no experimental studies were conducted to date. Few studies within the SEI programme included psychometric measures (e.g. Furnham & Fong, 2000; Furnham & Mottabu, 2004; Furnham & Rawles, 1999; Holling & Preckel, 2005; Reilly & Mulhern, 1995) and those were included to investigate accuracy of subjective, i.e. SEI, estimates. Furthermore, contrary to Studies 1 to 10 that used predominantly university students, the majority of participants in Studies 11 to 15 are from the general public, making the results more generalisable and robust.

Thus, the second part of this thesis reports the results of the five experimental studies. Consequently, Study 11 introduces the experimental design and aims to confirm the existence of HHE on DMIQ in the pre-task (T1) and post-task (T2) estimation conditions. It also aims at demonstrating a decline in DMIQ estimates from

T1 to T2 as a possible effect of the TCAP/TSP task. While ascertaining sex differences in cognitive abilities is not the aim of this thesis, an occurrence of sex differences in TCAP will be examined. Similarly, assertions of gender differences in confidence will be tested through investigation of TSP. Finally, the role gender plays in the relationship between TCAP, TSP and DMIQ will be examined.

Studies 12 to 15 set to validate the findings of Study 11, while the task content and format differs in each study. In addition, Study 14 data were collected separately in three different conditions. Study 14 reports the combined results for the overall group. The results for the three individual conditions are reported in the Appendix.

7.2. Study 11

Gender, Task Success Probability Estimation (TSP) and Total Correct Aptitude Problems (TCAP) as Predictors of the Domain-Masculine Intelligence Type

7.2.1. Introduction

This is the first of five experimental studies that sets out to corroborate the existence of HHE on DMIQ at both pre- and post-task estimation conditions. Repeated measures are included to validate assertions that they influence behaviour and performance and as such change mood and confidence (Bartsch & Nesselroade, 1973; Ryckman et al., 1971). Gender-stereotypes and self-confidence are likely to play a role in HHE or the display of male hubris and female humility in estimation of abilities. Therefore, subjects were asked to undertake a gender-stereotype inducing task, i.e. numerical and reasoning aptitude problems that are likely to increase hubris and humility (cf. Betsworth, 1999; Beyer, 1990, 1998; Dar-Nimrod, 2007; Ehrlinger & Dunning, 2003; Hoffman & Hurst, 1990; Steele & Aronson, 1995) as well as task-success estimates or confidence probes that will enable the assessment of confidence (cf. Burson et al., 2006; Carr et al., 2008; Dunning et al., 1990; Pallier, 2003). The task in Study 11 contains fifteen numerical and reasoning aptitude problems that were offered in five blocks of three problems. After each block, participants were asked to estimate their task-success confidence.

Thus, it was predicted that HHE will be confirmed on DMIQ at the pre-task (T1) and post-task (T2) estimating conditions (H1) and that there will be a significant decrease in DMIQ estimates from T1 to T2 following the gender-stereotype inducing task (H2).

Existing literature suggests that males have higher self-confidence, despite being inaccurate about their (math) skills or underperforming, whereas females are lacking confidence, while being accurate or outperforming males (e.g. Carr et al., 2008; Eccles-Parsons et al., 1984; Pallier, 2003). Consequently, males are expected to provide significantly higher task-success probability estimations (TSP) than females (H3).

Given the ample evidence about sex differences in cognitive abilities (cf. Halpern et al., 2007; Hyde, Fennema, & Lamon, 1990a; Jackson & Rushton, 2006; Lynn & Irwing, 2004; Ogle et al., 2003; Novell & Hedges, 1998; Voyer, Voyer, & Bryden, 1995, sex differences are expected on the numerical and reasoning problems (TCAP), with males providing more correct answers than females (H4).

As in Studies 1, 2, 5, 6, and 10, gender is expected to be the best predictor of DMIQ1 (H5) and DMIQ2 (H6) over and above TSP and TCAP. Finally, gender is presumed to influence the relationship between TSP and DMIQ1 (H7) and DMIQ2 (H8). Gender is also expected to affect the relationship between TCAP and DMIQ1 (H9) and DMIQ2 (H10).

7.2.2. Method

Participants

A total of four hundred and eighty-eight participants took part in this experimental online study. There were 326 females (67%) and 164 males. Their age ranged from 17 to 70 (M = 22.33, SD = 6.86) years. All participants were fluent in English and no language or other problems were reported. 50% had completed A-levels, 21%

achieved BA/BSc level, and 7% MA/MSc/MBA or equivalent level of education. As their favourite subject at school, 21% reported Languages/Literature, 14% reported Maths, 13.9% reported Biology and Sciences, 10.8% history, 8.5% Chemistry and Physics, 6.8% reported Psychology/Political Science/Sociology, 6.8% reported Arts/Drama/Photography/Theatre, 3.7% reported Music, 3% Computer Science/IT/IT Design, 2.8% Sports/PE, 2.6% Geography, 2% Philosophy and Religion, and reported 1.4% Accounting/Business Studies/Economics.

Measures

Repeated Measure of Domain-Masculine Intelligence Type (DMIQ).

Participants were asked to estimate their mathematical/logical and spatial intelligences on two occasions, prior (T1) and post (T2) completing a psychometric task (TCAP) and estimating their task-success confidence (TSP). As in the previous studies, individual scores for DMIQ were computed. Alpha for DMIQ1 was .82 and for DMIQ2 .88.

Psychometric Aptitude Task

Total Correct Aptitude Problems (TCAP)

Numerical and Reasoning Problems (Bryon, 2006)

Fifteen numerical and reasoning problems that were taken from an intelligence test training book were presented in five blocks of three analogous problems (Bryon, 2006). See Appendix for an overview of the problems. Participants were informed that items in each block varied in difficulty level, ranging from elementary to difficult. A time limit of 90 seconds was given for each block of problems. Participants were advised to leave unanswered problems blank, in order not to exceed the time limit, or be disqualified. The time limit was set to reflect a real-life intelligence testing situation, with the entire task taking 7.5 minutes to complete. Correct answers were available at the end of the survey. Alpha for the fifteen items was .93.

Task Success Probability

Task Success Probability Estimation Measure (TSP) (Storek, 2007)

After each problem block, participants were asked to indicate how likely they felt they would succeed on a similar task but with increased difficulty, e.g.*" Using the scale, indicate how likely you are to succeed on the same task, but with increased difficulty"* using a rating scale where 1 was *Very Unlikely* and 5 *Very Likely.* The five task success probability statements made up the Task Success Probability measure, with individual scores computed for all participants. The alpha for the five-item measure was .82. As such, the measure was a calibration measure of individual differences.

Procedure

Participants were second year medical students and financial industry professionals. Medical students were recruited through word of mouth by a research assistant and a fellow medical student, who was participating in a psychology course at UCL. The financial industry professionals were recruited via a number of financial industry acquaintances of the main researcher, who used to work in the financial industry.

An email invitation, with a URL link (www.zoomerang.com) to the study and a background explanation of the study, was sent by the main researcher and the research assistant to a wide audience of acquaintances, friends, (ex)-colleagues and their colleagues. The snow-balling technique of participant recruitment was used, i.e. participants were asked to forward the study invitation and the URL link to as many acquaintances and friends as possible. In total, 798 individuals logged onto the site during the period of June to November 2007. The data was gathered through an online survey engine www.Zoomerang.com and participation was voluntary.

Detailed scoring instructions were given at the beginning of each measure, including timing instructions for the psychometric problems. Participants were aware that the study was approved by UCL Ethics Committee, meeting confidentiality and Data Protection requirements. Debrief feedback, correct answers and opportunity to leave comments about the survey were provided at the end.

7.2.3. Results

7.2.3.1. Domain-masculine intelligence and the Hubris and Humility Effect in T1 and T2

In order to test hypothesis 1 whether HHE will occur in DMIQ1 and DMIQ2, two independent samples t-tests were computed. Results are presented in Table 7.2.1. Significant gender differences, with males providing higher DMIQ estimates, were observed in both estimation conditions. Hypothesis 1 was confirmed.

Table 7.2.1: Overview of Independent t-Tests and Effect Sizes for DMIQ1 and DMIQ2

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
DMIQ1	120.64 (18.13) 138	108.55 (18.70) 249	6.16(385)***	12.09	8.23	15.95	.09	.66
DMIQ2	116.02 (21.58) 92	102.57 (21.14) 137	4.68(227)***	13.56	7.79	19.12	.09	.63

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Hedge's Adjustment *d* is adjusted for sample size and used in both tests.

A paired samples t-test[8] was conducted to test whether DMIQ estimates decreased significantly from T1 to T2. There was a statistically significant decrease in DMIQ from T1 (M = 113.49, SD = 19.40) to T2 (M = 108.21, SD = 22.04), t(224) = 5.66, p = .00, two-tailed, r = .78, p =.00. The mean decrease in domain-masculine intelligence self-estimates was 5.28 (14.00) with 95% confidence interval ranging from 3.44 to 7.12. Cohen's *d* statistic (.38) indicated a small effect size. Hypothesis 2 was confirmed.

7.2.3.2. Gender Differences in Task Success Probability Estimation (TSP) and Psychometric Aptitude Task (TCAP)

Table 7.2.2. gives an overview of independent-samples t-tests and effect sizes for the five individual TSP probes and the overall TSP measure. With the exception of TSP4, the independent-samples t-tests were significant, with males providing higher TSP estimates than females. The observed effect sizes were small. Inspection of the correlational results (see Table 7.2.4.) revealed a negative correlation between gender and TSP (TSP) (r = -.18, p <.01), with males providing higher TSP estimates than females (M_{Males} = 3.18, SD_{Males} = .80; $M_{Females}$ = 2.88, $SD_{Females}$ = .81).Hypothesis 3 was confirmed.

Equally, inspection of the correlational results (see Table 7.2.4.) revealed a small negative correlation between gender and Total Correct Aptitude Problems

---

[8] Paired t-test is used when the samples are dependent, i.e. when there is only one sample that has been tested twice (repeated measures) or when there are two samples that have been matched or "paired". The appropriate equation is $t = \bar{X}_D - \mu 0 / s_D / \sqrt{n}$. The differences between all pairs must be calculated. The pairs are either one person's pre-test and post-test scores or between pairs of persons matched into meaningful groups. The average (X_D) and standard deviation (s_D) of those differences are used in the equation. The constant μ_0 is non-zero if one needs to test whether the average of the difference is significantly different from μ_0. The degree of freedom used is $n-1$ (Field, 2005; Wikipedia, Effect Sizes, November, 2010).

(TCAP), (r =-.18, p =.00), with males correctly solving more problems then females. An independent-samples t-test for TCAP revealed significant gender differences t(307) = 3.96, p =.00, two-tailed between males (M_{Males} = 5.47, SD_{Males} = 4.60) and females ($M_{Females}$ = 3.77, $SD_{Females}$ = 4.27). The magnitude of the differences in the means (mean difference = .43, 95% CI: .86 to 2.55) was small (η^2 = .05; Hedge's Adjustment =.01).

Table 7.2.2: Independent t-tests and Effect Sizes for Task-Success Probability Estimation and 5 Individual TSP Probes

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Difference	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
Total TSP	3.18 (.80) 90	2.88 (.81) 132	2.75(220)**	.30	.09	.52	.03	.37
TSP1	3.61 (1.09) 99	3.32 (1.04) 154	2.11(251)*	.29	.02	.56	.02	.27
TSP2	2.81 (1.04) 110	2.54 (1.04) 150	2.01(248)*	.27	.01	.54	.02	.48
TSP 3	3.43 (1.02) 98	2.97 (1.10) 143	3.27(237)**	.46	.18	.73	.04	.43
TSP 4	3.40 (.91) 99	3.20 (1.09) 143	1.51(240)	.20	-.06	.46	.01	.20
TSP 5	2.67 (1.15) 96	2.31 (1.13) 140	2.38(234)*	.36	.06	.66	.02	.31

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Hedge's Adjustment or Cohen's d adjusted for sample size.

Table 7.2.3. gives an overview of the 2x2 χ^2 tests[9] and effect sizes for the 5x3 numerical and reasoning problem blocks. Out of fifteen problems, significant gender differences were observed on twelve problems. Despite the unequal gender distribution (67% of participants were females), more males solved correctly the psychometric problems. No significant gender differences were observed on problems

[9] $\chi^2_{(1)} = Z^2 = r^2+N$. Phi (φ) is the best measure of association for χ^2 test (2x2 contingency table); it estimates the extend of the relationship between the variables. For a 2x2 matrix the following formula is used: $\varphi = \sqrt{\chi^2 / N}$, where N is the number of subjects (Bartlett, 1954; Field, 2005, Pallant, 2007).

14C and 21. Phi coefficient effect sizes, using Cohen's effect size criteria (1988), were small. Hypothesis 4 was confirmed.

Table 7.2.3: 2 x 2 Chi Square Tests and Effect Sizes for 5 Numerical and Reasoning Problem Blocks (TCAP) – Per Gender and % Correct Answer

		Correct Answer		Total	Yates Continuity Correction Value for 2x2	Asymp. Sig. (2-sided)	Phi (φ) Coefficient
		Wrong	Right				
Block 1	Male N	70	94	164	10.51	.001	-.15**
Q12A	% Within Gender	43%	57%	100%			
	% Within Correct Answer	27%	41%	34%			
	% of Total	14%	19%	34%			
	Female N	190	134	324			
	% Within Gender	59%	41%	100%			
	% Within Correct Answer	73%	59%	66%			
	% of Total	39%	28%	66%			
	Total N	260	228	448			
	% Within Gender	53%	47%	100%			
Q12B	Male N	91	73	164	9.56	.002	-.14**
	% Within Gender	56%	45%	100%			
	% Within Correct Answer	29%	43%	34%			
	% of Total	19%	15%	34%			
	Female N	227	97	324			
	% Within Gender	70%	30%	100%			
	% Within Correct Answer	71%	57%	66%			
	% of Total	47%	20%	66%			
	Total N	318	170	488			
	% Within Gender	65%	35%	100%			
Q12C	Male N	94	70	164	4.27	.039	-.10*
	% Within Gender	57%	43%	100%			
	% Within Correct Answer	30%	40%	34%			
	% of Total	19%	14%	34%			
	Female N	218	106	324			
	% Within Gender	67%	33%	100%			
	% Within Correct Answer	70%	60%	66%			
	% of Total	45%	22%	66%			
	Total N	312	176	488			
	% Within Gender	64%	36%	100%			
Block 2	Male N	87	77	164	8.41	.004	-.14**
Q14A	% Within Gender	53%	47%	100%			
	% Within Correct Answer	29%	42%	34%			
	% of Total	18%	16%	34%			
	Female N	217	107	324			
	% Within Gender	67%	33%	100%			
	% Within Correct Answer	71%	58%	66%			
	% of Total	45%	22%	66%			
	Total N	304	184	488			
	% Within Gender	62%	38%	100%			
Q14B	Male N	82	82	164	11.55	.001	-.15***
	% Within Gender	50%	50%	100%			
	% Within Correct Answer	28%	43%	34%			

	% of Total	17%	17%	34%			
	Female N	215	109	324			
	% Within Gender	66%	39%	100%			
	% Within Correct Answer	72%	100%	66%			
	% of Total	44%	39%	66%			
	Total N	297	191	488			
	% Within Gender	61%	39%	100%			
Q14C	Male N	162	2	164	.03	.869	-.03
	% Within Gender	99%	1.2%	100%			
	% Within Correct Answer	34%	50%	34%			
	% of Total	33%	.4%	34%			
	Female N	322	2	324			
	% Within Gender	99%	.6%	100%			
	% Within Correct Answer	67%	50%	66%			
	% of Total	66%	.4%	66%			
	Total N	484	4	488			
	% Within Gender	99%	.8%	100%			
Block 3 Q16	Male N	68	96	164	19.54	.000	-.20***
	% Within Gender	42%	59%	100%			
	% Within Correct Answer	25%	44%	34%			
	% of Total	14%	20%	34%			
	Female N	204	120	324			
	% Within Gender	63%	37%	100%			
	% Within Correct Answer	75%	56%	66%			
	% of Total	42%	25%	66%			
	Total N	272	216	488			
	% Within Gender	56%	44%	100%			
Q17	Male N	80	84	164	18.97	.000	-.20***
	% Within Gender	49%	51%	100%			
	% Within Correct Answer	26%	46%	34%			
	% of Total	16%	17%	34%			
	Female N	225	99	324			
	% Within Gender	69%	31%	100%			
	% Within Correct Answer	74%	54%	66%			
	% of Total	46%	20%	66%			
	Total N	305	183	488			
	% Within Gender	63%	38%	100%			
Q18	Male N	164	--	164	--	--	--
	% Within Gender	100%	--	100%			
	% Within Correct Answer	34%	--	34%			
	% of Total	34%	--	34%			
	Female N	324	--	324			
	% Within Gender	100%	--	100%			
	% Within Correct Answer	66%	--	66%			
	% of Total	66%	--	66%			
	Total N	488	--	488			
	% Within Gender	100%	--	100%			
Block 4 Q20	Male N	79	85	164	10.89	.001	-.15**
	% Within Gender	48%	52%	100%			
	% Within Correct Answer	38%	42%	34%			
	% of Total	16%	17%	34%			
	Female N	208	116	324			
	% Within Gender	64%	36%	100%			
	% Within Correct Answer	73%	58%	66%			
	% of Total	43%	24%	66%			
	Total N	287	201	488			
	% Within Gender	59%	41%	100%			

Q21	Male N	148	16	164	.09	.762	.02
	% Within Gender	90%	10%	100%			
	% Within Correct Answer	34%	31%	34%			
	% of Total	30%	3%	34%			
	Female N	288	36	324			
	% Within Gender	89%	11%	66%			
	% Within Correct Answer	66%	69%	66%			
	% of Total	59%	7%				
	Total N	436	52	488			
	% Within Gender	89%	11%	100%			
Q22	Male N	95	69	164	5.23	.022	-.11*
	% Within Gender	58%	42%	100%			
	% Within Correct Answer	30%	41%	34%			
	% of Total	20%	14%	34%			
	Female N	223	101	324			
	% Within Gender	69%	31%	100%			
	% Within Correct Answer	70%	59%	66%			
	% of Total	46%	21%	66%			
	Total N	318	170	488%			
	% Within Gender	65%	100%	100%			
Block 5 Q24	Male N	95	69	164	5.56	.018	-.11*
	% Within Gender	58%	42%	100%			
	% Within Correct Answer	30%	41%	34%			
	% of Total	20%	14%	34%			
	Female N	224	100	324			
	% Within Gender	69%	31%	100%			
	% Within Correct Answer	70%	59%	66%			
	% of Total	46%	21%	66%			
	Total N	319	169	488			
	% Within Gender	65%	35%	100%			
Q25	Male N	144	20	164	5.79	.016	-.12*
	% Within Gender	88%	12%	100%			
	% Within Correct Answer	32%	53%	34%			
	% of Total	30%	4%	34%			
	Female N	306	18	324			
	% Within Gender	94%	6%	100%			
	% Within Correct Answer	68%	47%	66%			
	% of Total	63%	4%	66%			
	Total N	450	38	488			
	% Within Gender	92%	8%	100%			
Q26	Male N	104	60	164	9.17	.002	-.14**
	% Within Gender	63%	37%	100%			
	% Within Correct Answer	30%	44%	34%			
	% of Total	21%	12%	34%			
	Female N	249	75	324			
	% Within Gender	77%	23%	100%			
	% Within Correct Answer	71%	56%	66%			
	% of Total	51%	15%	66%			
	Total N	353	135	488			
	% Within Gender	72%	28%	100%			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

7.2.3.3. Gender, Task-Success Probability (TSP) and Total Correct Aptitude Problems (TCAP) as Predictors of DMIQ1 and DMIQ2

Firstly, the relationships between the DMIQ1 and DMIQ2, gender, TSP and TCAP were explored. Table 7.2.4. shows the results of the correlational and partial correlational analyses. DMIQ1 and DMIQ2 were strongly intercorrelated (r = .78, *p* =.00). Gender correlated negatively (r = -.30, *p* =.00), with DMIQ1 as well as DMIQ2 (r = -.30, *p* =.00), with females providing lower scores than males. A positive relationship was observed between DMIQ1 and TSP (r = .47, *p* =.00) and DMIQ2 and TSP (r = .62, *p* =.00). DMIQ1 also correlated positively with TCAP (r= .16, *p* <.01) as did DMIQ2 (r= .40, *p* =.00). Interestingly, the correlations between TSP, TCAP and DMIQ2 were stronger than with DMIQ1. A medium positive correlation was observed between TSP and TCAP (r =.43, *p* =.00).

Table 7.2.4: Correlations and Partial Correlations, Means and Standard Deviations between DMIQ1 and DMIQ2, Gender, TSP, TCAP and Age

	DMIQ1	DMIQ2	G	TSP	TCAP	A
	112.86	108.43	1.66	3.00	4.34	22.33
	(19.37)	(21.20)	(.47)	(.82)	(4.45)	(6.86)
DMIQ1						
DMIQ2	.78***					
Gender	-.30***	-.30***				
TSP	.47***	.62***	-.18**			
TCAP	.16**	.40***	-.18***	.43***		
Age	.08	.01	-.14**	-.06	.12*	
-Controlled For Age-						
DMIQ1						
DMIQ2	.78***					
Gender	-.29***	-.30***				
TSP	.48***	.63***	-.19**			
TCAP	.15**	.40***	-.17**	.44***		

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). *N between 198 and 487.*

As in previous studies, the role of age in the DMIQ estimation process was examined. Despite the wide age range (53 years), no significant relationships were observed between age and DMIQ1 and DMIQ2. A negative relationship was observed between age and gender ($r= -.14$, $p <.01$) indicating that females in this sample were younger than males. A positive relationship between age and TCAP ($r= .12$, $p =.01$) indicated that older participants solved more TCAP problems. This finding is contrary to assertions that fluid cognitive ability declines with age (e.g. Beier & Ackerman, 2001, 2003; Deary et al., 2001).

The correlations were re-run, with age partialled out. Preliminary analyses were performed to ensure no violation of the assumptions of normality, linearity and homoscedasticity. An inspection of the zero order correlation matrix suggested that controlling for age had little impact on the strength of the observed relationships, with values slightly higher.

Subsequently, the data was split per gender and the correlational analysis recomputed. The results are presented in Table 7.2.5. TSP displayed a strong positive relationship with DMIQ1 and DMIQ2 for both genders, with stronger correlations between TSP and DMIQ2 than between TSP and DMIQ1. Medium positive correlations were observed between TCAP and DMIQ2 for both genders, but no significant relationships were observed between TCAP and DMIQ1. These findings indicate that the relationships between TSP and TCAP and DMIQ became stronger following the task.

Table 7.2.5: Correlations, Means and Standard Deviations between DMIQ1 and DMIQ2, TSP, TCAP and Age – Per Gender

	Males		Females	
	DMIQ1	DMIQ2	DMIQ1	DMIQ2
	120.64	116.02	108.55	102.57
	(18.13)	(21.58)	(18.70)	(21.14)
DMIQ1				
DMIQ2	.64***		.83***	
TSP	.49***	.65***	.41***	.57***
TCAP	.14	.44***	.10	.31***
Age	.01	.08	.07	-.07

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
N between 47 and 321.

7.2.3.4. Gender as the best predictor of DMIQ1 and DMIQ2

To determine the best predictor of DMIQ1 and DMIQ2 the simultaneous multiple regressions were performed. The dependent variables were DMIQ1 and DMIQ2 and the independent variables were gender, TSP and TCAP. Results are reported in Table 7.2.6. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity.

The first model predicting DMIQ1 was significant $F(3,212) = 26.48$, $p = .00$, Adjusted $R^2 = .26$, $f^2 = .37$), with the overall model explaining 27% of total variance. Gender ($\beta = -.23$, $p = .00$, $r_{part} = -.22$) and TSP ($\beta = .46$, $p = .00$, $r_{part} = .41$) were significant predictors of DMIQ1,with gender accounting for 5% and TSP for 17% of variance. TCAP did not significantly contribute to the prediction of DMIQ1. Contrary to prediction, TSP and not gender was the best predictor of the DMIQ1. Hypothesis 5 was not supported.

The second model, predicting DMIQ2 was also significant $F(3,205) = 53.43$, $p = .00$, Adjusted $R^2 = .43$, $f^2 = .79$), with the overall model explaining 44% of total variance. Gender ($\beta = -.18$, $p < .01$, $r_{part} = -.17$), TSP ($\beta = .54$, $p = .00$, $r_{part} = .48$) and TCAP ($\beta = .14$, $p < .05$, $r_{part} = .12$) were significant predictors, explaining 3%, 23%

and 1% of variance respectively. As in DMIQ1, TSP, and not gender, was the best predictor of DMIQ2. Hypothesis 6 was also not supported.

Table 7.2.6: Beta coefficients for Simultaneous Multiple Regressions of Gender, TSP and TCAP onto DMIQ1 and DMIQ2

Dependent Variable	***DMIQ1*** β	*t*	***DMIQT2*** β	*t*
Gender	**-.23**	**-3.83*****	**-.18**	**-3.26****
TSP	**.46**	**7.07*****	**.54**	**9.17*****
TCAP	-.08	-1.20	**.14**	**2.34***
Regression Model	**F(3, 212) = 26.48*****		**F(3, 205) = 53.43*****	
R^2	.27		.44	
R^2 Change	.27		.44	
Adj. R^2	.26		.43	
f^2	.37		.79	

p <.05, ** *p* <.01, *** *p* <.001 *Note:* Significant values are in bold.

7.2.3.5. Impact of Gender on the Relationship between TSP and DMIQ1 and DMIQ2

TSP was collapsed into a categorical variable, with Group 1 containing individuals who had low confidence in their ability to successfully solve similar tasks in the future. Group 2 was made up of individuals who had an average confidence and Group 3 was made up of high confidence individuals. Results are presented in Table 7.2.7.

Table 7.2.7: Overview of TSP Banded

	TSP	n
Group 1	<=3	88
Group 2	3-4	77
Group 3	4+	57

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analysis of variance were conducted to explore whether gender influences the relationship between TSP and DMIQ1 and DMIQ2. Results are presented in Table 7.2.8. For DMIQ1, Levene's Test of Equality of Error Variance was significant (*p* <.05), indicating that the variance across the groups was not equal. As a result, a more stringent significance level, *p* =.01, was set for evaluating the results of the analysis. The interaction effect between gender and TSP estimation conditions was not significant, $F(2,210) = .30$, $p = .74$, $\eta_p^2 = .00$. There was a statistically significant main effect for TSP, $F(2,210) = 19.56$, $p = .00$, $\eta_p^2 = .16$ with large effect size. The main effect for gender was also significant, $F(1,210) = 13.26$, $p = .00$, $\eta_p^2 = .06$, with medium effect size.

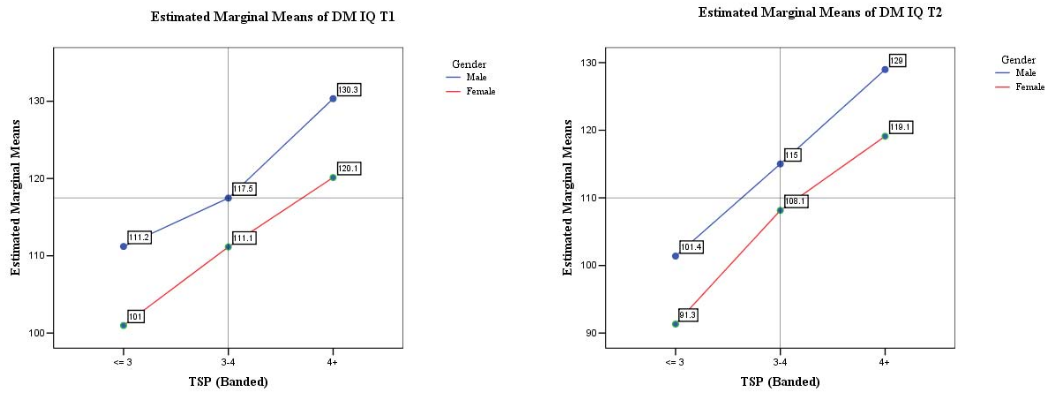
Table 7.2.8: Two-way ANOVAS (TSP and gender) on DMIQ1 and DMIQ2

Variable	TSP Groups	Mean Score (SD)			*F*-score		
		Total	Males	Females	TSP	Gender	TSP x Gender
DMIQ1	G1 (L)	104.43 (20.17)	111.21 (23.80)	100.98 (17.28)	19.56***	13.26***	.30
	G2 (M)	113.76 (16.17)	117.47 (16.23)	111.15 (15.78)			
	G3 (H)	125.33 (15.69)	130.34 (12.75)	120.13 (16.95)			
DMIQ2	G1 (L)	94.56 (23.04)	101.38 (27.69)	91.33 (19.97)	34.82***	11.10**	.16
	G2 (M)	111.01 (15.90)	115.02 (15.55)	108.14 (15.71)			
	G3 (H)	124.04 (16.24)	128.98 (13.05)	119.11 (17.78)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: DMIQ1 = Domain-Masculine Intelligence Type at pre-task estimation condition; DMIQ2 = Domain-Masculine Intelligence Type at post-task estimation condition. TSP = Task-success probability estimation condition.

Planned contrasts revealed significant differences between Group 1 and Group 2, (Contrast Estimate -13.68, $p = .00$) and between Group 2 and Group 3 (Contrast Estimate -10.93, $p = .00$). Post-hoc comparisons using the Games-Howell test indicated that the mean score for Group 1 (<= 3) was significantly different from Group 2 (3-4) as well as Group 3 (4+). The mean score for Group 2 was also significantly different from Group 3. Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 7 was partially confirmed.

Figure 7.2.1: Two 2-way ANOVAs (TSP and Gender) on DMIQ1 and DIMQ2



For DMIQ2, the interaction effect between gender and TSP was not significant, F(2,203) = .16, p = .86, η_p^2 = .00. There was a statistically significant main effect for TSP, F(2,203) = 34.82, p =.00, η_p^2 = .26, with large effect size, and for gender, F(1,203) = 11.10, p <.01, η_p^2 = .05, with medium effect size. Planned contrasts revealed significant differences between Group 1 and Group 2, (Contrast Estimate -21.46, p =.00) and between Group 2 and Group 3 (Contrast Estimate -12.47, p =.00). Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=3) was significantly different from Group 2 (3-4) as well as from Group 3 (4+). Group 2 mean scores were also significantly different from Group 3. Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 8 was partially confirmed.

7.2.3.6. Impact of Gender on the Relationship between TCAP and DMIQ1 and DMIQ2

Individual scores of total correctly solved 15 psychometric aptitude problems were computed, forming a new variable: Total Correct Aptitude Problems (TCAP).

TCAP was collapsed into a categorical variable with three groups, with Group 1 containing individuals who did not correctly solve any problems. Group 2 was made of individuals who solved average number of problems and Group 3 of individuals who correctly solved the majority of problems. Results are presented in Table 7.2.9.

Table 7.2.9: Overview of TCAP Banded

	TCAP	n
Group 1	<=0	232
Group 2	1-8	120
Group 3	9+	136

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender influences the relationship between TCAP and DMIQ1 and DMIQ2. Results are presented in Table 7.2.10. For DMIQ1, the interaction effect between gender and TCAP was significant, $F(2,381) = 3.26$, $p < .05$, $\eta^2 = .02$, with small effect size. The main effect for TCAP, $F(2,381) = 19.56$, $p = .00$, $\eta^2 = .09$, was also significant, with medium effect size. The main effect for gender $F(1,381) = 26.49$, $p = .00$, $\eta^2 = .07$ was also significant, with medium effect size. Planned contrasts revealed significant differences between Group 2 and Group 3, (Contrast Estimate -14.73, $p = .00$).

Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=0) was significantly different from Group 2 (1-8). Group 1 also significantly differed from Group 3 (9+). Group 2 mean scores were also significantly different from Group 3. This was confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets.

Table 7.2.10: Two 2-way ANOVAs (TCAP and gender) on DMIQ1 and DMIQ2

Variable	TCAP Groups	Mean Score (SD)			*F*-score		
		Total	Males	Females	TCAP	Gender	TCAP x Gender
DMIQ1	G1 (L)	111.84 (19.15)	122.50 (16.04)	107.65 (18.70)	19.56***	26.49***	3.26*
	G2 (M)	105.41 (20.68)	107.41 (19.70)	104.55 (21.15)			
	G3 (H)	120.53 (15.26)	126.73 (14.60)	114.69 (13.54)			
DMIQ2	G1 (L)	98.12 (22.44)	105.30 (18.66)	95.25 (23.29)	28.35***	12.99***	.01
	G2 (M)	115.34 (19.15)	120.71 (21.21)	110.21 (15.42)			
	G3 (H)	107.97 (22.27)	116.02 (21.58)	102.57 (21.14)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

Note: DMIQ1 = Domain-Masculine Intelligence Type at pre-task estimation condition; DMIQ2 = Domain-Masculine Intelligence Type at post-task estimation condition. TCAP = Total correct aptitude problems.

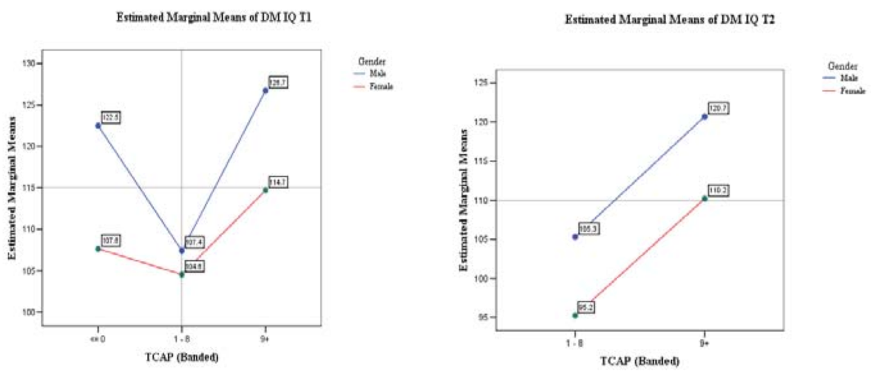
As the main interaction effect was significant, further investigation of the relationship was warranted. Simple effects analysis was conducted. The data was split per gender and two one-way between-groups analyses of variance were conducted. For males, the one-way between-groups analysis of variance for DMIQ1 was significant, F(2,135) = 16.01, p =.00, η^2 =.19, with large effect size. The robust tests of equality of means, Welch (2, 72) = 12.83, p =.00; Brown-Forsythe (2, 97) = 14.67, p =.00 were also significant. Post-hoc comparisons using the Tukey HSD and Bonferroni tests revealed significant differences in mean scores between Group 1 (<=0) (M = 122.50, SD = 16.05) and Group 2 (1-8) (M = 107.41, SD = 19.70) as well as between Group 2 (1-8) and Group 3 (9+) (M = 126.73, SD = 14.60).

The Levene's Test of Equality of Error Variance was significant (p <.05) in the female sub-sample. As a result, a more stringent significance level, i.e. p =.01, was set for evaluating the results of the analysis. For females, the one-way analysis of variance was also significant, F (2,246) = 5.87, p<.01, η^2 =.05, with medium effect size. The robust tests of equality of means, Welch (2, 160) = 7.55, p <.01; Brown-

Forsythe (2,227) = 6.14, p <.01 were significant. The post-hoc comparisons using the Games-Howell test revealed significant differences between Group 1 (<=0) (M = 107.65, SD = 18.70) and Group 3 (9+) (M = 114.69, SD = 13.38) and between Group 3 and Group 2 (1-8) (M = 114.69, SD = 13.54). Hypothesis 9 was confirmed.

For DMIQ2, the interaction effect between gender and TCAP was not significant, F(1,225) = .01, *p* = .94, η^2 = .00. The main effect for TCAP, F(1,225) = 28.35, *p* =.00, η^2 = .11 was significant, with medium effect size. The main effect for gender, F(1,225) = 12.99, *p* =.00, η^2 = .06 was significant with medium effect size. Planned contrasts revealed significant differences between Group 1 and Group 2, (Contrast Estimate -15.18, *p* =.00). Post-hoc comparisons were not computed as for TCAP only two categories were available, i.e. Group 2 and Group 3 were available. Hypothesis 10 was partially confirmed.

Figure 7.2.2: Two 2-way ANOVAs (TCAP and Gender) on DMIQ1 and DMIQ2



Thus, hypotheses 1, 2, 3, 4 and 9 were confirmed and hypotheses 5 and 6 were not supported. Hypotheses 7, 8 and 10 were partially confirmed.

7.2.4. Discussion

The aim of this study was to confirm the occurrence of HHE on DMIQ1 and DMIQ2. The results confirmed the existence of gender differences on the numerical-spatial factor of SEI (η^2 =.09, *d* =.66 for DMIQ1 and η^2 =.09, *d* =.63 for DMIQ2). Equally, a significant decrease in DMIQ estimates was observed from the pre-task to post-task estimation condition (*d* =.38). The results also revealed significant gender differences in the task-success probes, with males providing higher task-success estimates than females. Yet, males also solved correctly more psychometric problems than did females. The observed effect sizes for both TSP and TCAP were small.

The findings also revealed a stronger relationship between TSP, TCAP and DMIQ2, compared to DMIQ1. This pattern was also observed when the data was split per gender, with TSP and DMIQ2 having stronger relationship than TSP and DMIQ1. Interestingly, for both genders, TCAP only correlated with DMIQ2 and not with DMIQ1. These results indicate that although TSP and TCAP were not assessed during DMIQ1, TSP or task confidence already played a role in the estimation process, indicating the individuals rely on their confidence before they are prompted to do so.

As in previous studies, gender was expected to be the best predictor of DMIQ. The results failed to validate this claim, with TSP confirmed as the best predictor of DMIQ1 and DMIQ2, over and above gender and TCAP, explaining 17% and 23% of variance respectively. Thus, it appears that TSP or task confidence plays an important role in the prediction of the intelligence type.

Subsequently, the role gender plays in the relationship between TSP and DMIQ1 and DMIQ2 was investigated. For DMIQ1, results revealed significant task-success effect, with significant differences between the lowest, average and high task-success groups, with the lowest DMIQ1 estimates provided in the lowest TSP group, average estimates in the average TSP group and the highest DMIQ1 estimates in the highest TSP estimates group. Equally, a significant gender effect revealed that males were more confident than females across the three groups. These results provided further support for the role of confidence in the self-estimation process as well as for male hubris. The results were identical for DMIQ2.

Finally, gender’s role in the relationship between TCAP and DMIQ1 and DMIQ2 was examined. For DMIQ1, the results revealed a significant interaction effect as well as significant TCAP and gender effects. Significant differences between the three TCAP groups were observed; with lowest DMIQ1 estimates provided by the group that solved an average number of psychometric problems, average DMIQ1 estimates by the group that did not solve any problems and the highest estimates by the group that solved most psychometric problems. Identical estimation patterns were observed for males and females respectively. These results provided additional support for the role of BTAE and WTAE biases in the self-estimation process (e.g. Alicke et al., 1995; Dunning et al., 1999; Kruger & Dunning, 1999).

Still, males provided higher DMIQ1 estimates than females in all three groups. Further analyses showed that males’DMIQ1 estimates were significantly different in the lowest and medium TCAP groups as well as between the medium and the highest TCAP groups. Significant differences were also observed for females, with DMIQ1 estimates significantly different in the lowest and highest as well as between medium and highest TCAP groups.

For DMIQ2, the results revealed a significant TCAP effect, with findings identical to the DMIQ1 estimation pattern. Equally, a significant gender effect revealed that males provided higher DMIQ2 estimates than females across the three groups, providing further support for the hubris-humility effect in self-estimated intelligence.

Thus, while gender differences exist in self-estimated intelligence, and in particular in the domain-masculine intelligence type, one's confidence in ability to succeed on a gender stereotype-inducing task, was a better determinant of performance than gender itself. Equally, contrary to some assertions (Ehrlinger & Dunning, 2003; Johnson & Bouchard, 2007; Kruger & Dunning, 1999), the results demonstrated that individuals were capable of making accurate self-estimates that match their confidence levels. Likewise, the existence of the hubris-humility effect, and in particular of the male hubris, was established in the pre- and post- task conditions. As the psychometric task was likely to activate gender-stereotypical biases, it was unsurprising that the provided self-estimates did not match the number of correctly resolved problems, with only the most capable problem solvers providing accurately matching self-estimates, while inflated self-estimates were provided by the average and the least capable problem solvers. Hence, self-confidence seems to positively influence the accuracy of self-estimates, but the psychometric task that evokes cognitive stereotypical biases, seems to impact the accuracy of self-estimates.

7.3. Study 12

Gender, TSP and TCAP as predictors of the Domain-Masculine Intelligence Type

7.3.1. Introduction

This study set to validate the findings of Study 11. The study was identical in set-up and execution, with the following exception. Two numerical problems were dropped; Q14C because it had the lowest percentage of correct answers and Q18 since it yielded zero correct answers. The other measures were unchanged. In order to further substantiate the previous results, Study 12 ensured that the gender groups were homogeneous in size.

Thus, it is predicted that HHE would prevail on DMIQ1 and DMIQ2 (H1) and that a significant reduction will occur in DMIQ2 (H2). Males were expected to give significantly higher TSP estimations than females (H3). Sex differences are expected to be observed in the psychometric problems, with males providing more correct answers (H4). Further, gender was expected to be the best predictor of the DMIQ1 (H5) and DMIQ2 (H6), over and above TSP and TCAP. Based on previous findings, gender is expected to influence the relationship between TSP and DMIQ1 (H7) and DMIQ2 (H8). Gender was also expected to affect the relationship between TCAP and DMIQ1 (H9) and DMIQ2 (H10). Male and female DMIQ2 estimates were expected to differ in response to TSP probes, while DMIQ1 estimates are controlled for (H11). Equally, males and female DMIQ2 estimates are expected to differ in response to the psychometric problems, while DMIQ1 estimates are controlled for (H12).

7.3.2. Method

Participants

A total of one hundred and eighty-two participants took part in the second experimental online study. There were 92 females (50.5%) and 90 males (49.5%). Their age ranged from 17 to 50 (M = 22.84, SD = 6.51) years. All participants were fluent in English and no language or other problems were reported. 55% had completed A-levels, 21% achieved BA/BSc level, and 10% MA/MSc/MBA or equivalent level of education.

Measures

Repeated Measure of Domain-Masculine Intelligence Type (DMIQ).

See Study 11 (section 7.2.2). Alpha for DMIQ1 was .85 and DMIQ2 .88.

Psychometric Aptitude Task

Total Correct Aptitude Problems (TCAP)

Numerical and Reasoning Problems (Bryon, 2006)

Thirteen numerical and reasoning problems that were based on actual intelligence test items were presented in three blocks of three and two blocks of two analogous problems (Bryon, 2006). For an overview of the problems see Appendix. Participants were informed that items in each block varied in difficulty level, ranging from elementary to difficult. A time limit of 60 or 90 seconds was given for each block. Participants were advised to leave unanswered problems blank in order to not exceed the time limit, or be disqualified. The time limit was set to reflect a real-life testing situation, with the entire task taking 6.5 minutes. Correct answers were available at the end of the survey. Alpha for the thirteen items was .53.

Task Success Probability

Task Success Probability Estimation Measure (TSP) (Storek, 2007)

See Study 11 (section 7.2.2). The alpha for the five-item measure was .81.

Procedure

Participants were from the general public. They were recruited through an intensive mass email campaign by the main researcher. An email invitation, with a URL link (www.zoomerang.com) to the study and a background explanation of the study was sent to all participants. The snow-balling technique of participant recruitment was used, i.e. participants were asked to forward the study invitation and the URL link to as many acquaintances and friends as possible. In total, 230 individuals logged onto the site during the period of June to December 2007. The data was gathered through an online survey engine www.Zoomerang.com and participation was voluntary.

Detailed scoring instructions were given at the beginning of each measure, including timing instructions for the numerical and reasoning problems. Participants were aware that the study was approved by UCL Ethics Committee, meeting confidentiality and Data Protection requirements. Debrief feedback, correct answers and opportunity to leave survey feedback were provided at the end.

7.3.3. Results

7.3.3.1. HHE and DMIQ1 and DMIQ2

Two independent samples t-tests were computed to assess whether significant gender differences or HHE occurred on DMIQ1 and DMIQ2. Results are presented in Table 7.3.1. Significant gender differences, with males providing higher DMIQ estimates in T1 and T2 estimation conditions, were observed. Hypothesis 1 was confirmed.

Table 7.3.1: Overview of Independent t-Tests and Effect Sizes for DMIQ1 and DMIQ2

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI		Effect Size	
					L	U	η^2	*d*
DMIQ1	120.94 (6.06) 90	104.59 (18.46) 92	6.06(180)***	16.35	11.02	21.68	**.17**	**1.19**
DMIQ2	117.46 (18.10) 90	95.96 (19.13) 92	7.78(180)***	21.50	16.05	26.95	**.25**	**1.15**

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Cohen's *d*. Large effect sizes are in bold.

To test hypothesis 2 whether significant change occurred from DMIQ1 to DMIQ2 following the intervention task, a paired-samples t-test was conducted. There was a statistically significant decrease in DMIQ1 (M = 112.68, SD = 19.93) to DMIQ2 (M = 106.59, SD = 21.48), t(181) = 7.77, p = .00, two-tailed, r = .87, p =.00. The mean decrease in DMIQ was 6.09 (SD = 10.57) with 95% confidence interval ranging from 4.54 to 7.64. Cohen's *d* (.58) indicated a medium effect size. Hypothesis 2 was confirmed.

7.3.3.2. Gender Differences in Task Success Probability Estimation (TSP) and Psychometric Aptitude Task (TCAP)

Table 7.3.2. gives an overview of independent-samples t-tests and effect sizes for the five individual task-success probability (TSP) estimation probes and the Total TSP measure. The independent samples t-tests for the five TSP probes and the Total TSP measure were significant, with males providing higher TSP estimates than females. The observed effect sizes were small to medium. Inspection of the correlational results (see Table 7.3.4.) revealed a medium negative correlation

between gender and TSP (*r* = -.32, *p* =.00), with males providing higher TSP estimates than females (M_{Males} = 3.24, SD_{Males} = .79; $M_{Females}$ = 2.71, $SD_{Females}$ = .77). Hypothesis 3 was confirmed.

Table 7.3.2: Independent t-tests and Effect Sizes for Task-Success Probability Estimation and 5 Individual TSP Probes

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Difference	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
Total TSP	3.24 (.79) 82	2.71 (.77) 85	4.39(164)***	.53	.29	.77	.10	.68
TSP1	3.69 (1.03) 88	3.20 (1.05) 91	3.19(177)**	.50	.19	80	.05	.47
TSP2	2.82 (1.86) 88	2.36 (1.01) 89	2.95(175)**	.50	.15	.77	.05	.31
TSP 3	3.48 (1.02) 88	2.79 (1.12) 89	4.29(175)***	.69	.37	1.01	.10	.64
TSP 4	3.44 (.91) 90	3.13 (1.06) 89	2.10(177)*	.31	.02	.60	.02	.31
TSP 5	2.72 (1.16) 88	2.09 (1.08) 88	3.69(174)***	.63	.29	.96	.07	.56

* *p* < .05 ** *p* < .01 *** *p* < .001 (2-tailed). Note: *d* = Hedge's Adjustment or Cohen's d adjusted for sample size.

Inspection of the correlational results (see Table 7.3.4.) revealed a small negative correlation between gender and TCAP (r =-.26, *p* =.00), with males correctly solving more problems then females (M_{Males} = 9.04, SD_{Males} = 1.87; $M_{Females}$ = 7.95, $SD_{Females}$ = 2.24). Table 7.3.3. gives an overview of the 2x2 χ^2 tests and effect sizes for the thirteen psychometric problems. Significant gender differences were observed only on four problems, i.e. Q12A, Q16, Q17 and Q20, with males providing significantly more correct answers than females. This finding differs from the previous study where thirteen problems (87%) revealed significant gender differences. Phi coefficient values, using Cohen's effect size criteria (1988), were small. An independent samples t-test revealed significant gender differences on TCAP, t(180) = 3.60, p =. 00 two-tailed, with males (M = 9.04, SD = 1.87) correctly solving more psychometric problems than females (M = 7.95, SD = 2.24). The magnitude of the differences in the means (mean difference = 1.10, 95% CI: .50 to 1.70) was medium (η^2 = .07; Cohen's d =.53). Hypothesis 4 was confirmed.

Table 7.3.3: 2x2 Chi Square Tests and Effect Sizes for 5 Numerical and Reasoning Problem Blocks – Per Gender and % Correct Answer

		Correct Answer		Total	Yates Continuity Correction Value for 2x2	Asymp. Sig. (2-sided)	Phi (φ) Coefficient
		Wrong	Right				
Block 1	Male N	5	85	90	5.14	.013	-.19*
Q12A	% Within Gender	6%	94%	100%			
	% Within Correct Answer	24%	53%	50%			
	% of Total	3%	47%	50%			
	Female N	16	76	92			
	% Within Gender	17%	83%	100%			
	% Within Correct Answer	77%	47%	51%			
	% of Total	9%	42%	51%			
	Total N	21	161	182			
	% Within Gender	12%	89%	100%			
Q12B	Male N	25	65	90	2.15	n.s.	-.12
	% Within Gender	28%	72%	100%			
	% Within Correct Answer	41%	54%	50%			
	% of Total	14%	36%	50%			
	Female N	36	56	92			
	% Within Gender	39%	61%	100%			
	% Within Correct Answer	59%	46%	51%			
	% of Total	20%	31%	51%			
	Total N	61	121	182			
	% Within Gender	34%	67%	100%			
Q12C	Male N	25	65	90	.15	n.s.	-.04
	% Within Gender	28%	72%	100%			
	% Within Correct Answer	46%	51%	50%			
	% of Total	14%	36%	50%			
	Female N	29	63	92			
	% Within Gender	32%	69%	100%			
	% Within Correct Answer	54%	49%	51%			
	% of Total	16%	35%	51%			
	Total N	54	128	182			
	% Within Gender	30%	70%	100%			
Block 2	Male N	19	71	90	2.50	n.s.	-.13
Q14A	% Within Gender	21%	79%	100%			
	% Within Correct Answer	39%	53%	50%			
	% of Total	10%	39%	50%			
	Female N	30	62	92			
	% Within Gender	33%	67%	100%			
	% Within Correct Answer	61%	47%	51%			
	% of Total	17%	34%	51%			

	Total N	49	133	182			
	% Within Gender	27%	73%	100%			
Q14B	Male N	15	75	90	2.35	n.s.	-.13
	% Within Gender	17%	83%	100%			
	% Within Correct Answer	38%	53%	50%			
	% of Total	8%	41%	50%			
	Female N	25	67	92			
	% Within Gender	27%	78%	100%			
	% Within Correct Answer	63%	100%	51%			
	% of Total	14%	78%	51%			
	Total N	40	142	182			
	% Within Gender	22%	78%	100%			
Block 3 Q16	Male N	3	87	90	8.17	.004	-.23**
	% Within Gender	3%	97%	100%			
	% Within Correct Answer	16%	53%	50%			
	% of Total	2%	48%	50%			
	Female N	16	76	92			
	% Within Gender	17%	83%	100%			
	% Within Correct Answer	84%	47%	51%			
	% of Total	9%	42%	51%			
	Total N	19	163	182			
	% Within Gender	10%	90%	100%			
Q17	Male N	15	75	90	6.11	.013	-.20**
	% Within Gender	17%	83%	100%			
	% Within Correct Answer	33%	55%	50%			
	% of Total	8%	41%	50%			
	Female N	31	61	92			
	% Within Gender	34%	66%	100%			
	% Within Correct Answer	67%	45%	51%			
	% of Total	46%	34%	51%			
	Total N	46	136	182			
	% Within Gender	25%	75%	100%			
Block 4 Q20	Male N	12	78	90	3.27	n.s.	-.15*
	% Within Gender	13%	87%	100%			
	% Within Correct Answer	34%	53%	50%			
	% of Total	7%	43%	50			
	Female N	23	69	92			
	% Within Gender	25%	75%	100%			
	% Within Correct Answer	66%	47%	51%			
	% of Total	13%	38%	51%			
	Total N	35	147	182			
	% Within Gender	19%	81%	100%			
Q21	Male N	76	14	90	2.45	n.s.	.13
	% Within Gender	84%	16%	100%			
	% Within Correct Answer	53%	37%	50%			
	% of Total	42%	8%	50%			
	Female N	68	24	92			
	% Within Gender	74%	26%	100%			
	% Within Correct Answer	47%	63%	51%			
	% of Total	37%	13%	51%			
	Total N	144	38	182			
	% Within Gender	79%	21%	100%			
Q22	Male N	27	63	90	.00	n.s.	-.02
	% Within Gender	30%	70%	100%			
	% Within Correct Answer	48%	50%	50%			
	% of Total	15%	35%	50%			

	Female N	29	63	92			
	% Within Gender	32%	69%	100%			
	% Within Correct Answer	52%	50%	51%			
	% of Total	16%	35%	51%			
	Total N	56	126	182%			
	% Within Gender	31%	69%	100%			
Block 5	Male N	25	65	90	1.03	n.s.	-.09
Q24	% Within Gender	28%	72%	100%			
	% Within Correct Answer	43%	52%	50%			
	% of Total	14%	36%	50%			
	Female N	33	59	92			
	% Within Gender	36%	64%	100%			
	% Within Correct Answer	57%	48%	51%			
	% of Total	18%	32%	51%			
	Total N	58	124	182			
	% Within Gender	32%	68%	100%			
Q25	Male N	75	15	90	.48	n.s.	-.07
	% Within Gender	83%	17%	100%			
	% Within Correct Answer	48%	58%	50%			
	% of Total	41%	8	50%			
	Female N	81	11	92			
	% Within Gender	88%	12%	100%			
	% Within Correct Answer	52%	42%	51%			
	% of Total	45%	6%	51%			
	Total N	156	26	182			
	% Within Gender	86%	14	100%			
Q26	Male N	34	56	90	3.25	n.s.	-.15
	% Within Gender	38%	62%	100%			
	% Within Correct Answer	42%	56%	50%			
	% of Total	19%	31%	50%			
	Female N	48	44	92			
	% Within Gender	52%	48%	100%			
	% Within Correct Answer	59%	44%	51%			
	% of Total	26%	24%	51%			
	Total N	82	100	182			
	% Within Gender	45%	55%	100%			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

7.3.3.3. Gender, TSP and TCAP as Predictors of DMIQ1 and DMIQ2

Firstly, the relationship between the DMIQ1 and DMIQ2, gender, TSP, TCAP and age was explored. Table 7.3.4. shows the correlational results. DMIQ1 and DMIQ2 were strongly intercorrelated, which is not surprising (r = .87, p =.00). Gender correlated negatively with DMIQ1 (r = -.41, p =.00) and DMIQ2 (r = -.50, p =.00), with females providing lower scores than males (DMIQ1M_{Males} = 120.94, SD_{Males} = 17.96; DMIQ1$M_{Females}$ = 104.59, $SD_{Females}$ = 18.46; DMIQ2M_{Males} = 117.46, SD_{Males} = 18.10; DMIQ1$M_{Females}$ = 95.96, $SD_{Females}$ = 19.13).

Table 7.3.4: Correlations, Means and Standard Deviations between DMIQ1, DMIQ2, Gender, TSP, TCAP and Age

	DMIQ1	DMIQ2	G	TSP	TCAP	A
	112.68	106.59	1.51	2.97	8.49	22.84
	(19.93)	(21.48)	(.50)	(.82)	(2.13)	(6.51)
DMIQ1						
DMIQ2	.87***					
Gender	-.41***	-.50***				
TSP	.50***	.60***	-.32***			
TCAP	.45***	.51***	-.26***	.53***		
Age	.05	-.02	-.11	-.10	.10	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).   *N = between 167 and 182.*

Strong positive correlations were observed between TSP and DMIQ1 (r =.50, p =.00) and between TSP and DMIQ2 (r =.60, p =.00). Strong positive correlations were also observed between TCAP and DMIQ1 (r =.45, p =.00) and between TCAP and DMIQ2 (r =.51, p =.00). A strong positive relationship was observed between TSP and TCAP (r = .53, p =.00). These results are similar to results of Study 11, yet, the correlations between TSP, TCAP and DMIQ1 are even stronger.

As in previous studies and given the age range of the participants, i.e. 33 years, age was included in the analysis to explore whether it had an impact on DMIQ. No significant relationships were observed.

7.3.3.4. Gender as the best predictor of DMIQ1 and DMIQ2

To determine the best predictor of DMIQ1 and DMIQ2 two simultaneous multiple regressions were performed. The dependent variables were DMIQ1 and DMIQ2 and the independent variables were gender, TSP and TCAP. Results are reported in Table 7.3.5. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity.

The first model predicting DMIQ1 was significant $F(3,163) = 30.44$, $p = .00$, Adjusted $R^2 = .35$, $f^2 = .56$), with the overall model explaining 36% of total variance. Gender ($\beta = -.26$, $p = .00$, $r_{part} = -.24$), TSP ($\beta = .30$, $p = .00$, $r_{part} = .25$) and TCAP ($\beta = .23$, $p < .01$, $r_{part} = .19$) were significant predictors of DMIQ1, accounting for 6%, 6% and 4% of variance respectively. As in Study 11, TSP was the best predictor of the DMIQ1. Hypothesis 5 was not supported.

The second model, predicting DMIQ2 was also significant $F(3,163) = 55.74$, $p = .00$, Adjusted $R^2 = .50$, $f^2 = 1.04$), with the overall model explaining 51% of total variance. Gender ($\beta = -.32$, $p = .00$, $r_{part} = -.30$), TSP ($\beta = .38$, $p = .00$, $r_{part} = .31$) and TCAP ($\beta = .23$, $p < .01$, $r_{part} = .19$) were significant predictors, explaining 9%, 10% and 4% of variance respectively. As in DMIQ1 and identical to Study 11, TSP was the best predictor of DMIQ2. Hypothesis 6 was also not supported.

Table 7.3.5: Beta coefficients for Simultaneous Multiple Regressions of Gender, TSP and TCAP onto DMIQ1 and DMIQ2

Dependent Variable	***DMIQ1***		***DMIQT2***	
	β	*t*	*β*	*t*
Gender	**-.26**	**-3.83*** **	**-.32**	**-5.47*** **
TSP	**.30**	**3.98*** **	**.38**	**5.68*** **
TCAP	**.23**	**3.05** **	**.23**	**3.53** **
Regression Model	**F(3, 163) = 30.44*** **		**F(3, 163) = 55.74*** **	
R^2	.36		.51	
R^2 Change	.36		.51	
Adj. R^2	.35		.50	
f^2	.56		1.04	

p <.05, ** *p* <.01, *** *p* <.001 *Note:* Significant values are in bold.

7.3.3.5. Impact of Gender on the Relationship between TSP on DMIQ1 and DMIQ2

TSP was collapsed into a categorical variable, with Group 1 containing individuals with lowest confidence in their ability to successfully solve similar tasks in the future. Group 2 was made up of individuals that had an average confidence in their abilities, and Group 3 of individuals with high confidence in their abilities. Results are presented in Table 7.3.6.

Table 7.3.6: Overview of TSP Banded

	TSP	n
Group 1	<=3	56
Group 2	3-4	70
Group 3	4+	41

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender influences the relationship between TSP and DMIQ1 and DMIQ2. Results are presented in Table 7.3.7. For DMIQ1, the interaction effect between gender and TSP estimation conditions was not significant, $F(2,161) = 2.39$, $p = .10$, $\eta_p^2 = .03$. There was a statistically significant main effect for TSP, $F(2,161) = 16.12$, $p = .00$, $\eta_p^2 = .17$ with large effect size. The main effect for gender was also significant, $F(1,161) = 13.23$, $p = .00$, $\eta_p^2 = .08$, with medium effect size. Planned contrasts revealed significant differences between Group 1 and Group 2, (Contrast Estimate -16.21, $p = .00$) and between Group 2 and Group 3 (Contrast Estimate -9.39, $p < .01$). Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<= 3) was significantly different from Group 2 (3-4) as well as from Group 3 (4+). The mean score for Group 2 was also significantly different from Group 3. Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 7 was partially confirmed.

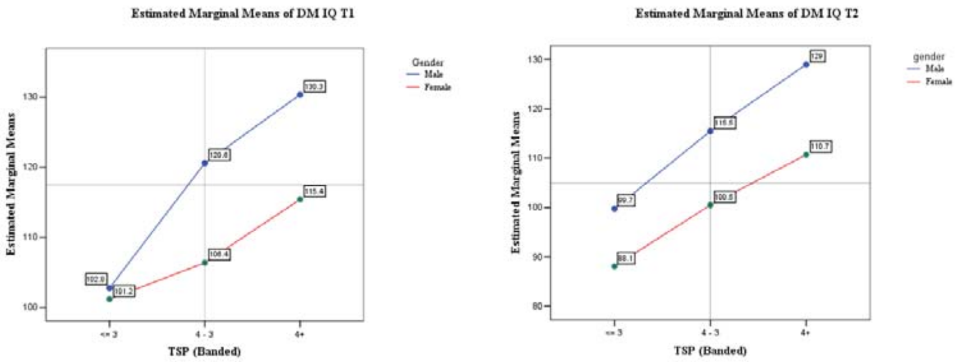
Table 7.3.7: Two 2-way ANOVAs (TSP and gender) on DMIQ1 and DMIQ2

Variable	TSP Groups	Mean Score (SD)			*F*-score		
		Total	Males	Females	TSP	Gender	TSP x Gender
DMIQ1	G1 (L)	101.62 (19.65)	102.75 (20.49)	101.16 (19.55)	16.12***	13.23***	2.39
	G2 (M)	114.08 (17.09)	120.58 (15.84)	106.36 (15.40)			
	G3 (H)	125.60 (14.86)	130.34 (12.75)	115.38 (14.30)			
DMIQ2	G1 (L)	91.42 (19.94)	99.75 (16.93)	88.09 (20.26)	24.53***	28.04***	.40
	G2 (M)	108.61 (17.72)	115.46 (16.60)	100.48 (15.63)			
	G3 (H)	123.18 (15.71)	128.98 (13.05)	110.69 (13.87)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: DMIQ1 = Domain-Masculine Intelligence Type at pre-task estimation condition; DMIQ2 = Domain-Masculine Intelligence Type at post-task estimation condition. TSP = Task-success probability estimation condition.

For DMIQ2, the interaction effect between gender and TSP estimation conditions was not significant, $F(2,161) = .40$, $p = .67$, $\eta_p^2 = .01$. There was a statistically significant main effect for TSP, $F(2,161) = 24.53$, $p =.00$, $\eta_p^2 = .23$, and for gender, $F(1,161) = 28.04$, $p =.00$, $\eta_p^2 = .15$, both with large effect sizes. Planned contrasts revealed significant differences between Group 1 and Group 2, (Contrast Estimate -19.93, $p =.00$) and between Group 2 and Group 3 (Contrast Estimate -11.87, $p =.00$). Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=3) was significantly different from Group 2 (3-4) as well as from Group 3 (4+). Group 2 mean scores were also significantly different from Group 3. Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 8 was partially confirmed.

Figure 7.3.1: Two 2-way ANOVAs (TSP and Gender) on DMIQ1 and DMIQ2



7.3.3.6. Impact of Gender on the Relationship between TCAP and DMIQ1 and DMIQ2

Individual scores of the total correctly solved fifteen psychometric problems were computed, forming a new variable: Total Correct Aptitude Problems (TCAP). TCAP was collapsed into a categorical variable, with Group 1 containing individuals that correctly solved fewest problems, Group 2 individuals that accurately solved average number of problems, and Group 3 individuals that correctly solved the most problems. Results are presented in Table 7.3.8.

Table 7.3.8: Overview of TCAP Banded

	TCAP	n
Group 1	<=8	81
Group 2	8-9	41
Group 3	10+	60

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender influences the relationship between TCAP and DMIQ1 and DMIQ2. Results are presented in Table 7.3.9. For DMIQ1, the interaction effect between gender and TCAP was not significant, F(2,176) = .29, *p* = .75, η_p^2 = .00. The main effect for TCAP, F (2,176) = 18.77, *p* =.00, η_p^2 = .17, was significant, with large effect size. The main effect for gender F(1,176) = 20.64, p =.00, η_p^2 = .11 was also significant, with medium effect size. Planned contrasts revealed significant differences between Group 1 and Group 2, (Contrast Estimate -14.75, p =.00). Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=8) was significantly different from Group 2 (8-9) as well as from Group 3 (10+). This was confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 9 was partially confirmed.

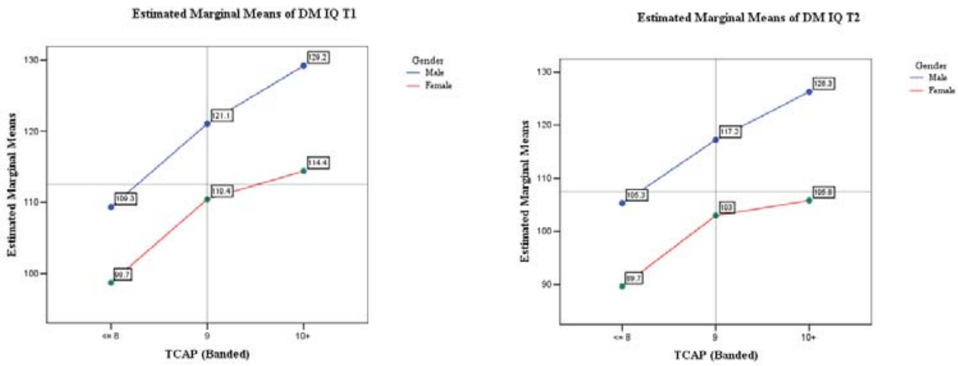
Table 7.3.9: Two 2-way ANOVAs (TCAP and gender) on DMIQ1 and DMIQ2

Variable	TCAP Groups	Mean Score (SD)			*F*-score		
		Total	Males	Females	TCAP	Gender	TCAP x Gender
DMIQ1	G1 (L)	102.39 (20.26)	109.34 (19.70)	98.72 (19.75)	18.77***	20.64***	.29
	G2 (M)	116.40 (15.92)	121.07 (14.39)	110.44 (16.17)			
	G3 (H)	124.03 (14.29)	129.21 (13.80)	114.40 (9.55)			
DMIQ2	G1 (L)	95.07 (20.93)	105.30 (18.66)	89.66 (20.17)	20.12***	39.19***	.48
	G2 (M)	110.98 (17.39)	117.24 (15.58)	102.97 (16.61)			
	G3 (H)	119.14 (16.14)	126.31 (13.79)	105.83 (11.02)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: DMIQ1 = Domain-Masculine Intelligence Type at pre-task estimation condition; DMIQ2 = Domain-Masculine Intelligence Type at post-task estimation condition. TCAP = Total Correct Aptitude Problems.

For DMIQ2, the Levene's Test of Equality of Error Variance was significant (p <.05), indicating the DMIQ2 variance across the groups was not equal. Firstly, the largest and the smallest standard deviations were squared. The largest squared SD was divided by the smallest squared SD, with resulting value of 1.41, which is smaller than the recommended value of 2, suggesting that the group variances were not unacceptably unequal. Equally, a more stringent significance level, p =.01, was set for evaluating the results of the analysis.

Figure 7.3.2: Two 2-way ANOVAs (TCAP and Gender) on DMIQ1 and DMIQ2



The interaction effect between gender and TCAP was not significant, F (2,176) = .48, p = .62, η_p^2 = .01. The main effect for TCAP, F (2,176) = 20.12, p =.00, η_p^2 = .19 was significant, with large effect size. The main effect for gender, F (1,176) = 39.19, p =.00, η_p^2 = .18 was also significant, with large effect size. Planned contrasts revealed significant differences between Group 1 and Group 2, (Contrast Estimate - 15.61, p =.00). Post-hoc comparisons using the Games-Howell test indicated that the mean score for Group 1 (<= 8) was significantly different from Group 2 (8-9) as well as from Group 3 (10+). Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 10 was partially confirmed.

7.3.3.7. Gender Differences in DMIQ2 Estimates in Response to TSP

A 2-by-2 between-groups analysis of covariance[10] was conducted to assess the influence of the TSP probes on the DMIQ2 estimates for males and females. The independent variables were TSP and gender. The dependent variable was DMIQ2. DMIQ1 was used as a covariate to control for individual differences. Preliminary checks were conducted to ensure that there was no violation of the assumptions of

[10] Analysis of covariance (ANCOVA) is recommended in situations with two-group pre-test/post-test design. The pre-test scores are treated as a covariate to control for pre-existing differences between the groups. Thus, ANCOVA is particularly useful in situations with small sample size and only small or medium effect sizes. (Pallant, 2007, p. 291).

normality, linearity, homogeneity of variances, homogeneity of regression slopes, and reliable measurement of the covariate.

Levene's Test of Equality of Error Variance was significant (p <.05), indicating the groups variances were not equal. An alternative check for comparing variances was used. Firstly, the largest and the smallest standard deviations were squared. The largest squared SD was divided by the smallest squared SD, with resulting value of 1.36, which is smaller than the recommended value of 2, suggesting that the group variances were not unacceptably unequal. Subsequently, a more stringent significance level, p =.01, was set for evaluating the results of the analysis. Homogeneity of regression slopes assumption was not violated, F(2,159) = 1.23, p =.29 for the TSP by DMIQ1 interaction, nor for the gender by DMIQ1 interaction, F(1,159) = .52, p =.47.

After adjusting for DMIQ1 estimates, there was a non-significant interaction effect between TSP and gender, F (2,160) = 1.80, p = .17, η_p^2 = .02. The main effect for TSP was significant, F(2,160) = 6.97, p < .01, η^2 = .08, with medium effect size. The main effect of gender was significant, F(1,160) = 14.94, p =.00, η_p^2 = .09, with medium effect size. The main effect for the covariate variable DMIQ1 was also significant, F(1,160) = 324.31, p =.00, η_p^2 = .67, with the covariate significantly and positively related to DMIQ2 and a large effect size.

Planned comparisons analysis revealed significant differences between Group 2 and Group 1, (Contrast Estimate 4.60, p <.05), between Group 3 and Group 1 (Contrast Estimate 8.75, p =.00) and between the genders (Contrast Estimate 6.56, p =.00). Males provided higher self-estimates of ability (Group 1: M_{Male} = 99.75, SD_{Male} =16.93; M_{Female} = 88.09, SD_{Female} =20.26; Group 2: M_{Male} = 115.46, SD_{Male} =16.60; M_{Female} = 100.48, SD_{Female} =15.63; Group 3: M_{Male} =128.98, SD_{Male} =13.05; M_{Female} =

110.69, SD_{Female} =13.87). The results confirmed that gender, and in particular male hubris plays, as well as task-success probability, a role in DMIQ2. Equally, DMIQ1 contributed to DMIQ2 estimations. Hypothesis 11 was partially confirmed.

7.3.3.8. Gender Differences in DMIQ2 in Response to TCAP

A 2-by-2 between-groups analysis of covariance was conducted to assess the influence of TCAP on DMIQ2 estimates for males and females. The independent variables were TCAP and gender. The dependent variable was DMIQ2. DMIQ1 was used as a covariate to control for individual differences. Preliminary checks were conducted to ensure that there was no violation of the assumptions of normality, linearity, homogeneity of variances, homogeneity of regression slopes, and reliable measurement of the covariate.

Homogeneity of regression slopes assumption was not violated for the TCAP by DMIQ1 assumption, $F\ (2,174) = .58$, p =.56 nor for the gender by DMIQ1 interaction, $F\ (1,174) = .36$, p =.55. After adjusting for DMIQ1 estimates, there was a non-significant interaction effect between TCAP and gender, $F\ (2,175) = .23$, $p = .80$, $\eta_p^2 = .00$. The main effect for TCAP was not significant, $F\ (2,175) = 2.30$, $p = .10$, $\eta_p^2 = .03$. The main effect for gender was significant, $F(1,175) = 17.20$, p =.00, $\eta_p^2 = .09$, with medium effect size. The main effect for the covariate variable DMIQ1 was significant, $F(1,175) = 330.60$, p =.00, $\eta_p^2 = .65$, with the covariate significantly and positively related to DMIQ2 and of very large effect size.

Planned comparisons analysis revealed significant differences between Group 3 and Group 1, (Contrast Estimate 4.01, $p < .05$) and between the genders (Contrast Estimate 6.94, p =.00). Males provided higher self-estimates of ability (Group 1: M_{Male} = 105.30, SD_{Male} =18.66; M_{Female} = 89.66, SD_{Female} =20.17; Group 2: M_{Male} =

117.24, SD_{Male} =15.58; M_{Female} = 102.97, SD_{Female} =16.61; Group 3: M_{Male} =126.31, SD_{Male} =13.79; M_{Female} = 105.83, SD_{Female} =11.02). The results confirmed that gender, and in particular male hubris play a role in DMIQ2 but TCAP did not. Equally, DMIQ1 contributed to DMIQ2 estimations. Hypothesis 12 was partially confirmed.

Thus, hypotheses 1, 2, 3, and 4 were confirmed and hypotheses 5 and 6 were not supported. Hypotheses 7, 8, 9, 10, 11 and 12 were partially supported.

7.3.4. Discussion

This study set out to validate the findings of Study 11. The results confirmed the existence of HHE on DMIQ1 (η^2 =.17, d =1.19 for DMIQ1 and on DMIQ2 (η^2 =.25, d =1.15). Equally, a significant decrease in DMIQ estimates was observed from the pre-task to post-task estimation condition (d =.58). The results also revealed significant gender differences in the task-success probes, with males providing higher task-success estimates than females. Males also correctly solved more psychometric problems than did females. The observed effect sizes for both TSP and TCAP were small to medium. Stronger relationships were also observed between TSP, TCAP and DMIQ2 than between TSP, TCAP and DMIQ1.

As in previous studies, gender was expected to be the best predictor of DMIQ. Results failed to validate this claim, with TSP confirmed as the best predictor of DMIQ1 and DMIQ2, over and above gender and TCAP, explaining 6% and 10% of variance respectively. As in Study 11, TSP or task confidence plays an important role in the prediction of the intelligence type.

The role that gender plays in the relationship between TSP and DMIQ1 and DMIQ2 was investigated next. For DMIQ1, results revealed a significant task-success effect, with significant differences between the lowest, average and high task-success groups, with the lowest DMIQ1 estimates provided in the lowest TSP group, average estimates in the average TSP group and highest DMIQ1 estimates in the highest TSP estimates group. Equally, a significant gender effect revealed that males were more confident than females across the three groups. These results provide added support for the role of task-confidence in the SEI estimation process and for the display of male hubris in the estimation process. Identical results pattern was observed for DMIQ2.

Subsequently, the role gender plays in the relationship between TCAP and DMIQ1 and DMIQ2 was investigated. For DMIQ1, results revealed a significant TCAP effect, with significant differences between the lowest, average and high TCAP groups, with the lowest DMIQ1 estimates provided by the group that solved fewest TCAP problems, average estimates by the average TCAP group and highest DMIQ1 estimates by the group that solved the most TCAP problems. Equally, significant gender effects revealed that males provided higher DMIQ1 estimates than females across the three groups. These results provide additional support for the assertion that individuals are aware of their abilities and thus capable of accurate self-assessment (e.g. Ackerman et al., 2002; Chamorro-Premuzic et al., 2010; Swim, 1994) of ability as well as for male hubris. An identical result pattern was observed for DMIQ2.

Lastly, two 2x2 between-groups analyses of covariance were conducted to assess whether males and females provided different DMIQ2 estimates in their response to TSP probes as well as the psychometric problems. Both analyses confirmed gender differences in DMIQ2 but not as a result of TSP probes or psychometric problems.

Thus, the results of this study replicated the findings of Study 11 in that the existence of the hubris-humility effect was confirmed on the domain-masculine intelligence type in both estimation conditions. Confidence in one's ability to succeed on a psychometric stereotype-inducing task was again the best predictor of the intelligence type. Equally, the results confirmed that the provided self-estimates accurately matched individuals' confidence levels. Contrary to Study 11, the supplied self-estimates were also accurately provided by subjects in all three ability groups, providing further support for the assertion that individuals are capable of accurate self-assessments of ability.

7.4. Study 13

Gender, TSP and TCAP as predictors of the Domain-Masculine Intelligence Type

7.4.1. Introduction

This study sets out to confirm the findings of Study 11 and 12. The study was identical in set-up and execution, with the following changes. Firstly, based on participants' feedback, the number of numerical and reasoning items was reduced and more time was given. In the previous two studies, several participants declared that they felt they did not have sufficient time to complete the problems and that this time pressure caused stress. Thus, the number of the numerical and reasoning problems was reduced to six, with problems Q12A, Q12B, Q14A and Q14B dropped as they were identical to other problems.

In addition, ten items that assessed crystallised intelligence (Gc) were added in order to assess whether the addition of Gc problems will impact the previously observed sex differences in TCAP (e.g. Ackerman, 2006; Lynn, Irwing, & Cammock, 2002; Novel & Hedges, 1998). Likewise, only three TSP probes were included, compared to five used in the previous two studies. As in Study 12, and to facilitate the validity of previous results by ensuring homogeneous gender groups, 80 participants were randomly selected from the overall sample and all analyses were computed with the smaller sample.

As in previous studies, HHE is expected to be observed on DMIQ1 and DMIQ2 (H1). Equally, significant reduction in DMIQ estimates from T1 to T2 is predicted (H2). Males were expected to give significantly higher TSP estimations than females

(H3). In agreement with previous findings, males were also expected to provide significantly more correct answers to the psychometric problems (H4). Gender is hypothesised to be the best predictor of DMIQ1 (H5) and DMIQ2 (H6), over and above TSP and TCAP. Conclusively, gender is expected to influence the relationship between TSP and DMIQ1 (H7) and DMIQ2 (H8) and between TCAP and DMIQ1 (H9) and DMIQ2 (H10).

7.4.2. Method

Participants

A total of one hundred and thirty-six participants took part in the third experimental online study. There were 96 females (71%) and 40 males (29%). However, as this study aimed to validate previous results using homogeneously sized gender groups, 56 female subjects were dropped and 40 males (50%) and 40 females (50%) were randomly selected, bringing the total number of participants to 80. All analyses were run with both samples, yielding alike results. This study uses the results of the smaller sample. Participants' age ranged from 17 to 60 (M = 26.65, SD = 10.21) years. 10% of participants completed GSCE/O-levels or similar level of education, 63% achieved A-levels and non-university level of education, 11% achieved BA/BSc level, and 15% achieved MA/MSc/MBA or equivalent level of education. 35% were the youngest child, 34% the oldest, 18% the middle child and 11% the only child, and 1% had a twin. 21% reported Arts/Drama/Music as their favourite subject at school, 16% reported English Language as their favourite subject, 14% reported mathematics, 8% reported History and 8% reported Biology, 6% reported Psychology/ Sociology/Philosophy, 5% reported Physics, 5% Geography and

5% PE/Sports, 4% reported Business/Economics and 4% IT/Media, 3% reported Sciences and 3% Music.

Measures

Repeated Measures Domain-Masculine Intelligence Type (DMIQ)

See Study 11 (section 7.2.2). Alpha for DMIQ1 was .73 and for DMIQ2 .85.

Psychometric Aptitude Task

Total Correct Aptitude Problems (TCAP)

Numerical and Reasoning Problems (Bryon, 2006)

Six numerical and reasoning problems that were taken from an intelligence test training book were presented in two blocks of three problems (Bryon, 2006). A time limit of 3 minutes was given. Participants were advised to leave unanswered problems blank, in order to not exceed the time limit, or be disqualified. The time limit was set to reflect a real-life testing situation. Compared to Studies 11 and 12, more time was given, incorporating previous feedback. Participants were instructed to note their answers, as correct answers were available at the end of the survey. An alpha for the six items was .54 and the inter-item correlation was r =.16. For an overview of the problems see Appendix.

Crystallised Knowledge Task

Crystallized Intelligence (Gc): General Knowledge: General Knowledge (GKT: Irwing, Cammock, & Lynn, 2001)

See Study 3 (section 3.2.2).Ten items from the 72-item questionnaire measuring general knowledge were selected, assessing knowledge of literature, general science, medicine, games, fashion and finance. A time limit of 2 minutes was given. Participants were advised to leave unanswered problems blank, in order to not exceed the time limit, or face disqualification. An overview of the items is in the Appendix. The alpha for the ten items was 81.

Task Success Probability

Task Success Probability Estimation Measure (TSP) (Storek, 2007)

See Study 11 (section 7.2.2). The alpha for the three-item measure was .66 and the inter-item correlation was = .39.

Procedure

The majority of participants were from the general public, although several UCL undergraduates and their acquaintances also took part. They were recruited through an intensive email campaign by the main researcher and four second year UCL students who participated in a mini-research study group in spring of 2008 that the main researcher was leading. An email invitation, with a URL link (www.zoomerang.com) to the study and a background explanation of the study was sent to all participants. The snow-balling technique of participant recruitment was used, i.e. participants were asked to forward the study invitation and the URL link to as many acquaintances as possible.

In total, 231 individuals logged onto the site during the period of February and April 2008. The data was gathered through an online survey engine www.Zoomerang.com and participation was voluntary. Detailed scoring instructions were given at the beginning of each measure, including time instructions for the psychometric problems. Participants were aware that the study was approved by UCL Ethics Committee, meeting confidentiality and Data Protection requirements. Debrief feedback, correct answers to the psychometric problems and opportunity to leave feedback were provided at the end of the survey. All participants were fluent in English and no language or other problems were reported.

7.4.3. Results

7.4.3.1. Domain-masculine intelligence and the Hubris and Humility Effect in T1 and T2

Two independent samples t-tests were computed to assess whether significant gender differences on DMIQ were observed in the pre-task (T1) and post-task (T2) estimation condition. Results (see Table 7.4.1.) corroborated the existence of HHE on DMIQ1 and DMIQ2, confirming hypothesis 1.

Table 7.4.1: Overview of Independent t-Tests and Effect Sizes for DMIQ1 and DMIQ2

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
DMIQ1	114.29 (15.45) 40	98.50 (10.26) 40	5.39(68)***	15.81	9.96	21.67	**.30**	**1.20**
DMIQ2	113.06 (17.22) 33	94.10 (12.92) 39	5.53(70)***	18.96	11.86	26.05	**.30**	**1.25**

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Cohen's *d*. Large effect sizes are in bold.

To test whether a significant decrease occurred in DMIQ2 from DMIQ1, a paired-samples t-test was conducted. There was a statistically significant decrease in DMIQ estimates from T1 (M = 106.74, SD = 15.92) to T2 (M = 102.79, SD = 17.70), t(71) = 4.87, p = .00, two-tailed, r = .92, p =.00. The mean decrease in DMIQ was 3.95 (SD = 6.89) with 95% confidence interval ranging from 2.33 to 5.57. Cohen's *d* (.57) indicated a medium effect size. Hypothesis 2 was confirmed.

7.4.3.2. Gender Differences in TSP and TCAP

Table 7.4.2. gives an overview of independent-samples t-tests and effect sizes for the three individual TSP probes and the Total TSP measure. The independent samples t-tests for Total TSP measure was significant, with medium (η^2 =.07) effect size. Male subjects provided higher TSP estimates than did females. Among the three individual TSP probes, only TSP3 that was asked after Gc questions was significant, with medium effect size (η^2 =.08), with males providing higher probability estimates than females.

Table 7.4.2: Independent t-Tests and Effect Sizes for Task-Success Probability Estimation and 3 Individual TSP Probes

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Difference	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
Total TSP	3.22 (.72) 31	2.81 (.76) 39	2.25(68)*	.40	.05	.76	.07	.55
TSP1	3.22 (.94) 32	2.90 (.94) 39	1.43(69)	.32	-.13	.77	.03	.34
TSP2	3.27 (.98) 33	2.90 (.97) 39	1.63(70)	.38	-.08	.83	.04	.38
TSP 3	3.24 (.97) 33	2.64 (1.04) 39	2.52(70)*	.60	.13	1.08	.08	.60

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Cohen's d.

Inspection of the correlational results (see Table 7.7.4.) revealed a small negative correlation between TSP and gender (r = -.26, p <.05), with males providing higher TSP estimates than females (M_{Males} = 3.22, SD_{Males} = .72; $M_{Females}$ = 2.81, $SD_{Females}$ = .76). In order to investigate whether the TSP correlation pattern differed for males and females, the data was split per gender and the correlations rerun. For males no significant relationships were observed. For females, medium strength positive correlations were observed between TSP and DMIQ1 (r = .32, p <.05) and between TSP and DMIQ2 (r = .34, p <.05). Hypothesis 3 was confirmed.

Table 7.4.3. gives an overview of the 2x2 χ^2 tests and effect sizes for six numerical and reasoning and ten crystallised knowledge problems. No significant gender differences were observed on the sixteen problems. A small significant effect size was found on Q24 (Phi coefficient =.23, p <.05). These findings differ notably from previous results. In order to investigate whether TCAP correlated differently in male and female subsamples, the data was split per gender and the correlations re-ran. For males the data revealed a medium strength positive relationship between TCAP and DMIQ1 (r = .41, p <.01). No other relationships were observed. For females, a medium strength positive correlation was observed between TCAP and DMIQ2 (r = .46, p <.01). A medium strength positive relationship was also observed between TCAP and TSP (r = .46, p <.01). Inspection of the correlational results (see Table 7.4.4.) revealed no significant relationship between TCAP and gender (r =.04, p =.76) and nor was an independent samples t-test for TCAP significant, t(67) -.31, p =.76; M_{Males} = 7.25, SD_{Males} = 4.30; $M_{Females}$ = 7.50, $SD_{Females}$ = 2.79; Mean Differences = -.25, 95 CI from -1.87 to 1.37; η^2 = .00; Cohen's d = .07 Hypothesis 4 was not supported.

Table 7.4.3: 2 x 2 Chi Square Tests and Effect Sizes for 6 Numerical and Reasoning and 10 Crystallised Intelligence Problem Blocks – Per Gender and % Correct Answer

		Correct Answer Wrong	Correct Answer Right	Total	Yates Continuity Correction Value for 2x2	Asymp. Sig. (2-sided)	Phi (φ) Coefficient
Block 1	Male N	11	29	40	.51	.47	-.11
Q12	% Within Gender	28%	73%	100%			
	% Within Correct Answer	42%	54%	50%			
	% of Total	14%	36%	50%			
	Female N	15	25	40			
	% Within Gender	38%	63%	100%			
	% Within Correct Answer	58%	46%	50%			
	% of Total	19%	31%	50%			
	Total N	26	54	182			
	% Within Gender	32%	68%	100%			
Q13	Male N	14	26	40	.00	1.00	.03
	% Within Gender	35%	65%	100%			
	% Within Correct Answer	52%	49%	50%			
	% of Total	18%	33%	50%			
	Female N	13	27	40			
	% Within Gender	33%	68%	100%			
	% Within Correct Answer	48%	51%	50%			
	% of Total	16%	34%	50%			
	Total N	27	53	182			
	% Within Gender	34%	66%	100%			
Q14	Male N	22	18	40	.80	.37.	.13
	% Within Gender	55%	45%	100%			
	% Within Correct Answer	56%	44%	50%			
	% of Total	28%	23%	50%			
	Female N	17	23	40			
	% Within Gender	43%	58%	100%			
	% Within Correct Answer	44%	56%	50%			
	% of Total	21%	29%	50%			
	Total N	39	41	80			
	% Within Gender	49%	51%	100%			
Block 2	Male N	16	24	40	.22	.64	.08
Q16	% Within Gender	40%	60%	100%			
	% Within Correct Answer	55%	47%	50%			
	% of Total	20%	30%	50%			
	Female N	13	27	40			
	% Within Gender	33%	68%	100%			
	% Within Correct Answer	45%	53%	50%			
	% of Total	16%	34%	50%			
	Total N	29	51	80			
	% Within Gender	36%	64%	100%			
Q17	Male N	33	7	40	1.10	.29	.15
	% Within Gender	83%	183%	100%			
	% Within Correct Answer	54%	37%	50%			
	% of Total	41%	9%	50%			
	Female N	28	12	40			
	% Within Gender	70%	30%	100%			
	% Within Correct Answer	46%	63%	50%			
	% of Total	35%	15%	50%			
	Total N	61	19	80			
	% Within Gender	76%	24%	100%			

Q18	Male N	22	18	40	.45	.50	.10
	% Within Gender	55%	45%	100%			
	% Within Correct Answer	55%	45%	50%			
	% of Total	28%	23%	50%			
	Female N	18	22	40			
	% Within Gender	45%	55%	100%			
	% Within Correct Answer	45%	55%	50%			
	% of Total	23%	28%	50%			
	Total N	40	40	80			
	% Within Gender	50%	50%	100%			
Block 3	Male N	38	2	40	.00	1.00	.05
GNK	% Within Gender	95%	5%	100%			
Q20	% Within Correct Answer	51%	40%	50%			
	% of Total	48%	3%	50%			
	Female N	37	3	40			
	% Within Gender	93%	8%	100%			
	% Within Correct Answer	49%	60%	50%			
	% of Total	46%	4%	50%			
	Total N	75	5	80			
	% Within Gender	94%	6%	100%			
Q21	Male N	31	9	40	.08	.78	-.06
	% Within Gender	78%	23%	100%			
	% Within Correct Answer	48%	56%	50%			
	% of Total	39%	11%	50%			
	Female N	33	7	40			
	% Within Gender	83%	18%	100%			
	% Within Correct Answer	52%	44%	50%			
	% of Total	41%	9%	50%			
	Total N	64	16	80			
	% Within Gender	80%	20%	100%			
Q22	Male N	17	23	90	.00	1.00	.00
	% Within Gender	43%	58%	100%			
	% Within Correct Answer	50%	50%	50%			
	% of Total	21%	29%	50			
	Female N	17	23	92			
	% Within Gender	43%	58%	100%			
	% Within Correct Answer	50%	50%	50%			
	% of Total	21%	29%	50%			
	Total N	34	46	80			
	% Within Gender	43%	58%	100%			
Q23	Male N	38	2	40	.18	.67	.10
	% Within Gender	95%	5%	100%			
	% Within Correct Answer	51%	33%	50%			
	% of Total	48%	3%	50%			
	Female N	36	4	40			
	% Within Gender	90%	10%	100%			
	% Within Correct Answer	49%	67%	50%			
	% of Total	45%	5%	50%			
	Total N	74	6	80			
	% Within Gender	93%	7%	100%			
Q24	Male N	14	26	40	3.27	.07	.23*
	% Within Gender	35%	65%	100%			
	% Within Correct Answer	70%	43%	50%			
	% of Total	18%	33%	50%			
	Female N	6	34	40			
	% Within Gender	15%	85%	100%			
	% Within Correct Answer	30%	57%	50%			
	% of Total	8%	43%	50%			

	Total N	20	60	80%			
	% Within Gender	25%	75%	100%			
Q25	Male N	28	12	40	.06	.80	-.06
	% Within Gender	70%	30%	100%			
	% Within Correct Answer	48%	55%	50%			
	% of Total	35%	15%	50%			
	Female N	30	10	40			
	% Within Gender	75%	25%	100%			
	% Within Correct Answer	52%	46%	50%			
	% of Total	38%	13%	50%			
	Total N	58	22	80			
	% Within Gender	73%	27%	100%			
Q26	Male N	19	21	40	.81	.37	-.13
	% Within Gender	48%	53%	100%			
	% Within Correct Answer	44%	57%	50%			
	% of Total	24%	26%	50%			
	Female N	24	16	40			
	% Within Gender	60%	40%	100%			
	% Within Correct Answer	56%	43%	50%			
	% of Total	30%	20%	50%			
	Total N	43	37	80			
	% Within Gender	54%	46%	100%			
Q27	Male N	17	23	40	.00	1.00	-.03
	% Within Gender	43%	58%	100%			
	% Within Correct Answer	49%	51%	50%			
	% of Total	21%	29%	50%			
	Female N	18	22	40			
	% Within Gender	45%	55%	100%			
	% Within Correct Answer	51%	49%	50%			
	% of Total	23%	28%	50%			
	Total N	35	45	80			
	% Within Gender	44%	56%	100%			
Q28	Male N	17	23	40	2.46	.12.	-.20
	% Within Gender	43%	58%	100%			
	% Within Correct Answer	41%	61%	50%			
	% of Total	21%	29%	50%			
	Female N	25	15	40			
	% Within Gender	63%	38%	100%			
	% Within Correct Answer	60%	40%	50%			
	% of Total	31%	19%	50%			
	Total N	42	38	80			
	% Within Gender	53%	48%	100%			
Q29	Male N	13	27	40	.24	.62	.08
	% Within Gender	33%	68%	100%			
	% Within Correct Answer	57%	47%	50%			
	% of Total	16%	34%	50%			
	Female N	10	30	40			
	% Within Gender	25%	75%	100%			
	% Within Correct Answer	44%	53%	50%			
	% of Total	13%	38%	50%			
	Total N	23	57	80			
	% Within Gender	29%	71%	100%			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

7.4.3.3. Gender, TSP, and TCAP as Predictors of DMIQ1 and DMIQ2

Firstly, the relationships between DMIQ1, DMIQ2, gender, TSP and TCAP were explored. Table 7.4.4. reveals the correlational results. DMIQ1 and DMIQ2 were strongly interrelated (r = .92, *p* =.00). Gender correlated negatively with DMIQ1 (r = -.52, *p* =.00) as well as DMIQ2 (r = -.54, *p* =.00), with females providing lower DMIQ estimates than males.

Table 7.4.4: Correlations, Means and Standard Deviations between DMIQ1,DMIQ2 Gender ,TSP, TCAP and Age

	DMIQ1	DMIQ2	G	TSP	TCAP	A
	106.38	102.79	1.50	2.99	7.38	26.65
	(15.27)	(17.70)	(.50)	(.77)	(3.60)	(10.21)
DMIQ1						
DMIQ2	.92***					
Gender	-.52***	-.54***				
TSP	.35**	.36**	-.26*			
TCAP	.29**	.37**	.04	.38**		
Age	-.09	-.15	-.03	.12	-.00	

* *p* < .05   ** *p* < .01   *** *p* < .001 (2-tailed).   *N = between 70 and 80.*

Medium strength positive correlations were observed between TSP and DMIQ1 (r = .35, *p* <.01) and DMIQ2 (r =.36, *p* <.01) as well as between TCAP and DMIQ1 (r = .29, *p*=<.01) and DMIQ2 (r =.37, *p* <.01). Gender correlated negatively with TSP (r = -.26, p <.05) and there was also a positive medium correlation between TSP and TCAP (r =.38, *p* <.01). Given participants' age range (43 years), age was included in the correlational analysis to explore whether it had an impact on DMIQ estimates. Age did not correlate with any of the variables.

7.4.3.4. Gender as the best predictor of DMIQ1 and DMIQ2

To determine the best predictor of DMIQ1 and DMIQ2 two simultaneous multiple regressions were performed. The dependent variables were DMIQ1 and DMIQ2 and the independent variables were gender, TSP and TCAP. Results are reported in Table 7.4.5. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity.

The first model predicting DMIQ1 was significant $F(3,66) = 13.27$, $p = .00$, Adjusted $R^2 = .35$, $f^2 = .61$), with the overall model explaining 38% of total variance. Gender ($\beta = -.50$, $p = .00$, $r_{part} = -.48$) and TCAP ($\beta = .26$, $p < .05$, $r_{part} = .24$) were significant predictors of DMIQ1, accounting for 23% and 6% of variance respectively. TSP did not significantly contribute to the prediction of DMIQ1. Contrary to results of Study 11 and 12, but in support of Hypothesis 5, gender was the best predictor of the DMIQ1.

The second model, predicting DMIQ2 was also significant $F(3,66) = 17.77$, $p = .00$, Adjusted $R^2 = .42$, $f^2 = .82$), with the overall model explaining 45% of total variance. Gender ($\beta = -.53$, $p = .00$, $r_{part} = -.50$) and TCAP ($\beta = .36$, $p < .01$, $r_{part} = .33$) were significant predictors of DMIQ2, explaining 25% and 11% of variance respectively. TSP did not significantly contribute to the prediction of DMIQ2. Thus, the results were identical to DMIQ1, with gender confirmed as the best predictor. Hypothesis 6 was confirmed.

Table 7.4.5: Beta Coefficients for Simultaneous Multiple Regressions of Gender, TSP and TCAP onto DMIQ1 and DMIQ2

Dependent Variable	***DMIQ1***		***DMIQT2***	
	β	*t*	*β*	*t*
Gender	**-.50**	**-4.90*****	**-.53**	**-5.49*****
TSP	.12	1.06	.09	.84
TCAP	**.26**	**2.46***	**.36**	**3.56****
Regression Model	**F(3, 66) = 13.27*****		**F(3, 66) = 17.77*****	
R^2	.38		.45	
R^2 Change	.38		.45	
Adj. R^2	.35		.42	
f^2	.61		.82	

p <.05, ** *p* <.01, *** *p* <.001 *Note:* Significant values are in bold.

7.4.3.5. Impact of Gender on the Relationship between TSP and DMIQ1 and DMIQ2

TSP was collapsed into a categorical variable with Group 1 containing individuals with lowest confidence in their ability to successfully solve similar tasks in future, Group 2 made of individuals that had an average confidence, and Group 3 made of highly confident individuals. Results are presented in Table 7.4.6.

Table 7.4.6: Overview of TSP Banded

	TSP	n
Group 1	<=3	25
Group 2	3-4	18
Group 3	4+	27

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender influences the relationship between TSP and DMIQ1 and DMIQ2. Results are presented in Table 7.4.7. For DMIQ1, the Levene's Test of Equality of Error Variance was significant (*p* <.05), indicating the DMIQ2 variance across the groups was not equal. As a result, a more stringent significance level, *p* =.01, was set for evaluating the results of the analysis.

The interaction effect between gender and TSP estimation conditions was not significant, F(2,64) = .01, p = .99, η_p^2 = .00. The main effect for TSP, F(2,64) = 1.93, p =.15, η_p^2 = .06, was also not significant.

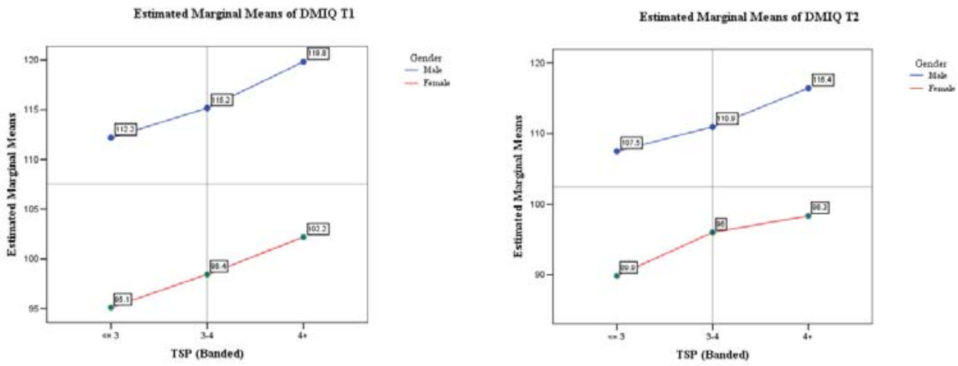
Table 7.4.7: Two 2-way ANOVAs (TSP and gender) on DMIQ1 and DMIQ2

Variable	TSP Groups	Mean Score (SD)			*F*-score		
		Total	Males	Females	TSP	Gender	TSP x Gender
DMIQ1	G1 (L)	100.58 (13.33)	112.19 (11.76)	95.12 (10.35)	1.93	27.85***	.01
	G2 (M)	106.81 (16.50)	115.17 (17.25)	98.44 (11.11)			
	G3 (H)	111.33 (16.41)	119.82 (17.44)	102.19 (8.90)			
DMIQ2	G1 (L)	95.50 (16.41)	107.50 (14.58)	89.85 (14.32)	2.01	20.41***	.06
	G2 (M)	103.47 (17.27)	110.94 (19.67)	96.00 (11.01)			
	G3 (H)	107.72 (17.42)	116.43 (17.91)	98.35 (11.25)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: DMIQ1 = Domain-Masculine Intelligence Type at pre-task estimation condition; DMIQ2 = Domain-Masculine Intelligence Type at post-task estimation condition. TSP = Task-success probability estimation condition.

The main effect for gender was significant, F(1,64) = 27.85, p =.00, η_p^2 = .30, with a very large effect size. Planned contrasts revealed no significant differences between the groups. Post-hoc comparisons using the Games-Howell test indicated that the mean score for Group 1 (<= 3) was significantly different from Group 3 (4+). No other differences were observed. Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 7 was partially confirmed.

Figure 7.4.1: Two 2-way ANOVAs (TSP and Gender) on DMIQ1 and DMIQ2



For DMIQ2, the interaction effect between gender and TSP estimation conditions was not significant, F(2,64) = .06 $p = .94$, $\eta_p^2 = .00$. The main effect for TSP, F(2,64) = 2.01, $p = .14$, $\eta_p^2 = .06$, was also not significant. There was a statistically significant main effect for gender, F(1,64) = 20.41, $p = .00$, $\eta_p^2 = .24$, with a very large effect size. Planned contrasts revealed no significant differences between the groups. Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=3) was significantly different from Group 3. No other significant differences were observed. Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 8 was partially confirmed.

7.4.3.6. Impact of Gender on the Relationship between TCAP and DMIQ1 and DMIQ2

Individual scores for the sixteen correctly solved psychometric problems were computed, creating a new variable TCAP. TCAP was collapsed into a categorical variable, with Group 1 made of individuals who correctly solved fewest problems, Group 2 of individuals who solved an average number of problems and Group 3 of individuals that correctly solved the most psychometric problems. Results are presented in Table 7.4.8.

Table 7.4.8: Overview of TCAP Banded

	TCAP	n
Group 1	<=7	35
Group 2	8-9	23
Group 3	10+	22

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender influences the relationship between TCAP and DMIQ1 and DMIQ2. Results are presented in Table 7.4.9. For DMIQ1, the interaction effect between gender and TCAP was not significant, $F(2,74) = .52$, $p = .60$, $\eta^2 = .01$. The main effect for TCAP, $F(2,74) = 9.33$, $p = .00$, $\eta^2 = .20$, was significant, with large effect size. The main effect for gender $F(1,74) = 34.28$, $p = .00$, $\eta^2 = .32$ was also significant, with a very large effect size. Planned contrasts revealed significant differences between Group 1 and Group 2, (Contrast Estimate -11.52, $p = .00$). Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=7) was significantly different from Group 2 (8-9) as well as from Group 3 (10+).This was confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 9 was partially confirmed.

Table 7.4.9: Two 2-way ANOVAs (TCAP and gender) on DMIQ1 and DMIQ2

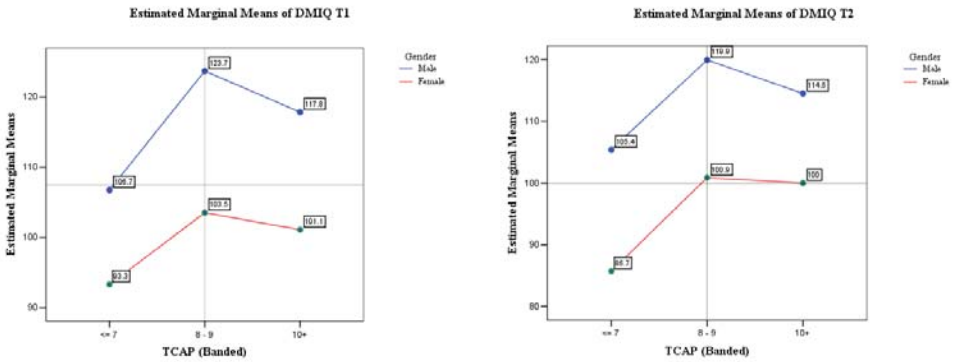
Variable	TCAP Groups	Mean Score (SD)			*F*-score		
		Total	Males	Females	TCAP	Gender	TCAP x Gender
DMIQ1	G1 (L)	99.81 (14.00)	106.74 (14.71)	93.28 (9.75)	9.33***	34.28***	.52
	G2 (M)	110.52 (13.02)	123.69 (10.24)	103.50 (7.88)			
	G3 (H)	124.03 (14.29)	129.21 (13.80)	114.40 (9.55)			
DMIQ2	G1 (L)	93.00 (17.94)	105.40 (19.98)	85.71 (12.08)	7.40***	26.32***	.21
	G2 (M)	107.50 (13.69)	119.94 (13.75)	100.87 (8.03)			
	G3 (H)	109.50 (16.27)	114.50 (15.99)	100.00 (12.75)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

Note: DMIQ1 = Domain-Masculine Intelligence Type at pre-task estimation condition; DMIQ2 = Domain-Masculine Intelligence Type at post-task estimation condition. TCAP = Total Correct Aptitude Problems.

For DMIQ2, the Levene's Test of Equality of Error Variance was significant (p <.05), indicating the DMIQ2 variance across the groups was not equal. As a result, a more stringent significance level, p =.01, was set for evaluating the results of the analysis.

Figure 7.4.2: Two 2-way ANOVAs (TCAP and Gender) on DMIQ1 and DMIQ2



The interaction effect between gender and TCAP was not significant, F(2,66) = .21, p = .81, η^2 = .01. The main effect for TCAP, F(2,66) = 7.40, p =.00, η^2 = .18 was significant, with large effect size. The main effect for gender, F(1,66) = 26.32, p =.00, η^2 = .29 was significant with a very large effect size. Planned contrasts revealed significant differences between Group 1 and Group 2, (Contrast Estimate -13.27, p =.00). Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=7) was significantly different from Group 2 (8-9) as well as from Group 3 (10+). This was confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 10 was partially confirmed.

Thus, hypotheses 1, 2, 3, 5 and 6 were confirmed and hypotheses 7, 8, 9 and 10 were partially confirmed. Hypothesis 4 was not confirmed.

7.4.4. Discussion

This study set out to confirm the findings of Study 11 and Study 12. The results confirmed the existence of HHE on DMIQ1 (η^2 =.30, d =1.20 and DMIQ2 (η^2 =.30, d =1.25). Equally, a significant decrease in DMIQ estimates was observed from the pre-task to post-task estimation condition (d =.57), following the TCAP/TSP task. Study 13 used only three TSP probes. Gender differences were confirmed on the Total TSP measure and on one individual TSP probe, endorsing previous findings and existing literature. Although the gender differences on TSP were less pronounced than in previous studies, males provided higher TSP estimates than females. The observed effect sizes were medium. Contrary to prior findings, no gender differences were observed on the sixteen psychometric problems. Thus, the inclusion of ten crystallised problems and a reduction of the numerical and reasoning problems had an impact on the observed gender differences. These results challenge the male advantage in crystallised intelligence claims and in particular, in the General Knowledge Test (e.g. Lynn & Irwing, 2002, 2004; Lynn, Irwing, & Cammock, 2002). Moreover, stronger relationships were also observed between TSP, TCAP and DMIQ2 than between TSP, TCAP and DMIQ1.

As in previous studies, gender was expected to be the best predictor of DMIQ. Results confirmed gender as the best predictor of DMIQ1 and DMIQ2, explaining 23% and 25% of variance respectively. Contrary to preceding studies, TSP did not play a role in the prediction of the intelligence type, but TCAP did. It appears that the psychometric task content change influenced male and female ability beliefs and performance perceptions, reducing the importance of task confidence. Equally, it seems that the reduction in TSP probes, from five (as in Study 11 and 12) to three, was sufficient for the TSP probes to cease having an affect in the prediction of the intelligence type.

The role that gender plays in the relationship between TSP and DMIQ1 and DMIQ2 was investigated next. The results were identical for DMIQ1 and DMIQ2, revealing only a significant gender effect, with males being more confident about their ability to succeed on a similar task than females, across the three TSP groups.

Subsequently, the role gender plays in the relationship between TCAP and DMIQ1 and DMIQ2 was investigated. For DMIQ1, the results revealed significant TCAP effects, with significant differences between the lowest, average and high TCAP groups, with the lowest DMIQ1 estimates provided by the group that solved fewest TCAP problems, average estimates by the average TCAP group and highest DMIQ1 estimates by the group that solved the most TCAP problems. Equally, significant gender effects revealed that males provided higher DMIQ1 estimates than females across the three groups.

Results for DMIQ2 were identical, with significant differences between the lowest, average, and highest TCAP groups and males providing higher DMIQ2 estimates in all three groups.

As in Study 12, these results provide support for the claim that individuals' ability self-insights are accurate and that they are capable of accurate self-assessments (e.g. Ackerman et al., 2002; Chamorro-Premuzic et al., 2010; Swim, 1994).

7.5. Study 14

Gender, TCAP, and Three TSP Conditions as Predictors of the Domain-Masculine Intelligence Type

7.5.1. Introduction

Study 14 builds on the previous three studies. It is similar in content and execution, with multiple measurement, psychometric task and assessment of task-success probability. Yet, it differs from earlier studies as it focuses on the role of task-success probability estimation or task confidence.

In fact, when Studies 11 and 12 used five TSP probes, TSP was the best predictor of DMIQ1 and DMIQ2. Yet, Study 13 that used only three TSP probes found gender as the best predictor of DMIQ1 and DMIQ2. Thus, it is possible that the reduction in TSP probes impacted on the observed results. In addition, the psychometric task in Study 13 differed from the earlier studies in that it included crystallised intelligence items and reduced the number of numerical and reasoning problems. This was done in order to accommodate participants' feedback as well as to test the assertions about the role of crystallised intelligence in DMIQ. Nevertheless, it is not clear whether the observed difference in the best predictor of the intelligence type was caused by the TSP probe decrease or the content change of the psychometric task.

Three experimental TSP conditions are introduced to investigate whether an increase, a decrease or a lack of the TSP probes will impact on the role TSP plays in the prediction of DMIQ1 and DMIQ2. Thus, the first condition increases the number of TPS probes to seven. The second condition decreases the number of TSP probes to four and the third condition uses zero probes. Each condition is assessed with an independent population sample. Study 14 reports the results of the combined total sample, i.e. one hundred and fifty-seven participants. The details of the three independent TSP condition analyses, i.e. Studies 14A, 14B and 14C, are reported in the Appendix (pp. 399-433).

As in precedent studies, HHE is expected to be observed on DMIQ1 and DMIQ2 (H1), while a significant decrease in the type estimates from T1 to T2 is predicted (H2). Gender is expected to influence the relationship between TSP and DMIQ1 (H3) and DMIQ2 (H4). Gender is also expected to influence the relationship between TCAP and DMIQ1 (H5) and DMIQ2 (H6). Gender is predicted as the best predictor of DMIQ1 (H7) and DMIQ2 (H8), over and above TSP and TCAP. Lastly, gender differences in DMIQ2 are expected to be observed in the TCAP intervention response, while DMIQ1 is controlled for (H9).

7.5.2. Method

Participants

One hundred and fifty-seven participants took part in this study. There were 81 females (52%) and 76 males. Their age ranged from 17 to 60 (M = 24.17, SD = 8.12) years. Participants were from the general public.

Measures

Repeated Measure of Domain-Masculine Intelligence Type (DMIQ)

See Study 11 (section 7.2.2).

Psychometric Aptitude Task

Total Correct Aptitude Problems (TCAP); Numerical and Spatial Psychometric Aptitude Problems (University of Kent, Career Services, 2009; http://www.kent.ac.uk/careers/test.htm)

Fifteen numerical reasoning and nine spatial problems that are in the public domain for online psychometric training purposes by the Career Services of University of Kent were adopted and used in the three TSP experimental conditions. The twenty-four problems were identical to the psychometric aptitude tests used by corporations in graduate recruitment processes for entry to graduate training programmes and job schemes. The problems were offered in two identical sections in all three TSP experimental conditions. In the seven TSP probe condition, the first fifteen numerical reasoning problems were broken down into three blocks of three problems, with each block followed by a TSP probe. The fourth block of six problems was followed by the fourth TSP probe. In the four TSP condition, only two TSP probes were asked. A time limit of 8 minutes was given. Participants were advised to leave unanswered problems blank, in order to not exceed the time limit, or face disqualification. The second section, which contained nine spatial problems, was offered in three blocks of three problems. In the seven TSP probe condition, three TSP probes were asked and in the four TSP condition two TSP probes were asked. A time limit of 5 minutes was given. Same instructions were used as in section one. Time limits were set to reflect a real-life testing situation, with sufficient limits to complete all problems. Participants were instructed to note their own answers as correct answers were given at the end of the survey. The number of correctly solved numerical reasoning and spatial psychometric aptitude problems, or Total Correctly Solved Aptitude Problems (TCAP) per individual was computed. Alpha in this study was .79.

Task Success Probability

Task Success Probability Estimation Measure (Storek, 2007)

See Study 11 (section 7.2.2).

7.5.3. Results

7.5.3.1. HHE on DMIQ1 and DMIQ2

Two independent samples t-tests were computed to assess whether significant gender differences occurred on DMIQ1 and DMIQ2. Results are presented in Table 7.5.1. HHE was confirmed for both DMIQ1 and DMIQ2 estimation conditions, with very large effect sizes. Hypothesis 1 was confirmed.

Table 7.5.1: Overview of Independent Samples t-Tests and Effect Sizes for DMIQ1 and DMIQ2

	Males M (SD) n	Females M (SD) N	*t(df)*	Mean Diff.	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
DMIQ1	119.32 (13.00) 76	104.48 (13.77) 81	6.93(155)***	14.83	10.61	19.06	**.24**	**1.11**
DMIQ2	115.09 (15.64) 76	95.24 (15.65) 81	7.94(155)***	19.85	14.91	24.78	**.29**	**1.27**

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Cohen's *d*. Large effect sizes are in bold.

To test whether a significant decrease in DMIQ estimates took place, a paired-samples t-test was conducted. There was a statistically significant decrease in DMIQ from T1 (M = 111.66, SD = 15.29) to T2 (M = 104.85, SD = 18.50), t(156) = 8.38, p = .00, two-tailed, r = .83, p =.00, N = 157. The mean decrease in DMIQ was 6.82 (SD = 10.20) with 95% confidence interval ranging from 5.21 to 8.42. The Cohen's *d* statistic (.67) indicated a medium effect size. Hypothesis 2 was confirmed.

7.5.3.2. Impact of Gender on the Relationship between TSP and DMIQ1 and DMIQ2

TSP was collapsed into a categorical variable with Group 1 containing individuals who were offered zero TSP probes and one individual who expressed no confidence in his/her ability to succeed on a similar task. Group 2 was made up of individuals that were offered four TSP probes and four individuals who had low confidence in their ability to succeed. Group 3 was made of individuals who were offered seven TSP probes or had high confidence in their ability to succeed on similar tasks in future. Results are presented in Table 7.5.2.

Table 7.5.2: Overview of TSP Banded

	TSP	n
Group 1	<=0	49
Group 2	1-16	65
Group 3	17+	43

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender affects the relationship between TSP and DMIQ1 and DMIQ2. Results are presented in Table 7.5.3.

Table 7.5.3: Two 2-way ANOVAs (TSP and gender) on DMIQ1 and DMIQ2

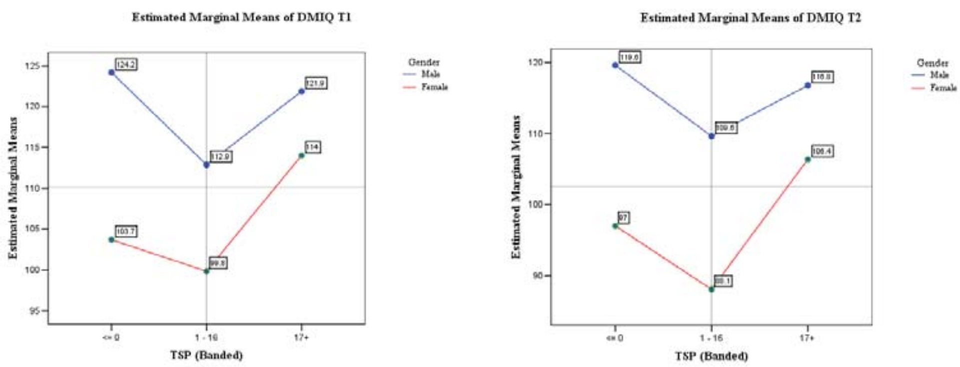
Variable	TSP Groups	Mean Score (SD)			*F*-score		
		Total	Males	Females	TSP	Gender	TSP x Gender
DMIQ1	G1 (L)	114.14 (17.37)	124.18 (14.39)	103.69 (13.78)	12.08***	46.67***	3.04
	G2 (M)	105.45 (11.73)	112.88 (8.43)	99.84 (10.78)			
	G3 (H)	118.22 (14.23)	121.87 (13.25)	114.03 (14.49)			
DMIQ2	G1 (L)	108.54 (21.56)	119.62 (21.93)	97.00 (13.98)	11.23***	58.96***	2.48
	G2 (M)	97.37 (16.17)	109.64 (9.80)	88.08 (13.67)			
	G3 (H)	104.85 (18.50)	115.09 (15.64)	95.24 (15.65)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

Note: DMIQ1 = Domain-Masculine Intelligence Type at pre-task estimation condition; DMIQ2 = Domain-Masculine Intelligence Type at post-task estimation condition. TSP = Task-success probability estimation condition.

For DMIQ1, the interaction effect between gender and TSP was not significant, F(2,151) = 3.04, $p > .05$, $\eta_p^2 = .04$. There was a statistically significant main effect for gender, F(1,151) = 46.67, $p = .00$, $\eta_p^2 = .24$, with large effect size and for TSP, F(2,151) = 12.08, $p = .00$, $\eta_p^2 = .14$, also with large effect size. Planned contrasts revealed significant differences between Group 2 and Group 3, (Contrast Estimate -11.59, $p = .00$). Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=0) was significantly different from Group 2. Group 2 mean score was also significantly different from Group 3 (17+). Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 3 was partially confirmed.

Figure 7.5.1: Two 2-way ANOVAs (TSP and Gender) on DMIQ1 and DMIQ2



For DMIQ2, the interaction effect between gender and TSP estimation conditions was not significant, F(2,151) = 2.48, $p = .09$, $\eta_p^2 = .03$. There was a statistically significant main effect for gender, F(1,151) = 58.96, $p = .00$, $\eta_p^2 = .28$, with large effect size, and for TSP, F(2,151) = 11.23, $p = .00$, $\eta_p^2 = .13$, with medium effect size. Planned contrasts revealed significant differences between Group 2 and Group 3, (Contrast Estimate -12.72, $p = .00$). Post-hoc comparisons using the Tukey

HSD and Bonferroni tests indicated that the mean score for Group 1 (<=0) was significantly different from Group 2 (1-16). Group 2 mean score was also significantly different from Group 3 (17+). Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 4 was partially confirmed.

7.5.3.3. Impact of Gender on the Relationship between TCAP and DMIQ1 and DMIQ2

TCAP was collapsed into a categorical variable with three groups, with Group 1 containing individuals who solved the lowest numbers of problems, Group 2 of individuals who solved average number of problems and Group 3 of individuals who correctly solved most problems. Results are presented in Table 7.5.4.

Table 7.5.4: Overview of TCAP Banded

	TCAP	n
Group 1	<=13	57
Group 2	14-17	57
Group 3	18+	43

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender influences the relationship between TCAP and DMIQ1 and DMIQ2. Results are presented in Table 7.5.5.

Table 7.5.5: Two 2-way ANOVAs (TCAP and gender) on DMIQ1 and DMIQ2

Variable	TCAP Groups	Mean Score (SD)			*F*-score		
		Total	Males	Females	TCAP	Gender	TCAP x Gender
DMIQ1	G1 (L)	106.25 (17.24)	122.38 (18.85)	99.95 (11.75)	1.54	39.57***	4.40**
	G2 (M)	112.78 (11.25)	115.86 (10.26)	108.54 (11.38)			
	G3 (H)	117.36 (15.11)	121.72 (11.19)	110.00 (18.17)			
DMIQ2	G1 (L)	96.96 (22.16)	117.81 (25.21)	88.83 (14.40)	1.95	53.88***	6.42***
	G2 (M)	107.33 (11.92)	110.62 (9.82)	102.81 (13.24)			
	G3 (H)	112.00 (16.79)	118.93 (13.24)	100.31 (15.91)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

Note: DMIQ1 = Domain-Masculine Intelligence Type at pre-task estimation condition; DMIQ2 = Domain-Masculine Intelligence Type at post-task estimation condition. TCAP = Total Correct Aptitude Problems.

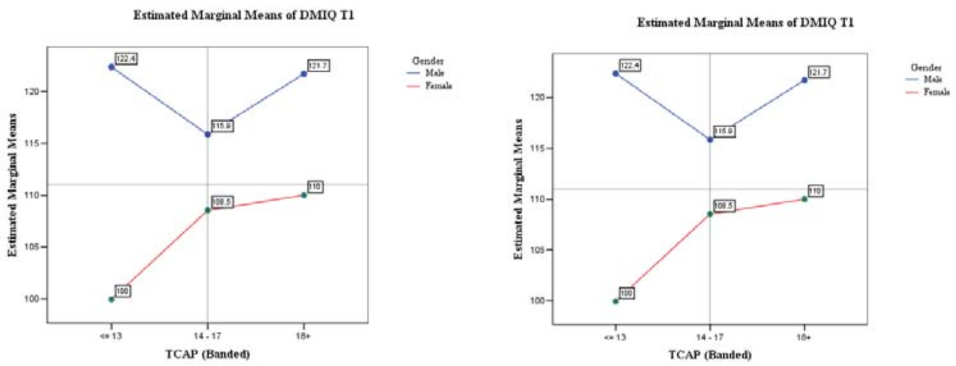
For DMIQ1, the interaction effect between gender and TCAP was significant, $F(2,151) = 4.40$, $p = .01$, $\eta_p^2 = .06$. The main effect for gender, $F(1,151) = 39.57$, $p = .00$, $\eta_p^2 = .21$ was also significant. The main effect for TCAP, $F(2,151) = 1.54$, $p = .22$, $\eta_p^2 = .02$ did not reach significance. Planned contrasts revealed no significant differences between the groups. Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=13) was significantly different from Group 2 (14-17) as well as from Group 3 (18+). This was confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. No other significant differences between the groups were observed.

As the main interaction effect was significant, further investigation of the relationship was warranted. Simple effects analysis was conducted. The data was split per gender and two one-way between-groups analyses of variance were conducted.

For males, the one-way between-groups analysis of variance for DMIQ1 was not significant, F $F(2,73) = 2.14$, $p = .13$, $\eta^2 = .05$. The robust tests of equality of means, Welch (2, 34) = 2.50, $p = .10$; Brown-Forsythe (2, 32) = 1.66, $p = .21$ were not significant. Post-hoc comparisons using the Tukey HSD and Bonferroni tests revealed no significant differences in mean scores between the three groups.

For females, there was a statistically significant difference in the three TCAP groups on DMIQ1, $F(2,78) = 5.00$, p=.01, η^2 =.11, with medium effect size. The robust tests of equality of means, Welch (2, 34) = 5.12, $p < .05$; Brown-Forsythe (2,35) = 4.07, $p < .05$ were significant. The post-hoc comparisons using the Tukey HSD and Bonferroni tests revealed significant differences between Group 1 (<=13) (M = 99.95, SD = 11.75) and Group 2 (14-17) (M = 108.54, SD = 11.38) and between Group 1 (<=13) and Group 3 (18+) (M = 104.48, SD = 13.77). No other significant mean score differences were observed between the groups. Hypothesis 5 was confirmed.

Figure 7.5.2: Two 2-way ANOVAs (TCAP and Gender) on DMIQ1 and DMIQ2



For DMIQ2, the interaction effect between gender and TCAP was also significant, $F(2,151) = 6.42$, $p = .00$, $\eta_p^2 = .08$, with medium effect size. The main effect for gender, $F(1,151) = 53.88$, p =.00, $\eta_p^2 = .26$ was significant. The main effect for TCAP, $F(2,151) = 1.95$, p =.15, $\eta_p^2 = .03$ was not significant. Planned contrasts revealed no significant differences between the groups. Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=13) was significantly different from Group 2 (14-17) as well as from Group 3 (18+).

As the main interaction effect was significant, further investigation of the relationship was warranted. Simple effects analysis was conducted. Data was split per gender and two one-way between-groups analysis of variance were conducted. The one-way between-groups analysis of variance for DMIQ2 was not significant for males, $F(2,73) = 2.50$, $p = .09$, $\eta^2 = .06$. The homogeneity of variance assumption was violated (Levene Statistic $p < .05$), indicating the groups variances were not equal. An alternative check for comparing variances was used. Firstly, the largest and the smallest standard deviations were squared. The largest squared SD was divided by the smallest squared SD, with resulting value of 6.59, which is bigger than the recommended value of 2, suggesting that the group variances were inadequate. The significance level was adjusted to $p < .01$. Post-hoc comparisons, using the Games-Howell procedure revealed significant differences between Group 2 (M = 110.62, SD = 9.82) and Group 3 (M = 119.93, SD = 13.24). The robust tests of equality of means, Welch (2, 32) = 3.84, $p < .05$; Brown-Forsythe (2,26) = 1.76, $p = .19$ revealed mixed results about the right to reject the null hypothesis.

For females, there was a statistically significant difference in the three TCAP groups on DMIQ2, $F(2,78) = 8.40$, $p = .00$, $\eta^2 = .18$, with large effect size. The post-hoc comparisons using Tukey HSD and Bonferroni tests revealed significant differences between Group 1 (<=13) (M = 88.83, SD = 14.40) and Group 2 (14-17) (M = 102.81, SD = 13.24) and between Group 1 (<=13) and Group 3 (18+) (M = 100.31, SD = 15.91). Hypothesis 6 was confirmed.

7.5.3.4. Gender as the best predictor of DMIQ1 and DMIQ2

To investigate whether gender was the best predictor of DMIQ1 and DMIQ2, two simultaneous multiple regressions were performed. The dependent variables were DMIQ1 and DMIQ2. Gender, TSP and TCAP were the independent variables. Results are presented in Table 7.5.6. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity.

The first model that used DMIQ1 as a dependent variable was significant ($F(3,153) = 18.56$, $p = .00$, Adjusted $R^2 = .25$, $f^2 = .37$), with the overall model explaining 27% of total variance. Gender ($\beta = -.44$, $p = .00$, $r_{part} = -.42$) and TCAP ($\beta = .16$, $p < .01$, $r_{part} = .15$) were significant predictors of DMIQ1,with gender accounting for 18% and TCAP for 2% of variance. TSP did not significantly contribute to the prediction. Thus, gender was confirmed as the best predictor of DMIQ1, in support of Hypothesis 7.

The second model that used DMIQ2 as the dependent variable was also significant ($F(3,153) = 24.46$, $p = .00$, Adjusted $R^2 = .31$, $f^2 = .47$), with the overall model explaining 32% of total variance. Gender ($\beta = -.48$, $p = .00$, $r_{part} = -.46$) and TCAP ($\beta = .11$, $p < .01$, $r_{part} = .18$) were significant predictors, explaining 21% and 3% of variance respectively. TSP again failed to reach significance. Gender was the best predictor of DMIQ2. Hypothesis 8 was supported.

Table 7.5.6: Beta coefficients for Simultaneous Multiple Regressions of Gender, TSP and TCAP onto DMIQ1 and DMIQ2

Dependent Variable N = 157	***DMIQ1*** β	*t*	***DMIQ2*** β	*t*
Gender	**-.44**	**-6.09*** **	**-.48**	**-6.97*** **
TSP	.07	.99	.03	.61
TCAP	**.16**	**2.22** **	**.11**	**2.72** **
Regression Model	**F(3, 153) = 18.56*** **		**F(3, 153) = 24.46*** **	
R^2	.27		.32	
R^2 Change	.27		.32	
Adj. R^2	.25		.31	
f^2	.37		.47	

Note: Significant values are in bold.

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

To further examine the relationships between these variables, correlations were computed with DMIQ1, DMIQ2, gender, TSP4, TSP7, TCAP and age. A correlational pattern similar to previous studies was observed, with strong negative correlations between gender and DMIQ1 (r =-.49, p=.00) and DMIQ2 (r =-.54, p =.00) and a medium negative correlation between gender and TCAP (r =-.27, p <.01), indicating that females provided lower DMIQ1 and DMIQ2 estimates and solved less TCAP problems. Medium positive correlations were also observed between TCAP and DMIQ1 (r =.34, p =.00) and DMIQ2 (r =.36, p =.00). Interestingly, the only significant correlation between the intelligence type and TSP was between TSP7 and DMIQ1 (r =.18, p.<.05). No significant relationships were observed for TSP7 and DMIQ2 and between TSP4 and DMIQ1 and DMIQ2. Thus, TSP probe reduction impacted the DMIQ relationship. In addition, TSP7 but not TSP4, correlated positively with TCAP (r =.21, p<.01). Age did not significantly contribute to the analysis.

7.5.3.5. Role of Gender in DMIQ2 in Response to TCAP Task

A 2-by-2 between-groups analysis of covariance was conducted to assess the influence of TCAP on DMIQ2 for males and females. The independent variables were the three TCAP groups and gender. The dependent variable was DMIQ2. DMIQ1 was used as a covariate to control for individual differences. Preliminary checks were conducted to ensure that there was no violation of the assumptions of normality, linearity, homogeneity of variances, homogeneity of regression slopes, and reliable measurement of the covariate. The homogeneity of regression slopes assumption was not violated for the TCAP by DMIQ1 interaction, $F(2,149) = .35$, $p=.71$ as well as for the gender by DMIQ1 interaction, $F(1,149) = 1.74$, $p = .19$.

After adjusting for DMIQ1 estimates, there was a non-significant interaction effect, $F(2,150) = 2.49$, $p = .09$, $\eta_p^2 = .03$. The main effect for TCAP was also not significant, $F(2,150) = .93$, $p = .40$, $\eta_p^2 = .01$. The main effect for gender was significant, $F(1,150) = 12.26$, $p = .00$, $\eta_p^2 = .08$, with medium effect size. The main effect for the covariate variable DMIQ1 was also significant, $F(1,150) = 200.76$, $p = .00$, $\eta_p^2 = .57$, with large effect size and the covariate significantly related to DMIQ2. Planned contrasts revealed no significant group means differences for TCAP, but significant differences between males and females (Contrast Estimate = 6.50, $p < .01$) in DMQI T2 were observed, with males providing higher self-estimates of ability (Group 1: $M_{Male} = 117.81$, $SD_{Male} = 25.21$; $M_{Female} = 88.83$, $SD_{Female} = 14.40$; Group 2: $M_{Male} = 110.62$, $SD_{Male} = 9.82$; $M_{Female} = 102.81$, $SD_{Female} = 13.24$; Group 3: $M_{Male} = 118.93$, $SD_{Male} = 13.24$; $M_{Female} = 100.31$, $SD_{Female} = 15.91$). The results confirm that gender, and in particular male hubris, plays a role in DMIQ2 but TCAP does not. Equally, DMIQ1 contributes to DMIQ2 estimations. Hypothesis 9 was partially confirmed.

7.5.4. Discussion

This study differed from earlier studies in that it set out to examine the role TSP probes play in the estimation process. Studies 11 and 12 that used seven TSP probes found TSP as the best predictor of the intelligence type at both T1 and T2. However, Study 13 that used only three TSP probes confirmed gender as the best predictor of DMIQ1 and DMIQ2. In addition, the TCAP content in Study 13 was changed by including crystallised intelligence items and reducing numerical and reasoning items. To ascertain whether gender's role as the best predictor of the intelligence type was caused by the decrease in TSP probes or TCAP content change, the three experimental conditions for TSP were designed, with seven, four and zero TSP probes. TCAP content was adapted to resemble the task from Studies 11 and 12. The reported results are for the combined overall sample, while the three experimental conditions are reported in the Appendix.

As in all previous studies, the existence of HHE on DMIQ1 (η^2 =.24, d =1.11) and DMIQ2 (η^2 =.29, d =1.27) was affirmed. Equally, a significant decrease in DMIQ estimates was observed from the pre-task to post-task estimation condition (d =.67).

The role that gender plays in the relationship between TSP and DMIQ1 and DMIQ2 was investigated next. For DMIQ1 the results revealed a significant TSP effect, with significant differences between the three TSP groups, with highest DMIQ1 estimates provided by the most confident group, average DMIQ1 estimates by group with lowest confidence and lowest DMIQ1 estimates by the group with average confidence. As in previous studies, males provided significantly higher DMIQ1 estimates than females across all three groups.

For DMIQ2, the results revealed a significant TSP effect, with significant differences between the three TSP groups, with highest DMIQ2 estimates provided by the group with lowest TSP scores, lowest DMIQ2 estimates by the average TSP group and average DMIQ2 estimates by group with highest TSP scores. Males' scores were higher than females' across all three groups. These findings provide mixed results about participants' ability to accurately estimate their abilities but further support for the existence of male hubris in SEI.

Subsequently, the role gender plays in the relationship between TCAP and DMIQ1 and DMIQ2 was investigated. For DMIQ1, the results revealed a significant interaction effect between TCAP and gender and a significant gender effect. Significant differences were observed for males, with highest DMIQ1 estimates provided by the group that solved the least TCAP problems, lowest DMIQ1 estimates by the average TCAP problem solving group and average DMIQ1 estimates by the group solving most TCAP problems. For females a different pattern was observed, with lowest DMIQ1 estimates provided by group that solved fewest problems, average estimates by group that solved average number of TCAP problems and highest DMIQ1 estimates by group that solved most TCAP problems. As in all previous analyses, males provided higher DMIQ1 estimates than females across the three groups.

For DMIQ2 the results also revealed a significant interaction between TCAP and gender and a significant gender effect. For males, the findings showed that highest DMIQ2 estimates were provided by the group that solved the most TCAP problems, average estimates by the group that solved the fewest TCAP problems and lowest DMIQ2 estimates by the group that solved the average number of TCAP problems. For females, lowest DMIQ2 estimates were provided by the group solving the fewest TCAP problems, average DMIQ2 estimates by the group solving the most

TCAP problems and highest DMIQ2 estimates by the average TCAP group. Equally, males provided higher DMIQ2 estimates in all three groups.

These findings, albeit complex and varied, provide support for the assertion that females, despite holding more self-handicapping and negative views and having lower self-confidence (Beyer, 1998; Ehrlinger & Dunning, 2003; Sleeper & Nigro, 1987) are more accurate judges of their performance, especially on gender-stereotyped tasks (e.g. Carr et al., 2008). These results also support the claims that most overconfident estimations occur on difficult tasks (Dunning et al., 1990; Jonsson & Allwood, 2003). The results are further supported by the findings that gender, and in particular the male hubris, plays a role in DMIQ2 but TCAP does not.

As in previous studies, gender was expected to be the best predictor of DMIQ. Results confirmed gender as the best predictor of DMIQ1 and DMIQ2, explaining 18% and 21% of variance respectively. As in Study 13, TSP did not play a role in the prediction of the intelligence type, but TCAP did. Thus, it seems that the TSP manipulation negatively affected the perceived importance of task confidence and the previous content change in TCAP did not. In Study 14 TCAP was qua content similar to problems used in Study 11 and 12. Moreover, positive relationships were observed between TCAP and DMIQ1 and DMIQ2. TSP7, i.e. the condition using seven probes, only correlated with DMIQ1 and no relationship was observed between TSP4 and DMIQ1 and DMIQ2, providing further evidence about the impact of TSP manipulation.

7.6. Study 15

Gender, TSP, TCAP as Predictors of the Domain-Masculine Intelligence Type

7.6.1. Introduction

The last experimental study aims to confirm the earlier findings as well as further examine whether content change in the psychometric task (TCAP) will impact on the results (see Study 13).

Thus, as in Studies 11 to 14, HHE is expected to be observed on both DMIQ1 and DMIQ2 (H1). A significant decrease in DMIQ estimates from T1 to T2 is also predicted (H2). In order to validate findings of Studies 11 to 13, gender differences are expected on the numerical and spatial psychometric problems, with males solving more problems correctly than females (H3). Based on the results of Study 13 and the lack of agreement about sex differences in general intelligence (e.g. Colom & Garcia-Lopez, 2002; Halpern et al., 2007; Lynn, 1999; Spelke, 2005), gender differences were not expected on the crystallised intelligence (Gc) items (H4). In order to corroborate the findings of Studies 11, 12 and 13, gender differences were expected to be observed in TSP, with males more task confident than females (H5).To validate the findings of Study 13 and 14, gender was hypothesised as the best predictor of DMIQ1 (H6) and DMIQ2 (H7), over and above TSP and TCAP. Gender is also expected to moderate the relationship between TSP and DMIQ1 (H8) and DMIQ2 (H9) and between TCAP and DMIQ1 (H10) and DMIQ2 (H11).

7.6.2. Method

Participants

Fifty-four participants took part in this study. There were 27 females (50%) and 27 males. Their age ranged from 21 to 50 (M = 35.39, SD = 6.89) years. 59% of the participants were French native speakers, 28% were English native speakers, 4% were native Chinese speakers and 4% were Brazilian Portuguese native speakers, 2% were native Arabic speakers, 2% were native Swedish speakers and 2% were native Guajarati native speakers. 63% of participants completed MA/MSc level of education, 19% completed a MBA level of education, 11% achieved BA/BSc level of education 6% had a Phd/Doctorate degree and 2% stated to have earned a non-university professional degree.

Measures

Repeated Measure of Domain-Masculine Intelligence Type (DMIQ)

See Study 11 (section 7.2.2) Alpha for the DMIQ1 was .83 and DMIQ2 .85.

Psychometric Aptitude Task

Total Correct Aptitude Problems (TCAP); Numerical Reasoning, Spatial and Crystallised Intelligence Aptitude Problems (University of Kent, Career Services, 2009; http://www.kent.ac.uk/careers/test.htm)

Fifteen aptitude problems that are in the public domain for online psychometric training purposes by the Career Services of University of Kent were adopted and used. Ten numerical reasoning and spatial problems, used in Studies 11 to 14, were used. Five crystallised intelligence (Gc) questions were adopted from the General Knowledge Test (Irwing, Cammock, and Lynn, 2001), covering general knowledge, science, literature, geography.

The problems were offered in three blocks of five problems. For the first section, containing five numerical reasoning problems, a five-minute time limit was given. For the second section, containing five spatial problems, a four-minute limit and for the third section with five Gc questions, a two-minute limit. Each section was followed by a TSP probe. Participants were advised to leave unanswered problems blank, in order to not exceed the time limit, or face disqualification.

As with the other online surveys, time limits were set to reflect a real-life testing situation. Participants were instructed to note their answers and check the correct answers at the end of the survey. The number of correctly solved aptitude problems, or Total Correct Aptitude Problems (TCAP) per individual was computed. The alpha for the fifteen numerical reasoning items was .66 and the inter-item correlation was .11. Alpha for the five numerical reasoning problems was .62 and the inter-item correlation was r =.26. Alpha for the five spatial problems was .35 and the inter-item correlation r =.10. Alpha for the five crystallised intelligence items was .58 and the inter-item r =.22. Alpha for the Total TCAP was .83. For the overview of the problems see Appendix.

Task Success Probability (TSP)

Task Success Probability Estimation Measure (Storek, 2007)

See Study 11 (section 7.2.2). The alpha for the three-item measure was .69 and the inter-item correlation was r= .43.

Procedure

Participants were from the general public and were all members of the social networking site Facebook. The main researcher sent an email invitation to the survey to her Facebook contacts, with background explanation of the survey and the URL (www.zoomerang.com). Participants were also requested to share the URL link with their 'Friends'. Participation was voluntary and participants were aware that they could withdraw from the study at any moment. Participants were given detailed scoring instructions at the beginning of each measure, including timing instructions. Participants were aware that the study was approved by UCL Ethics Committee, meeting confidentiality and Data Protection requirements.

Debrief feedback, correct answers to the aptitude problems, together with a feedback box, were provided at the end of the survey. The study was done over the course of four weeks in July 2010. In total, 138 individuals logged onto the site during this period and 54 completed the survey. All participants were fluent in English and no language or other problems were reported.

7.6.3. Results

7.6.3.1. HHE on DMIQ1 and DMIQ2

Two independent samples t-tests were computed to assess whether significant gender differences were observed on DMIQ1 and DMIQ2. Results are presented in Table 7.6.1. Significant gender differences, with males providing higher self-estimates than females, were observed on both the pre-task and post-task estimation conditions, further affirming the existence of HHE on DMIQ. Hypothesis 1 was confirmed.

To test whether significant change in DMIQ estimates occurred following the task, a paired-samples t-test was conducted. There was a statistically significant decrease in DMIQ from T1 (M = 106.91, SD = 15.40) to T2 (M = 102.41, SD = 16.96), t(53) = 4.50, $p < .01$, two-tailed, $r = .75$, p =.00, N=54. The mean decrease in DMIQ estimates was 4.50 (SD = 11.65) with 95% confidence interval ranging from

1.32 to 7.68. Cohen's *d* (.39) indicated a small effect size. Hypothesis 2 was confirmed.

Table 7.6.1:Overview of Independent Samples t-Tests and Effect Sizes for DMIQ1 and DMIQ2

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
DMIQ1	112.96 (11.99) 27	100.85 (16.23) 27	3.12(52)**	12.11	4.32	19.90	**.16**	**.85**
DMIQ2	108.15 (13.24) 27	96.67 (18.52) 27	2.62(52)*	11.48	2.69	20.27	.12	.78

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Cohen's *d*. Large effect sizes are in bold.

7.6.3.2. Gender Differences in TCAP and TSP

Table 7.6.2. gives an overview of the 2x2 χ^2 tests and effect sizes for the fifteen psychometric aptitude problems. Significant gender differences were observed only on two problems, i.e. Q8 and Q11 with males providing more correct answers than females. Medium sized negative significant effect sizes were observed on Q8, $\varphi = -.34$, $p <.05$ and on Q11, $\varphi = -.35$, $p <.01$.

Table 7.6.2: 2 x 2 Chi Square Tests and Effect Sizes for 5 Numerical Reasoning, 5 Spatial and 5 Crystallised intelligence (Gc) Problems – Per Gender and % Correct Answer

		Correct Answer Wrong	Correct Answer Right	Total	Yates Continuity Correction Value for 2x2	Asymp. Sig. (2-sided)	Phi (φ) Coefficient
Block 1 Numerical Reasoning Q8	Male N	1	26	27	1.69	.19	-.24
	% Within Gender	4%	96%	100%			
	% Within Correct Answer	17%	54%	50%			
	% of Total	2%	48%	50%			
	Female N	5	22	27			
	% Within Gender	19%	82%	100%			
	% Within Correct Answer	83%	46%	50%			
	% of Total	9%	41%	50%			
	Total N	6	48	54			
	% Within Gender	11%	89%	100%			
Q9	Male N	6	21	27	4.99	.03*	-.34*
	% Within Gender	22%	78%	100%			
	% Within Correct Answer	29%	64%	50%			
	% of Total	11%	39%	50%			
	Female N	15	12	27			
	% Within Gender	56%	44%	100%			
	% Within Correct Answer	71%	36%	50%			
	% of Total	28%	22%	50%			
	Total N	21	33	54			
	% Within Gender	39%	61%	100%			
Q10	Male N	14	13	27	1.99	.16	-.23
	% Within Gender	52%	48%	100%			
	% Within Correct Answer	41%	65%	50%			
	% of Total	26%	24%	50%			
	Female N	20	7	27			
	% Within Gender	74%	26%	100%			
	% Within Correct Answer	59%	35%	50%			
	% of Total	37%	13%	50%			
	Total N	34	20	54			
	% Within Gender	63%	37%	100%			
Q11	Male N	0	27	27	4.69	.03	-.35**
	% Within Gender	0%	100%	100%			
	% Within Correct Answer	0%	56%	50%			
	% of Total	0%	50%	50%			
	Female N	6	21	27			
	% Within Gender	22%	78%	100%			
	% Within Correct Answer	100%	44%	50%			
	% of Total	11%	39%	50%			
	Total N	6	48	54			
	% Within Gender	11%	89%	100%			
Q12	Male N	4	23	27	2.41	.12	-.25
	% Within Gender	15%	85%	100%			
	% Within Correct Answer	29%	58%	50%			
	% of Total	73%	43%	50%			
	Female N	10	17	27			
	% Within Gender	37%	63%	100%			
	% Within Correct Answer	71%	43%	50%			
	% of Total	19%	32%	50%			

	Total N	14	40	54			
	% Within Gender	26%	74%	100%			
Block 2 Spatial Q14	Male N	12	15	27	1.19	.28	-.19
	% Within Gender	44%	56%	100%			
	% Within Correct Answer	41%	60%	50%			
	% of Total	22%	28%	50%			
	Female N	17	10	27			
	% Within Gender	63%	37%	100%			
	% Within Correct Answer	59%	40%	50%			
	% of Total	32%	19%	50%			
	Total N	29	25	54			
	% Within Gender	54%	46%	100%			
Q15	Male N	9	18	27	.00	1.00	.04
	% Within Gender	33%	67%	100%			
	% Within Correct Answer	53%	49%	50%			
	% of Total	17%	33%	50%			
	Female N	8	19	27			
	% Within Gender	30%	70%	100%			
	% Within Correct Answer	47%	51%	50%			
	% of Total	15%	35%	50%			
	Total N	17	37	54			
	% Within Gender	32%	69%	100%			
Q16	Male N	21	6	27	.37	.54	.12
	% Within Gender	78%	22%	100%			
	% Within Correct Answer	54%	40%	50%			
	% of Total	39%	11%	50%			
	Female N	18	9	27			
	% Within Gender	67%	33%	100%			
	% Within Correct Answer	46%	60%	50%			
	% of Total	33%	17%	50%			
	Total N	39	15	54			
	% Within Gender	72%	28%	100%			
Q17	Male N	11	16	27	.67	.41	-.15
	% Within Gender	41%	59%	100%			
	% Within Correct Answer	42%	57%	50%			
	% of Total	20%	30%	50%			
	Female N	15	12	27			
	% Within Gender	56%	44%	100%			
	% Within Correct Answer	58%	43%	50%			
	% of Total	28%	22%	50%			
	Total N	26	28	54			
	% Within Gender	48%	52%	100%			
Q18	Male N	4	23	27	.46	.50	-.14
	% Within Gender	15%	85%	100%			
	% Within Correct Answer	36%	54%	50%			
	% of Total	7%	43%	50%			
	Female N	7	20	27			
	% Within Gender	26	74%	100%			
	% Within Correct Answer	64%	47%	50%			
	% of Total	13%	37%	50%			
	Total N	11	43	54			
	% Within Gender	20%	80%	100%			
Block 3 Gc Q20	Male N	21	6	27	1.32	.25	-.21
	% Within Gender	78%	22%	100%			
	% Within Correct Answer	46%	75%	50%			
	% of Total	39%	11%	50%			

	Female N	25	2	27			
	% Within Gender	93%	7%	100%			
	% Within Correct Answer	54%	25%	50%			
	% of Total	46%	4%	50%			
	Total N	46	8	54			
	% Within Gender	85%	15%	100%			
Q21	Male N	17	10	27	.34	.56	-.12
	% Within Gender	63%	37%	100%			
	% Within Correct Answer	46%	59%	50%			
	% of Total	32%	19%	50%			
	Female N	20	7	27			
	% Within Gender	74%	26%	100%			
	% Within Correct Answer	54%	41%	50%			
	% of Total	37%	13%	50%			
	Total N	37	17	54			
	% Within Gender	69%	32%	100%			
Q22	Male N	6	21	27	.00	1.00	-.04
	% Within Gender	22%	78%	100%			
	% Within Correct Answer	46%	51%	50%			
	% of Total	11%	39%	50%			
	Female N	7	20	27			
	% Within Gender	26%	74%	100%			
	% Within Correct Answer	54%	49%	50%			
	% of Total	13%	37%	50%			
	Total N	13	41	54			
	% Within Gender	24%	76%	100%			
Q23	Male N	11	16	27	.08	.78	.08
	% Within Gender	41%	59%	100%			
	% Within Correct Answer	55%	47%	50%			
	% of Total	20%	30%	50%			
	Female N	9	18	27			
	% Within Gender	33%	67%	100%			
	% Within Correct Answer	45%	53%	50%			
	% of Total	17%	33%	50%			
	Total N	20	34	54			
	% Within Gender	37%	63%	100%			
Q24	Male N	8	19	27	.00	1.00	-.04
	% Within Gender	30%	70%	100%			
	% Within Correct Answer	47%	51%	50%			
	% of Total	15%	35%	50%			
	Female N	9	18	27			
	% Within Gender	33%	67%	100%			
	% Within Correct Answer	53%	49%	50%			
	% of Total	17%	33%	50%			
	Total N	17	37	54			
	% Within Gender	32%	68%	100%			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

In order to test whether significant gender differences occurred on TCAP, the three blocks, i.e. numerical, spatial and crystallised as well as on Total TSP and the three independent TSP probes, independent samples t-tests were computed. Results are presented in Table 7.6.3. Significant gender differences occurred on TCAP, with males correctly solving more problems than did females. In addition, males also provided significantly more correct answers to the numerical problems than did females. Contrary to prediction, no significant gender differences were observed on the spatial problems. Hypothesis 3 was partially confirmed. In agreement with Hypothesis 4, no significant gender differences were observed on the crystallised intelligence problems, providing further support to the claim that no sex differences exist in general intelligence.

Table 7.6.3: Overview of Independent t-Tests and Effect Sizes for TCAP and 3 Blocks of Psychometric Aptitude Problems and for Total TSP and 3 TSP Probes

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
TCAP	9.63 (2.39) 27	7.93 (2.92) 27	2.35(52)*	1.70	.25	3.16	.10	.64
NR	4.07 (.83) 27	2.93 (1.47) 27	3.54(41)**	1.15	.49	1.80	**.23**	**.96**
Sp	2.89 (1.25) 27	2.59 (1.22) 27	.88(52)	.30	-.38	.97	.02	.24
Gc	2.67 (1.49) 27	2.41 (1.22) 27	.70(52)	.26	-.49	1.00	.00	.19
TSP	9.22 (1.87) 27	8.74 (3.02) 27	.71(52)	.48	-.89	1.85	.01	.19
TSP 1	3.52 (.75) 27	3.07 (1.11) 27	1.73(52)	.44	-.07	.96	.05	.48
TSP 2	2.78 (1.05) 27	3.00 (1.78) 27	-.73(52)	-.22	-.83	.39	.01	.15
TSP 3	2.93 (1.00) 27	2.67 (1.23) 27	.86(52)	.26	-.35	.86	.01	.23

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Cohen's *d*. Large effect sizes are in bold. NR = Numerical Reasoning; Sp = Spatial, Gc = Crystallised Intelligence problems.

Next, independent samples t-tests were run for the Total TSP and the three independent TSP probes to investigate whether significant gender differences occurred on these variables. Results are presented in Table 7.6.3. Contrary to prediction, no significant gender differences were observed on the total TSP measure or the individual TSP probes. Hypothesis 5 was not confirmed.

7.6.3.3. Gender, TSP and TCAP as Predictors of DMIQ1 and DMIQ2

Firstly, the relationship between the DMIQ1 and DMIQ2, gender, TSP, and TCAP was investigated. Results of the correlational analysis are presented in Table 7.6.4.

Table 7.6.4: Correlations, Means and Standard Deviations between DMIQ1, DMIQ2, Gender, TSP, TCAP and Age

	DMIQ1	DMIQ2	G	TSP	TCAP	A
	106.91	102.41	1.50	8.98	3.54	
	(15.40)	(16.96)	(.51)	(2.50)	(.69)	
DMIQ1						
DMIQ2	.75***					
Gender	-.40**	-.34*				
TSP	.31*	.47***	-.10			
TCAP	.22	.34*	-.31*	.37**		
Age	.14	.05	.08	-.05	-.07	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).   *N = 54.*

As in previous studies, DMIQ1 and DMIQ2 were strongly intercorrelated (r = .75, p =.00). Equally, negative correlations were observed between gender and DMIQ1 (r = -.40, p<.01) and DMIQ2 (r = -.34, p <.05), with females providing lower scores than males. Positive correlations were observed between TSP and DMIQ1 (r = .31, p <.05) and DMIQ2 (r =.47, p =.00). No significant relationship was observed between TCAP and DMIQ1 but a positive relationship was observed between TCAP and DMIQ2 (r =.34, p <.05) as well as between TCAP and TSP (r =.37, p <.01). A negative correlation between TCAP and gender (r =-.31, p <.05) suggested that females solved correctly less psychometric problems than males. As in earlier studies, age was included to examine its role in DMIQ estimation. Age did not significantly correlate with any of the variables.

To investigate whether gender was the best predictor of DMIQ1 and DMIQ2, two simultaneous multiple regressions were computed. The dependent variables were DMIQ1 and DMIQ2. Gender, TSP and TCAP were the independent variables. Results are reported in Table 7.6.5.

Gender, TSP and TCAP were regressed on DMIQ1 and DMIQ2. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity.

The first model was significant $F(3,50) = 5.09$, $p < .01$, Adjusted $R^2 = .19$, $f^2 = .30$), with the overall model explaining 23% of total variance. Gender ($\beta = -.37$, $p < .01$, $r_{part} = -.35$) and TSP ($\beta = .28$, $p < .05$, $r_{part} = .26$) were significant predictors of DMIQ1,with gender accounting for 12% and TSP for 7% of variance. Gender was the best predictor of DMIQ1, followed by TSP. TCAP did not significantly contribute to the prediction. Hypothesis 6 was confirmed.

The second model was also significant $F(3,50) = 7.68$, $p = .00$, Adjusted $R^2 = .27$, $f^2 = .47$), with the overall model explaining 32% of total variance. TSP ($\beta = .40$, $p < .01$, $r_{part} = .37$) and gender ($\beta = -.27$, $p < .05$, $r_{part} = -.26$) were significant predictors, explaining 14% and 7% of variance respectively. Thus, TSP was the best predictor of DMIQ2, followed by gender. As in DMIQ1, TCAP did not significantly contribute to the prediction. Hypothesis 7 was not confirmed.

Table 7.6.5: Beta coefficients for Simultaneous Multiple Regressions of Gender, TSP and TCAP – Total and 3 Blocks of 5 Problems onto DMIQ1 and DMIQ2

Dependent Variable N = 54	***DMIQ1*** *β*	*t*	***DMIQ2*** *β*	*t*
Gender	**-.37**	**-2.83****	**-.27**	**-2.19***
TSP	**.28**	**2.07***	**.40**	**3.16****
TCAP	.01	.03	.11	.83
Regression Model	**F(3, 50) = 5.09****		**F(3,50) = 7.68*****	
R^2	.23		.32	
R^2 Change	.23		.32	
Adj. R^2	.19		.27	
f^2	.30		.47	
Dependent Variable N = 54	***DMIQ1*** *β*	*t*	***DMIQ2*** *β*	*t*
Gender	**-.31**	**-2.22***	-.20	-1.51
TSP	**.28**	**2.10***	**.39**	**3.06****
Numerical Reasoning	.18	1.07	.27	1.74
Spatial	-.17	-1.14	-.05	-.34
Crystallised Knowledge	.02	.14	-.06	-.49
Regression Model	**F(5, 48) = 3.38***		**F(5,48) = 5.22****	
R^2	.26		.35	
R^2 Change	.26		.35	
Adj. R^2	.18		.29	
f^2	.35		.54	

Note: Significant values are in bold.

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

As TCAP's content was changed and included crystallised intelligence measures, as in Study 13, the impact of TCAP on DMIQ1 and DMIQ2 was the focus of this study. Yet, TCAP was not a significant predictor in the previous regressions. Thus, the variable was split into three TCAP blocks, i.e. numerical, spatial and crystallised intelligence (Gc) to investigate whether they were significant predictors of the intelligence type. The simultaneous multiple regressions were re-computed for DMIQ1 and DMIQ2. Results are presented in Table 7.6.5.

The first model was significant $F(5,48) = 3.38$, $p < .05$, Adjusted $R^2 = .18$, $f^2 = .35$), with the overall model explaining 26% of variance. Gender ($\beta = -.31$, $p < .05$, $r_{part} = -.28$) and TSP ($\beta = .28$, $p < .05$, $r_{part} = .26$) were significant predictors of DMIQ1,with gender accounting for 8% and TSP for 7% of variance. As in the previous regression for DMIQ1, gender was affirmed as the best predictor of DMIQ1, follwed by TSP. Beta values were smaller than in the previous analysis. None of the three TCAP blocks reached significance.

The second model was also significant $F(5,48) = 5.22$, $p <.01$, Adjusted $R^2 = .29$, $f^2 = .54$), with the overall model explaining 35% of variance. Task-success prediction ($\beta = .39$, $p <.01$, $r_{part} = .36$) was the only significant predictor, explaining 13% of variance. As with the previous regression for DMIQ2, TSP was the best predictor. Gender or the three TCAP blocks did not significantly contribute to the prediction.

7.6.3.4. Impact of Gender and TSP on DMIQ1 and DMIQ2

TSP was collapsed into a categorical variable with Group 1 made of individuals with lowest task-confidence, Group 2 of individuals with average task-confidence and Group 3 of individuals with highest task-confidence. Results are presented in Table 7.6.6.

Table 7.6.6: Overview of TSP Banded

	TSP	n
Group 1	<=8	21
Group 2	9	15
Group 3	10+	18

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender influences the relationship between TSP and DMIQ1 and DMIQ2. Results are presented in Table 7.6.7.

For DMIQ1, the interaction effect between gender and TSP estimation conditions was not significant, $F(2,48) = 1.99$, $p = .50$, $\eta_p^2 = .08$. There was a statistically significant main effect for gender, $F(1,48) = 9.65$, $p <.01$, $\eta_p^2 = .17$, with large effect size. The main effect for TSP, F(2,48) = .74, p =.48, η_p^2 = .03 was not significant. Planned contrasts revealed no significant differences between the groups. Post-hoc comparisons using the Tukey HSD and Bonferroni tests revealed no significant differences in mean scores between the three TSP groups. This was confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets.

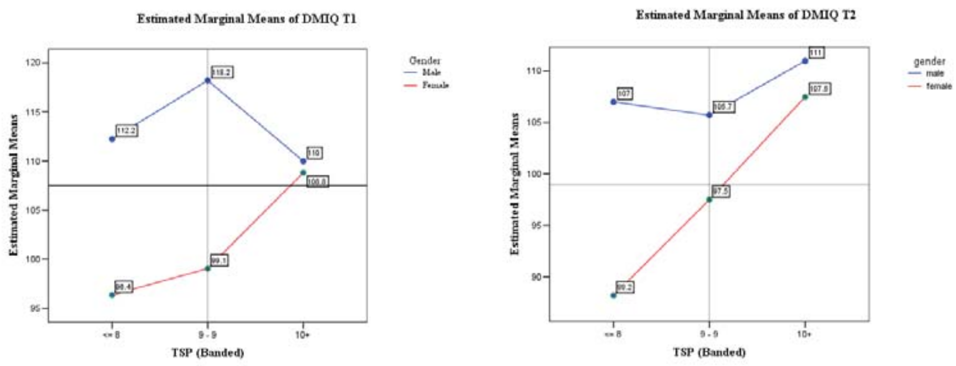
Table 7.6.7: Two 2-way ANOVAs (TSP and gender) on DMIQ1 and DMIQ2

Variable	TSP Groups	Mean Score (SD)			*F*-score		
		Total	Males	Females	TSP	Gender	TSP x Gender
DMIQ1	G1 (L)	103.93 (14.27)	112.25 (4.63)	96.36 (15.98)	.74	9.65**	1.99
	G2 (M)	108.00 (20.62)	118.21 (18.69)	99.06 (18.85)			
	G3 (H)	109.47 (11.44)	110.00 (11.37)	108.81 (12.27)			
DMIQ2	G1 (L)	97.14 (17.93)	107.00 (12.29)	88.18 (17.93)	2.75	5.64*	1.24
	G2 (M)	101.33 (17.01)	111.00 (10.42)	107.50 (17.93)			
	G3 (H)	109.44 (13.89)	110.00 (10.42)	107.50 (17.93)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: DMIQ1 = Domain-Masculine Intelligence Type at pre-task estimation condition; DMIQ2 = Domain-Masculine Intelligence Type at post-task estimation condition. TSP = Task-success probability estimation condition.

For DMIQ2, the interaction effect between gender and TSP estimation conditions was not significant, F(2,48) = 1.24, p = .30, η_p^2 = .05. There was a statistically significant main effect for gender, F(1,48) = 5.64, p <.05, η_p^2 = .11, with medium effect size. The main effect for TSP, F(2,48) = 2.75, p =.07, η_p^2 = .10 did not reach significance. Planned contrasts revealed no significant differences between the groups. Post-hoc comparisons using the Tukey HSD test revealed that the mean score for Group 1 (<= 8) was significantly different from Group 3 (10+). However, the more stringent Bonferroni test revealed no significant differences between the means of the three groups and no differences were found on the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypotheses 8 and 9 were partially confirmed.

Figure 7.6.1: Two 2-way ANOVAs (TSP and Gender) on DMIQ1 and DMIQ2



7.6.3.5. Impact of Gender and TCAP on the DMIQ1 and DMIQ2

TCAP was collapsed into a categorical variable with Group 1 made of individuals who solved fewest psychometric problems, Group 2 of individuals who solved an average number of problems and Group 3 of individuals who solved the most problems. Results are presented in Table 7.6.8.

Table 7.6.8: Overview of TCAP Banded

	TCAP	n
Group 1	<=8	23
Group 2	9-10	15
Group 3	11+	16

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender affects the relationship between TCAP and DMIQ1 and DMIQ2. Results are presented in Table 7.6.9.

For DMIQ1, the interaction effect between gender and TCAP was significant, $F(2,48) = 3.43$, $p < .05$, $\eta_p^2 = .13$, with medium effect size. The main effect for gender, F(1,48) = 4.49, p <.05, η_p^2 = .09 was also significant, with medium effect size. The main effect for TCAP, F(2,48) = 1.51, p =.23, η_p^2 = .06 was not significant. Planned contrasts revealed no significant differences between the groups. Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=8) was significantly different from Group 2 (9-10). This was confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. No other significant mean score differences between the groups were observed.

Table 7.6.9: Two 2-way ANOVAs (TCAP and gender) on DMIQ1 and DMIQ2

Variable	TCAP Groups	Mean Score (SD)			*F*-score		
		Total	Males	Females	TCAP	Gender	TCAP x Gender
DMIQ1	G1 (L)	100.65 (17.29)	115.36 (16.36)	94.22 (13.65)	1.51	4.49*	3.43*
	G2 (M)	112.50 (12.50)	111.39 (11.73)	114.17 (14.55)			
	G3 (H)	110.66 (12.13)	112.73 (9.84)	106.10 (16.05)			
DMIQ2	G1 (L)	92.93 (16.87)	105.71 (10.97)	87.34 (16.11)	4.69*	1.53	2.83
	G2 (M)	108.50 (13.95)	106.94 (15.80)	110.83 (11.58)			
	G3 (H)	102.41 (16.96)	108.15 (13.24)	96.67 (18.52)			

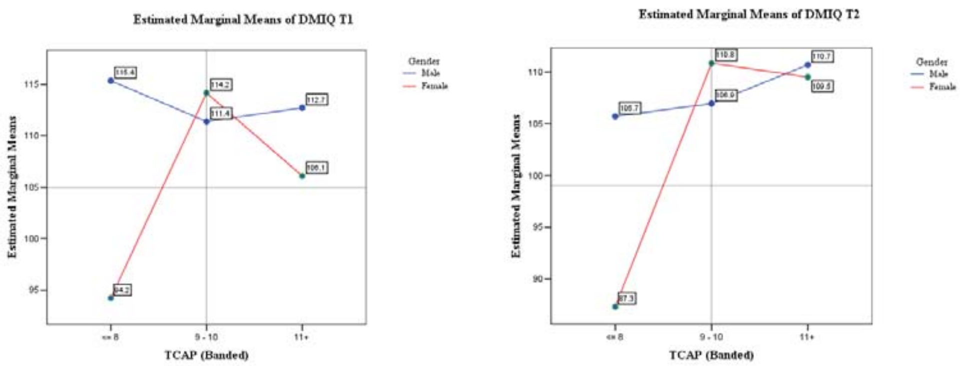
* p < .05 ** p < .01 *** p < .001 (2-tailed).
Note: DMIQ1 = Domain-Masculine Intelligence Type at pre-task estimation condition; DMIQ2 = Domain-Masculine Intelligence Type at post-task estimation condition. TCAP = Total Correct Aptitude Problems.

As the main interaction effect was significant, a further investigation of the relationship was warranted. Simple effects analysis was conducted. Data was split per gender and two one-way between-groups analysis of variance were conducted.

For males, the one-way between-groups analysis of variance for DMIQ1 was not significant, F(2,24) = .21, p =.82, η^2 =.02. The robust tests of equality of means, Welch (2, 13) = .14, p =.87; Brown-Forsythe (2,15) = .18, p =.84 were not significant.

Post-hoc comparisons using the Tukey HSD and Bonferroni tests revealed no significant differences in mean scores between the three groups for males.

Figure 7.6.2: Two 2-way ANOVAs (TCAP and Gender) on DMIQ1 and DMIQ2



For females, there was a statistically significant difference in the three TCAP groups on DMIQ1, $F(2,24) = 4.63$, p=.02, $\eta^2 = .28$, with large effect size. The robust tests of equality of means, Welch (2, 8) = 4.28, $p > .05$; Brown-Forsythe (2,12) = 4.12, $p < .05$ revealed mixed results about the right to reject the null hypothesis. The post-hoc comparisons using the Tukey HSD and Bonferroni tests revealed significant differences between Group 1 (<=8) (M = 94.22, SD = 13.65) and Group 2 (9-10) (M = 114.17, SD = 14.55). No other significant mean score differences were observed between the groups. Hypothesis 10 was confirmed.

For DMIQ2, the interaction effect between gender and TCAP was not significant, $F(2,48) = 2.83$, $p = .07$, $\eta_p^2 = .11$. The main effect for gender, $F(1,48) = 1.53$, p =.22, $\eta_p^2 = .03$ was not significant. The main effect for TCAP, $F(2,48) = 4.69$, $p < .05$, $\eta_p^2 = .16$ was significant, with medium effect size. Planned contrasts revealed significant differences between Group 1 and Group 3 (Contrast Estimate -12.96, $p < .01$). Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean score for Group 1 (<=8) was significantly different from Group 2 (9-10) as well as from Group 3 (11+). This was confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Hypothesis 11 was not confirmed.

Thus, hypotheses 1, 2, 4, 6, and 10 were confirmed and hypotheses 3, and 9 were partially confirmed, whilst hypotheses 5, 7, and 11 were not confirmed.

7.6.4. Discussion

Study 15 aimed to confirm the findings of Studies 11 to 14. It also set out to examine whether a content change in the psychometric measure, i.e. inclusion of Gc items, would impact pm the results as in Study 13.

As in all previous studies, HHE was observed on both DMIQ1 (η^2 =.16, *d* =.85) and DMIQ2 (η^2 =.12, *d* =.78) and a significant decrease in DMIQ from T1 to T2 was observed (*d* =.38). In line with previous results, males provided more correct answers on the numerical and spatial psychometric problems than females. The results also provided further support for the claim of no sex differences in general intelligence, with no differences observed on the crystallised items (e.g. Colom & Garcia-Lopez, 2002; Halpern et al., 2007; Lynn, 1999; Spelke, 2005). Yet, contrary to earlier results and literature reports (e.g. Carr et al., 2008), no gender differences were observed on the task-success probability estimation measure and probes, indicating the males and females in this sample did not differ in task-confidence beliefs.

Equally, in support of findings of Study13 and 14, gender was confirmed as the best predictor of DMIQ1, accounting for 12% of explained variance. However, gender failed to be the best predictor of DMIQ2, with TSP accounting for 14% of explained variance in the type. This result was identical to results of Study 11 and 12 where task-success probability was the best predictor of DMIQ2.

The role that gender plays in the relationship between TSP and DMIQ1 and DMIQ2 was investigated next. For DMIQ1 the results revealed only a significant effect for gender, with significantly higher DMIQ1 estimates provided by males on all three TSP groups. As in earlier studies, this finding provides further support for the existence of male hubris in the estimation process. For DMIQ2, the results revealed an identical estimation pattern, with male hubris across the three TSP groups.

Finally, gender's role in the relationship between TCAP and DMIQ1 and DMIQ2 was examined. For DMIQ1, the results revealed significant interaction effect between TCAP and gender and a significant gender effect. Lowest DMIQ1 estimates were provided by the group that solved fewest psychometric problems, average DMIQ1 estimates by the group that solved most problems and highest DMIQ1 estimates by the group that solved average number of problems.

For gender, males provided higher DMIQ1 estimates in the lowest and highest groups. However, for the first time, females provided higher DMIQ1 estimates than males. This occurred in the group that solved average number of problems, although the observed effect size was medium. In addition, males' highest DMIQ1 estimates were provided by the group that solved fewest psychometric problems, average estimates by group solving average number of problems and lowest DMIQ1 estimates by the group that solved the highest number of psychometric problems. However, the three groups did not significantly differ from each other. The male results resemble the male TSP estimation pattern observed in Study 14, and provide additional support for the claims that males are over-confident but inaccurate estimators of their math abilities (Meece et al., 2006; Meelissen & Luyten, 2008).

The female estimation pattern resembled the overall pattern, with lowest DMIQ1 estimates provided by the least capable group, highest estimates by the average group and average DMIQ1 estimates by the group solving most problems. For females, there was a significant difference in estimates between the lowest and average TCAP groups.

For DMIQ2, the only significant effect observed was for TCAP, with lowest DMIQ2 estimates provided by the group that solved least problems, highest DMIQ2 estimates by the average solving group and average DMIQ2 estimates by the group solving most psychometric problems.

Thus, the results of the analyses of variance in this study differ from the earlier studies in that females provided higher DMIQ estimates than males in one estimation condition and that no gender effect was observed between TCAP and DMIQ2.

7.7. Summary

To date, no experimental studies have been conducted in the SEI research programme and only a few SEI studies used 'objective' or psychometric measures to compare the accuracy and validity of SEI estimates (e.g. Batey et al., 2009; Chamorro-Premuzic, Moutafi, & Furnham, 2005; Furnham & Fong, 2000; Furnham & Mottabu, 2004; Furnham, Moutafi, & Chamorro-Premuzic, 2005; Furnham & Rawles, 1999; Holling & Preckel, 2005; Reilly & Mulhern, 1995). Likewise, the majority of SEI studies were conducted with university students. Participants in the experimental studies reported here were predominantly from the general public, making the results more generalisable and robust.

Chapter 7 contains five experimental studies that explored the impact of the repeated measurement of DMIQ and of the psychometric task and task-success probes on the occurrence of HHE on DMIQ. As in the first part of this thesis, gender was expected as the best predictor of DMIQ. The experimental design allowed for in-depth examination of the role gender plays in the repeated measurement of DMIQ as well as in the relationships between DMIQ and TCAP and DMIQ and TSP. Equally, gender differences in TCAP and TSP were examined in an attempt to understand the conflicting claims in current literature and to clarify whether they have any bearing on the gender differences in the intelligence type.

Although all five studies were identical in overall design and execution, the content and format of the psychometric task and the number of task-success probes differed per study. This was done to test whether alternating numerical, reasoning, spatial and crystallised knowledge problems and varying the number of TSP probes impacts on the DMIQ estimation process, the hubris-humility effect and the role of gender herein. In addition, TCAP content alternation was expected to be gender-stereotype inducing as it contained items that are perceived as domain masculine, especially by females.

The repeated measurement of DMIQ aimed to ascertain that HHE can be manipulated or reduced following the psychometric and task-success task, based on the assertions that repeated measures affect mood, confidence and behaviour (Bartsch & Nesselroade, 1973; Ryckman et al., 1971). The results of all five studies confirmed the existence of HHE in the pre- and post-task DMIQ estimates as well as significant reduction in the intelligence type estimates from pre- to post-task estimation condition. The effect sizes for HHE's occurrence on DMIQ1 and DMIQ2 ranged from medium to very large and the effect sizes for the DMIQ estimate reduction ranged from small to medium. These results validated the findings of the first ten studies as well as providing further support for the role gender plays in HHE and DMIQ.

The gender-stereotype literature has provided abundant evidence for female underperformance on domain-masculine tasks (e.g. Dar-Nimrod, 2007; Ehrlinger &

Dunning, 2003; Hyde et al., 1990a,b). The results here established that the psychometric and task-confidence task caused both genders to lower their post-task estimates, although female estimates were lower than males'. These findings are surprising as the existing literature shows that men have higher self-confidence and report higher self-perceived ability on domain-masculine tasks, e.g. mathematics (Meece et al., 2996; Meelissen & Luyten, 2008). Thus, the task seems to have affected both genders similarly, impacting on male and female self-perceptions and ability beliefs and causing both genders to reduce their post-task estimates. In other words, the task brought about skill and ability realisation that in turn affected self-perceptions.

Gender was expected to be the best predictor of DMIQ1 and DMIQ2. Table 7.7.1. provides a summary of significant predictors of the intelligence type in Studies 11 to 15. Gender was confirmed as the best predictor of DMIQ1 in Study 13, 14, and 15 and these results are in line with the findings of the first ten studies. However, gender was the best determinant of DMIQ2 in only two studies (13 and 14). Unexpectedly, TSP was twice the best predictor of DMIQ1 (Study 11and 12) and three times of DMIQ2 (Study 11, 12 and 15). The role of TSP as the best predictor of DMIQ2 was unforeseen, and revealed that the task-confidence probes or participants' perceived task-success, had the biggest impact on the post-task estimates. These results provide additional support for the impact of the psychometric and task confidence task, and in particular TSP probes, on the DMIQ estimation pattern by both genders.

Table 7.7.1: Summary of Significant Predictors of DMIQ1 and DMIQ2 – Studies 11 to 15

Study (n)	Significant Predictors of the Regression Analyses DMIQ1	DMIQ2
11 (488)	**TSP (.46)**	**TSP (.54)**
	G (-.23)	G (-.18)
12 (182)	**TSP (.30)**	**TSP (.38)**
	G (-.26)	G (-.32)
	TCAP (.23)	TCAP (.23)
13 (80)	**G (-.50)**	**G (-.53)**
	TCAP (.26)	TCAP (.36)
14 (157)	**G (-.44)**	**G (-.48)**
	TCAP (.16)	TCAP (.11)
15 (54)	**G (-.37)**	**TSP (.40)**
	TSP (.28)	G (-.27)

Legend: G = Gender; TSP = Task-success probability, TCAP = Total Correct Aptitude Problems. In brackets are β values. All values are significant and best predictors are in bold.

Gender differences in math achievement, attitudes and affect have been extensively researched and documented (cf. Halpern et al., 2007), with females displaying more negative or self-handicapping math attitudes, having lower math self-confidence, stereotyping math as domain-masculine, underperforming on standardised math tests, and opting out of STEM careers (Crombie et al., 2005; Beyer, 1990, 1998; Hyde et al., 1990a,b; Linn & Hyde, 1989; Meelissen & Luyten, 2008; Sax & Harper, 2007; The College Board, 1998). On the other hand, males perceive math as a domain-masculine and are more self-confident about their math abilities (Meece et al., 2006; Meelissen & Luyten, 2008; van der Sluis et al., 2010).

Thus, males were expected to do better on the psychometric task and be more confident about their success. The results confirmed these claims, with males correctly solving significantly more psychometric problems in Studies 11, 12, and 15 and providing higher task-success probability estimates in Studies 11, 12, and 13. An overview of the results is in Table 7.7.2. Study 14 did not examine gender differences in either TCAP or TSP, although the three individual studies did (see appendix for details of the three individual studies).

Table 7.7.2: Overview of Significant Gender Differences in TSP and TCAP- Studies 11 to 15

Overview of Significant Gender Differences in TSP and TCAP Overview of 5 Experiments				
Study (n)	TSP		TCAP	
	Total TSP Measure (sig. GD differences)	Number of Probes with sig. GD (Out of Total)	Number of Prbs. With sig. GD (Out of Total)	TCAP Content
11 (488)	Yes	4 (5)	Yes, 12 (15)	**N+R**
12 (182)	Yes	5 (5)	Yes, 4 (13)	**N+R**
13 (80)	Yes	1 (3)	No, 0 (16)	N+R+Gc
14 (157)	N/A	N/A	N/A	N/A
15 (54)	No	0 (3)	Yes, 2 (15)	**N+R**+Gc

Legend: N = Numerical Problems, R = Reasoning Problems, Gc = Crystallised Intelligence Problems. GD = Gender Differences. Significant gender differences in TCAP (content) are in bold.

No gender differences in TCAP were observed in Study 13. This result was attributed to the fact that crystallised intelligence items that are shown to yield same results in males and females (e.g. Colom & Garcia-Lopez, 2002), were included in the task. Results of Study 15 further validated this assumption as no gender differences were contained in TCAP, and in particular in crystallised intelligence. In fact, Study 15 demonstrated that gender differences or the male advantage in the psychometric task were contained to the numerical problems. These results provided additional support for the claim that tasks perceived as most domain-masculine trigger gender stereotypical responses and active negative self-perceptions of ability, especially in females. In regards to gender differences in task confidence, Studies 11, 12 and 13, confirmed male advantage on task-success probability, providing additional support for higher task confidence in males.

To better understand the role gender plays in TSP and TCAP in both estimation conditions, a series of analyses of variance were conducted. Table 7.7.3. summarises the results of Studies 11 to 15. For TSP, no interaction affects between TSP and gender were observed in both estimation conditions. However, a significant gender effect was observed in all ten analyses and a significant TSP effect in six analyses. No significant TSP effect was observed in Studies 13 and 15, that had changed the psychometric task content and included Gc items. In all ten analyses males provided higher DMIQ estimates across all three TSP groups.

Table 7.7.3: Overview of Variables with Significant Effects and Significant Interactions in 2-way ANOVAs (TSP and Gender and TCAP and Gender) on DMIQ1 and DMIQ2- Studies 11 to 15

Study	TSP		TCAP	
(n)	DMIQ1	DMIQ2	DMIQ1	DMIQ2
11 (488)	Gender and TSP	Gender and TSP	Gender, TCAP, TCAP x Gender (I)	Gender and TCAP
12 (182)	Gender and TSP	Gender and TSP	Gender and TCAP	Gender and TCAP
13 (80)	Gender	Gender	Gender and TCAP	Gender and TCAP
14 (157)	Gender and TSP	Gender and TSP	Gender, TCAP, TCAP x Gender (I)	Gender and TCAP
15 (54)	Gender	Gender	Gender, TCAP, TCAP x Gender (I)	TCAP

Legend: TSP = Task-success probability; TCAP = Total Correct Aptitude Problems. I = significant interaction effect.

The accuracy of DMIQ1 and DMIQ2 estimates by the three TSP groups was notable. Overall, males and females provided accurate or matching DMIQ1 and DMIQ2 estimates, i.e. low DMIQ estimates by low task-success probability group, average estimates by average group and high DMIQ estimates by high task-success probability group. The only exception was Study 14, where participants provided miscalibrated DMIQ1 and DMIQ2 estimates. The results are also presented in an overview below.

However, the results were very different for TCAP. Significant interaction effects between TCAP and gender were observed three times for DMIQ1 but not for DMIQ2. Significant gender effects were observed in nine out of ten analyses, with males across all three TCAP groups providing higher DMIQ estimates than females. The only exception was Study 15, where females in the average psychometric group provided higher DMIQ2 estimates than did males. Yet, significant TCAP effects were observed in all ten analyses.

The accuracy of DMIQ1 and DMIQ2 estimates by the three TCAP groups differed remarkably from TSP results. See Table 7.7.4. for an overview. Overall, the estimates were less accurate, apart from Studies 12 and 13 for both DMIQ1 and DMIQ2 and Study 11 for DMIQ1. In particular, the interaction effect estimates provided by males and females in the three TCAP groups were inaccurate in three out of four instances, with the exception of female DMIQ1 estimates in Study 14.

Table 7.7.4: Overview of the DMIQ Estimation Patterns in 2-way ANOVAs (TSP and Gender and TCAP and Gender) – Studies 11 to 15

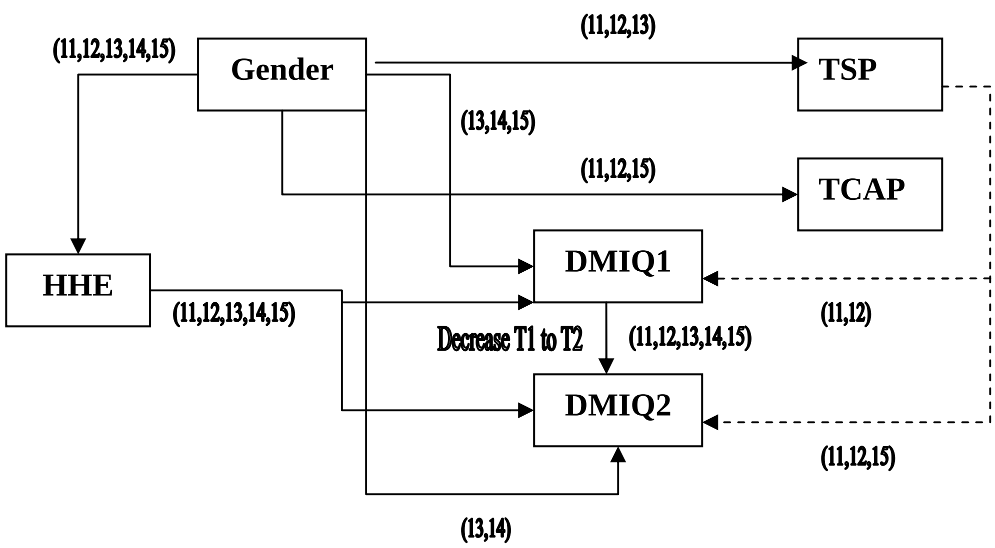
Study	TSP	TSP		TCAP	TCAP					
	Group	DMIQ1	DMIQ2	Group	DMIQ1			DMIQ2		
					TCAP	TCAP x G		TCAP	TCAP x G	
						M	F		M	F
11	L (1)	**L**	**L**	L (1)	**L**	**A**	**A**	**L**		
	A (2)	**A**	**A**	A (2)	**A**	**L**	**L**	**H**		
	H (3)	**H**	**H**	H (3)	**H**	**H**	**H**	**A**		
12	L (1)	**L**	**L**	L (1)	**L**			**L**		
	A (2)	**A**	**A**	A (2)	**A**			**A**		
	H (3)	**H**	**H**	H (3)	**H**			**H**		
13	L (1)	**L**	**L**	L (1)	**L**			**L**		
	A (2)	**A**	**A**	A (2)	**A**			**A**		
	H (3)	**H**	**H**	H (3)	**H**			**H**		
14	L (1)	**A**	**H**	L (1)		**H**	**L**		**A**	**L**
	A (2)	**L**	**L**	A (2)		**L**	**A**		**L**	**H**
	H (3)	**H**	**A**	H (3)		**A**	**H**		**H**	**A**
15	L (1)	**L**	**L**	L (1)	**L**	**H**	**L**	**L**		
	A (2)	**A**	**A**	A (2)	**H**	**L**	**H**	**H**		
	H (3)	**H**	**H**	H (3)	**A**	**A**	**A**	**A**		

Legend: G = Gender; TSP = Task-success probability; TCAP = Total Correct Aptitude Problems. In bold are DMIQ estimates, i.e. L = low, A = average, H = high. Shaded cells indicate inaccurate or mismatched DMIQ estimations by participants in the three TSP/TCAP groups. For analyses with significant interaction affects (TCAP x Gender) scores for male and female participants are given. M = males, F = females.

As the TCAP and TSP tasks were devised to also validate the claims that individuals overestimate their ability on easy tasks and underestimate their abilities on difficult tasks (e.g. Alicke et al., 1995; Burson et al., 2006; Guenther & Alicke, 2010; Moore & Small, 2007), leading them to make inaccurate performance judgements (Ehrlinger & Dunning, 2003; Kahneman & Tversky, 2000), the observed results are particularly interesting. Based on the observed data, individuals were capable of more accurate intelligence estimates in the task-success probability conditions then in the psychometric conditions. In particular, the TSP results support the assertions that individuals are capable of accurate self-assessments of ability (e.g. Ackerman et al., 2002; Chamorro-Premuzic et al., 2010; Hall & Carter, 1999; Swim, 1994) but not in the psychometric task condition. Equally, the observed male hubris in DMIQ estimates that was observed on all but one occasion, provided support for the literature in the field. Thus, gender influenced the relationship between TSP and DMIQ as well as between TCAP and DMIQ.

The results of Studies 11 to 15 are represented in Figure 7.7.1. The single-pointed arrows symbolize a direct relationship between two variables. The dashed arrows (i.e. between TSP and DMIQ1 and DMIQ2) represent relationships that were not predicted but observed. In brackets are studies with same results. Variables that exhibited a relationship with DMIQ1 and DMIQ2 are in bold. The direction of the arrows implies causality that is based on results of Studies 11 to 15.

Figure 7.7.1: Pictorial representation of the results of the experimental studies 11 to 15



Chapter 8: Conclusions

8.1. Summary of findings and implications

The aim of this thesis was to investigate the potential determinants of gender differences in the self-estimated intelligence model (SEI) in order to contribute to the SEI research programme. Specifically, the largest gender differences were expected to be observed on the numerical-spatial factor of SEI (e.g. Beloff, 1992; Bennet, 1997; Bond, 1991; Furnham, 2001; Furnham & Baguma, 1999; Furnham & Fukumoto, 2008;Holling & Preckel, 2005; Pallier, 2003; Rammstedt & Rammsayer, 2002a,b; Swami et al., 2006; Szymanowicz et al., 2011, unpublished manuscript), or the *'domain-masculine intelligence type'* (DMIQ), which is a novel variable introduced in this thesis. Equally, *hubris-humility effect'* (HHE)*,* i.e. male overestimation and female underestimation of cognitive abilities (Furnham, 2001; von Stumm et al., 2009), was anticipated to occur on DMIQ.

The existence of gender differences in DMIQ as well as the occurrence of HHE on DMIQ was validated in all fifteen studies (see Table 8.1.1. for more details), with males providing higher DMIQ1 and DMIQ2 estimates than females in all individual studies. The observed effect sizes ranged from medium to very large, with the largest effect size observed in Study 4 (η^2 =.32) and the largest effect sizes for an experimental study observed in Study 14B, which was the second individual condition of Study 14, with η^2 =.38 for DMIQ1 and η^2 =.50 for DMIQ2 respectively. The smallest effect size (η^2 =.08) was observed in the Czech Republic sample (Study 8), which is unsurprising given earlier results (Furnham, Rakow et al., 1999) with a comparable culture (Slovakia) that found no gender differences.

Studies 1 and 2 affirmed that HHE was most pronounced on the DMIQ type among the ten self-estimated intelligences, providing further support for the notion of male-normativeness of intelligence (Furnham, 2000). Indeed, HHE was stronger on the DMIQ type than it was on the mathematical/logical and spatial intelligences individually (see section 2.4. for more detailed discussion). Thus, the findings of the present research provide strong evidence for the confinement of gender differences in self-estimates abilities to the mathematical/logical/spatial factor of the SEI model, or the DMIQ type.

Table 8.1.1: Summary Statistics and Effect Sizes for Gender Differences in DMIQ or HHE on DMIQ for All Studies – Total Sample and Per Gender

DMIQ	Total M (SD) n	Males M (SD) n	Females M (SD) n	t(df)	Mean Diff.	95% CI	η^2
Part I							
Study 1	110.97 (13.34) 129	117.72 (13.72) 52	106.41 (11.01) 77	-5.18(127)***	-11.31	-15.64 - -6.99	**.17**
Study 2	106.81 (9.55) 115	111.04 (9.22) 38	104.73 (9.06) 77	-3.49(113)**	-6.31	-.9.89 - -2.73	.10
Study 3	103.20 (14.08) 71	115.96 (17.10) 12	100.60 (11.97) 59	3.75(69)***	15.36	7.19 – 23.52	**.17**
Study 4	108.41 (15.01) 121	120.64 (14.34) 39	102.59 (11.45) 82	7.46(119)***	18.05	13.26 – 22.84	**.32**
Study 5	109.72 (10.66) 102	120.17 (8.01) 23	106.67 (9.34) 79	-6.29(100)***	-13.50	-17.77 - -9.24	**.28**
Study 6	111.35 (11.24) 143	116.82 (10.68) 64	106.92 (9.66) 79	-5.81(141)***	-9.90	-13.27 - -6.53	**.19**
Study 7	108.05 (12.82) 139	114.20 (11.73) 61	103.24 (11.57) 78	-5.51(137)***	-10.96	-14.90 - -7.03	**.18**
Study 8	102.59 (11.20) 116	107.66 (10.61) 31	100.75 (10.89) 85	-3.05(114)**	-6.91	-11.41 - -2.42	.08
Study 9 Colombia	106.13 (11.28) 50	110.36 (10.93) 28	100.75 (9.43) 22	-3.27(48)**	-9.61	-15.51 - -3.71	**.18**
UK	109.93 (11.19) 52	114.37 (9.21) 26	105.50 (11.38) 26	-3.09(50)**	-8.87	-14.63 - -3.10	**.16**
Study 10	139.31 (14.35)	143.92 (12.53)	134.43 (14.58)	5.56(243)***	9.49	6.13 – 12.85	.11

	255	131	124				
Part II							
Study 11	112.86	120.64	108.55				
DMIQ1	(19.37)	(18.13)	(18.70)	6.16(385)***	12.09	8.23 – 15.95	.09
	387	138	249				
DMIQ2	107.97	116.02	102.57				
	(22.27)	(21.58)	(21.14)	4.68(227)***	13.46	7.79 – 19.12	.09
	229	92	137				
Study 12	112.68	120.94	104.59				
DMIQ1	(19.93)	(17.96)	(18.46)	6.06(180)***	16.35	11.02 – 21.68	**.17**
	182	90	92				
DMIQ2	106.59	117.46	95.96				
	(21.48)	(18.10)	(19.13)	7.78(180)***	21.50	16.05 – 26.95	**.25**
	182	90	92				
Study 13	106.38	114.29	98.50				
DMIQ1	(15.27)	(15.45)	(10.26)	5.39(68)***	15.81	9.96 – 21.67	**.30**
	80	40	40				
DMIQ2	102.79	113.06	94.10				
	(17.70)	(17.22)	(12.92)	5.33(70)***	18.96	11.86 – 26.05	**.30**
	72	33	39				
Study 14	111.66	119.32	104.48				
DMIQ1	(15.29)	(13.00)	(13.77)	6.93(155)***	14.83	10.61 – 19.06	**.24**
	157	76	81				
DMIQ2	104.85	115.09	95.24				
	(18.50)	(15.64)	(15.65)	7.94(155)***	19.85	14.91 – 24.78	**.29**
	157	76	81				
Study 14A	113.35	119.78	108.77				
DMIQ1	(16.74)	(14.88)	(16.71)	2.35(46)*	11.01	1.59 – 20.43	.11
	48	20	28				
DMIQ2	105.85	114.68	99.55				
	(17.85)	(11.25)	(19.15)	3.16(46)**	15.12	5.48 – 27.76	**.18**
	48	20	28				
Study 14B	108.27	114.86	101.00				
DMIQ1	(11.31)	(8.76)	(9.22)	6.02(59)***	13.86	9.25 – 18.47	**.38**
	61	32	29				
DMIQ2	101.01	111.33	89.62				
	(15.55)	(10.87)	(11.48)	7.59(59)***	21.71	15.98 – 27.43	**.50**
	61	32	29				
Study 14C	114.28	124.88	103.69				
DMIQ1	(17.53)	(14.27)	(13.78)	5.23(46)***	21.19	13.04 – 29.34	**.37**
	48	24	24				
DMIQ2	108.72	120.44	97.00				
	(21.75)	(22.01)	(13.98)	4.40(46)***	23.44	12.72 – 34.15	**.30**
	48	24	24				
Study 15	106.91	112.96	100.85				
DMIQ1	(15.40)	(11.99)	(16.23)	3.12(52)**	12.11	4.32 – 19.90	**.16**
	54	27	27				
DMIQ2	102.41	108.15	96.67				
	(16.96)	(13.24)	(18.52)	2.62(52)*	11.48	2.69 – 20.27	.12
	54	27	27				
Summary							
All Studies	113.17	120.96	107.83				
DMIQ/1	(17.73)	(17.19)	(16.04)	17.83(1779)***	13.14	11.69 – 14.58	.13
	2137	869	1268				
DMIQ2	106.49	115.61	98.77				
	(18.92)	(16.70)	(17.18)	13.03(692)***	16.84	14.31 – 19.38	**.20**
	694	318	376				

* p < .05. ** p < .01. *** p < .001.

Note: Large effect sizes are in bold. DMIQ/1 = DMIQ estimates and DMIQ1 estimates taken at the pre-task estimation condition, DMIQ2 = estimates taken at the post-task estimation condition. *d* = Hedge's

Adjustment/Cohen's d adjusted for sample size or normal Cohen's *d* is used in all studies. Studies 1 to 7 and 10 to 15 = undergraduates of British universities and/or general public population; Study 8 = Czech Republic participants, Study 9 = Colombian and British participants.

As such these findings also provide support for the effect and role of (gender) stereotypical beliefs, biased self-evaluations, self-enhancement or self-derogatory biases, inflated and deflated performance beliefs, overconfidence and confidence bias, and gender differences in math achievement, attitudes and ability (Ackerman & Wolman, 2007; Bleeker & Jacobs, 2004; Chamorro-Premuzic & Arteche, 2008; Carr et al., 2008; Else-Quest, Hyde & Linn, 2010; Guimond et al., 2006; Kwan et al., 2008; Lytton & Romney, 1991) that were suspected to play a role in the self-estimates of intelligence (see sections 1.3.2., 1.3.3., 1.3.4., 1.3.5. and 1.3.5.1. for a more detailed explanation).

The second objective was to ascertain whether gender was the best determinant of DMIQ, over and above a number of potential determinants of gender differences in SEI, such as general intelligence ('*g*'), beliefs about intelligence, gender identity variables, self-constructs, and affect measures. Table 8.1.2. summarises the findings of the fifteen individual studies.

The selection of these determinants was based on previous evidence of a relationship or a role within the SEI model (e.g. Beyer, 1998, 1999; Chamorro-Premuzic & Arteche, 2008; Chamorro-Premuzic, Furnham, Moutafi, 2004; Duckworth & Seligman, 2005, 2006; Ehrlinger & Dunning, 2003; Guimond et al., 2006 ; Halpern et al., 2007; Hirchy & Morris, 2002; Kwan et al., 2008; Lippa, 2001; Petiprin & Johsnon, 1991) or based on literature assertions in the field (e.g. Ackerman & Wolman, 2007; Ambady et al., 2001; Carr et al., 2008; Dar-Nimrod, 2007; Dunning et al., 1990; Feingold, 1988, 1996; Gottfredson, 2000; Nosek, Banaji, &

Greenwald, 2002; Pallier, 2003; Sleeper & Nigro, 1982; Steele & Aronson, 1995; Watson & Tellegen, 1985). The role of age in DMIQ was also examined based on the evidence that it influences the provided self-estimates of intelligence (cf. Beier & Ackerman, 2001, 2003; Rammstedt & Rammsayer, 2002b).

In eight out of fifteen studies (53%) gender was confirmed as the best predictor of the intelligence type (see Table 8.1.2. and Figure 8.1.1. for more details). Table 8.1.2. summarises the findings of the regression analyses. Gender was also the best predictor of the post-task DMIQ estimation condition (DMIQ2) in Studies 13 and 14 (40%). Contrary to prediction, gender was not the only significant predictor of DMIQ.

Among the potential determinants, three determinants significantly contributed to the prediction of the intelligence type, '*g*', masculinity and task-success probability estimation probes. The Wonderlic Personnel Test (Wonderlic, 1992), which is a measure of general intelligence, was the best predictor of DMIQ in Studies 3 and 4 and masculinity was the best predictor of DMIQ in Studies 7, 8 and 9. Task-success or task confidence probes were twice (40%) the best predictor of DMIQ1 and three times (60%) of DMIQ2 in the five experimental studies. These findings contribute to the SEI research programme as well as to research in fields of gender identity and confidence biases and performance expectations.

Thus, when subjects were only offered general intelligence measures and asked to estimate their DMIQ, the intelligence measures accounted for most variance in the intelligence type, over and above gender. However, when other variables, such as gender identity and self-constructs were added, gender was the best predictor of the intelligence type, over and above intelligence measures, gender identity variables and self-construct measures. In studies where masculinity accounted for most variance in

DMIQ, gender identity variables and affect measures (Studies 8 and 9) and self-constructs (Study 7) were also included. Thus, when individuals are offered measures that are most likely to activate gender-stereotypical beliefs and attitudes, such as psychometric tests, and no other measures are offered that could *'dilute'* or *'divert'* these beliefs and attitudes, such as gender identity, affect and self-construct measures, individuals are *'primed'* and this is revealed through the cause, i.e. psychometric test(s), becoming the best predictor of the self-estimates of ability. In other words, when individuals are exposed to only gender-stereotype inducing activities and asked to provide self-estimates of ability in areas shown as the most gender stereotype sensitive, i.e. DMIQ, the activity itself becomes the most important predictor of the observed self-estimates of ability.

These results resemble the findings of research on stereotypical priming and stereotypical biases and threat (e.g. Ambady et al., 2001; Chatard et al., 2007; Dar-Nimrod, 2007; Dar-Nimrod & Heine, 2006; Kiefer & Sekaquaptewa, 2007; Rudman & Phelan, 2010; Steele, 1997; Wheeler & Petty, 2001) that have demonstrated that priming individuals with gender-stereotypical beliefs impacts on their subsequent behaviour and self-beliefs, evokes stereotype-consistent behaviours and increases stereotype susceptibility. This is in particular true for women working on mathematical and scientific tasks (Dar-Nimrod & Heine, 2006: Steele, 1997).

Interestingly, masculinity was the best predictor of DMIQ in three studies. Two of these studies were conducted with British samples and one in the Czech Republic. Since laymen foster a broader definition of masculinity that includes social roles, physical appearance, occupational choices and personality traits and is based on cultural norms and beliefs (Lippa, 2001), these results are not unanticipated. In fact the United Kingdom has a high masculinity score, compared to an average score in the Czech Republic (Hofstede, 2003). Moreover, the majority of the SEI research studies that reported significant gender differences have been conducted with British samples and none assessed the role of masculinity in the self-estimation process. Furthermore, domain-masculine activities, such as mathematics, have been shown to evoke gender stereotypical beliefs, effect learning and cause academic underperformance in females (Rydell et al., 2010; Steele & Aronson, 1995; Steele, 1997). Equally, as gender stereotypes are strongest in areas associated with masculinity and femininity (Brown & Josephs, 1999) and the domain-masculine intelligence type was shown as the most sensitive predictor of gender differences in SEI, these results seem logical and plausible.

As identical measures were used throughout the five experimental studies, it is tempting to assume that the impact of the composition of the psychometric task (TCAP) or the number of the TSP probes has contributed to the determination of the best predictor of the intelligence type in the individual studies. In Studies 11, 12 and 15 (only for DMIQ2), task-success probes accounted for most variance in DMIQ1 and DMIQ2, with 5, 5, and 3 TSP probes, respectively. Psychometric problems for Study 11 and 12 were almost identical, including numerical and reasoning problems. For Study 15, the task included numerical, spatial and crystallised intelligence problems.

For Studies 13 and 14, where gender was the best predictor of the intelligence type, 3 and 4/7/0 TPS probes were used. The content of the psychometric problems differed, for Study 13 numerical, spatial and crystallised intelligence problems were asked, whereas in Study 14 only numerical and spatial problems were given. These results are multifaceted and complex and suggest that participant composition in each study is likely to have contributed to the observed results. Equally, other variables that were not addressed, such as parental beliefs and attitudes towards math, cultural and religious norms, self-belief and stereotypical biases, might have contributed to the determination of the best predictor of the intelligence type in each study. This is particularly notable in the experimental studies where task confidence was the best overall predictor of DMIQ, following the gender-stereotype inducing psychometric task. Yet, in accordance with prediction, gender was overall the best predictor of the intelligence type providing additional support for the existence of gender differences in self-estimates of ability.

Table 8.1.2: Overview of Significant Predictors of DMIQ

Study (n)	Significant Predictors of the Regression Analyses DMIQ/DMIQ1	DMIQ2
1 (130)	**Gender (.43)**	N/A
2 (115)	**Gender (.33)**	N/A
3 (85)	**Gf (.48)** Gender (-.33)	N/A
4 (121)	**Gf (.49)** Gender (-.35)	N/A
5 (102)	**Gender (.48)**	N/A
6 (143)	**Gender (.46)**	N/A
7 (139)	**Masculinity (.36)** Gender (.32)	N/A
8 (116)	**Masculinity (.22)**	N/A
9 (102) UK	**Masculinity (.39)** Gender (.36)	N/A
10 (278)	**Gender (-.29)**	N/A
11 (488)	**TSP (.46)** Gender (-.23)	**TSP (.54)** Gender(-.18)
12 (182)	**TSP (.30)** Gender (-.26) TCAP (.23)	**TSP (.38)** Gender (-.32) TCAP (.23)
13 (80)	**Gender (-.50)** TCAP (.26)	**Gender (-.53)** TCAP (.36)
14 (157)	**Gender (-.44)** TCAP (.16)	**Gender (-.48)** TCAP (.11)
15 (54)	**Gender (-.37)** TSP (.28)	**TSP (.40)** Gender (-.27)

Legend: TSP = Task-success probability. TCAP = Total Correct Aptitude Problems. In brackets are β values of regression analyses. All values are significant and best predictors are in bold.

This research also contributed to the SEI research programme by conducting cross-cultural studies to confirm the existence of HHE or gender differences in DMIQ in cultures that were not previously tested: the Czech Republic and Colombia. Substantial evidence is available about the occurrence of HHE across cultures and geographies (e.g. Furnham & Chamorro-Premuzic, 2005; Furnham & Fong, 2000; Furnham, Hosoe, & Tang, 2001; Furnham & Mottabu, 2004; Furnham, von Stumm, et al., 2009). Therefore, the results that validated the existence of HHE on DMIQ in these cultures were not surprising.

However, in the Czech sample, masculinity was the best single predictor of DMIQ (see section 5.2.4. for more detailed explanation) and no significant predictors of DMIQ were found in the Colombian sample. These findings suggest that despite the validation of HHE on DMIQ, culture affected the determinant(s) of the intelligence type and provided additional support for the assertion that cultures differ in their understanding and meaning of intelligence as well as are prone to culturally specific stereotypical beliefs about intelligence (Furnham & Akande, 2004; Furnham, Rakow, & Mak, 2002; Furnham, Shahidi, & Baluch, 2002; Hofstede, 2003; Segall et al., 1999; Yang & Sternberg, 1997).

Although substantial literature exists on the differences and similarities between highly gifted and normal populations (e.g. Benbow et al., 2000; Lubinski & Humphreys, 1990; Roznowski, Reith, & Hong, 2000; Shea et al., 2001), no study in the SEI programme has to date investigated the occurrence of HHE and gender differences in a precocious sample. This research addressed this deficiency and provided evidence for the degree of embeddedness of gender stereotypical beliefs and gender differences in self-beliefs of ability and performance (Beyer, 1990, 1998; Ehrlinger & Dunning, 2003; Hoffman & Hurst, 1990; Kim et al., 2010; Steele & Aronson, 1995), such as the HHE, by affirming the existence of gender differences in DMIQ and HHE among highly gifted individuals.

Thus, even a population that is thoroughly knowledgeable about intelligence research findings as well as aware of their own intellectual superiority, displays beliefs that are biased, possibly damaging, and usually found among normal populations. These findings provide support for the assertions that precocious and normal populations are similar in their belief systems and life and career choices (Lubinski & Humphreys, 1990; Ferriman et al., 2009; Preckel et al., 2008; Roznowski et al., 2000).

Within this context, the findings of this research also contribute to the understanding of the role of (gender) stereotypical beliefs, cognitive biases, self-confidence, self-perceptions and self-beliefs of ability and performance in DMIQ. Likewise, the results of the initial experimental studies within the SEI research programme provided additional evidence about the role of gender in DMIQ, the accuracy of provided self-estimates, the role of task confidence in prediction of the DMIQ type, as well as the impact of repeated measurement and the effect of the gender-stereotype inducing psychometric task on gender differences in DMIQ.

The experimental studies were included to provide weight and objectivity to the SEI research programme by using a number of interventions that were not used previously, such as repeated measurement of DMIQ, specific psychometric task (TCAP) and task-success probes (TSP). Equally, these interventions were employed to test the role of (gender) stereotypical beliefs, cognitive biases, self-confidence, self-perceptions and self-beliefs of ability and performance in DMIQ. The results validated the main objectives of the experimental studies. Thus, HHE was observed in both pre- and post-task estimation conditions of the intelligence type in all five experiments.

Likewise, the intervention facilitated a significant size reduction in the hubris-humility effect, i.e. the DMIQ estimates provided after the intervention were smaller than the initial estimates. These results imply that the intervention brought about realisation or awareness about one's abilities and caused a downward correction. This 'correction' occurred in both males and females. However, male post-task estimates were still higher than female estimates. Equally, female estimates were lower than male estimates in pre- and post task estimation conditions. Thus, despite the downward correction in hubris, humility became even more modest.

It is likely that females perceived their performance more negatively, particularly since the task was gender stereotype inducing, and this in turn affected female confidence and led to self-handicapping behaviours (Kim et al., 2010; Roberts, 1991). Likewise, these results provide support for the assertion that self-beliefs are stronger determinants of future behaviour than objective feedback (Critcher & Dunning, 2009) as evidenced through the male and female estimates.

Equally, the results of this research provided evidence for the role of (over) confidence in the estimation process. Overconfidence has been shown to be responsible for the tendency to overestimate one's abilities and males have been shown to be significantly more overconfident than females (Burson et al., 2006; Carr et al., 2008; Pallier, 2003). The observed male hubris, in both pre- and post task estimation conditions affirms that overconfidence plays a role in male DMIQ estimates.

The role of confidence in the estimation process was further investigated in the experimental studies with task-success probes seen as proxy measures of confidence. In fact, the role of gender in the relationship between task- success probes and DMIQ was examined and revealed that gender played a role in all analyses. Predictably, males provided higher DMIQ estimates across all three task-success groups in all ten analyses (for detailed explanation of the results see section 7.7.). Yet, the most notable finding was the '*accuracy'* or the *'correspondence'* of the DMIQ estimates by participants in the three task confidence groups. These results support the earlier claims that individuals' self-estimates of ability are accurate and show understanding and awareness of one's ability (Ackerman et al., 2002; Ackerman & Wolman, 2007; Chamorro-Premuzic et al., 2010; Kornilova et al., 2009).

Males also provided higher DMIQ estimates in all but one psychometric task group. However, the DMIQ estimates provided by participants in the three psychometric task groups were less accurate (see section 7.7. for more detailed explanation). The results also revealed that individuals' DMIQ estimates were more accurate in the task-confidence condition than in the psychometric task condition. It seems likely that the psychometric task activated gender-stereotypical biases in participants and influenced the accuracy of the provided self-estimates by the participants in the three ability groups. These results are in line with literature in the field (e.g. Bonnot & Croizet, 2007; Dar-Nimrod, 2007; Chatard et al., 2007; Steele & Ambady, 2006; Wheeler & Petty, 2001).

Figure 8.1.1. shows the results of Studies 1 to 15 and represents a summary of the findings of this thesis. The single-pointed arrows symbolize a direct relationship between two variables. The dashed arrows (e.g. between '*g*' and DMIQ and masculinity and DMIQ) represent relationships that were not predicted but observed. Variables that exhibited significant relationships are in bold. The direction of the arrows implies causality, which is based on theoretical assumptions. These assumptions have been explained in detail in the individual studies, which examined these assumptions through numerous hypotheses. Studies that provided evidence in

support of the hypotheses or were observed despite being unpredicted are reported below.

In summary, the results of the fifteen studies confirmed the main aims of this thesis. The existence of gender differences in self-estimated intelligences and in particular in the domain-masculine intelligence type, was validated in all fifteen studies. Equally, the existence of the hubris-humility effect in the domain-masculine intelligence type was affirmed in all studies.

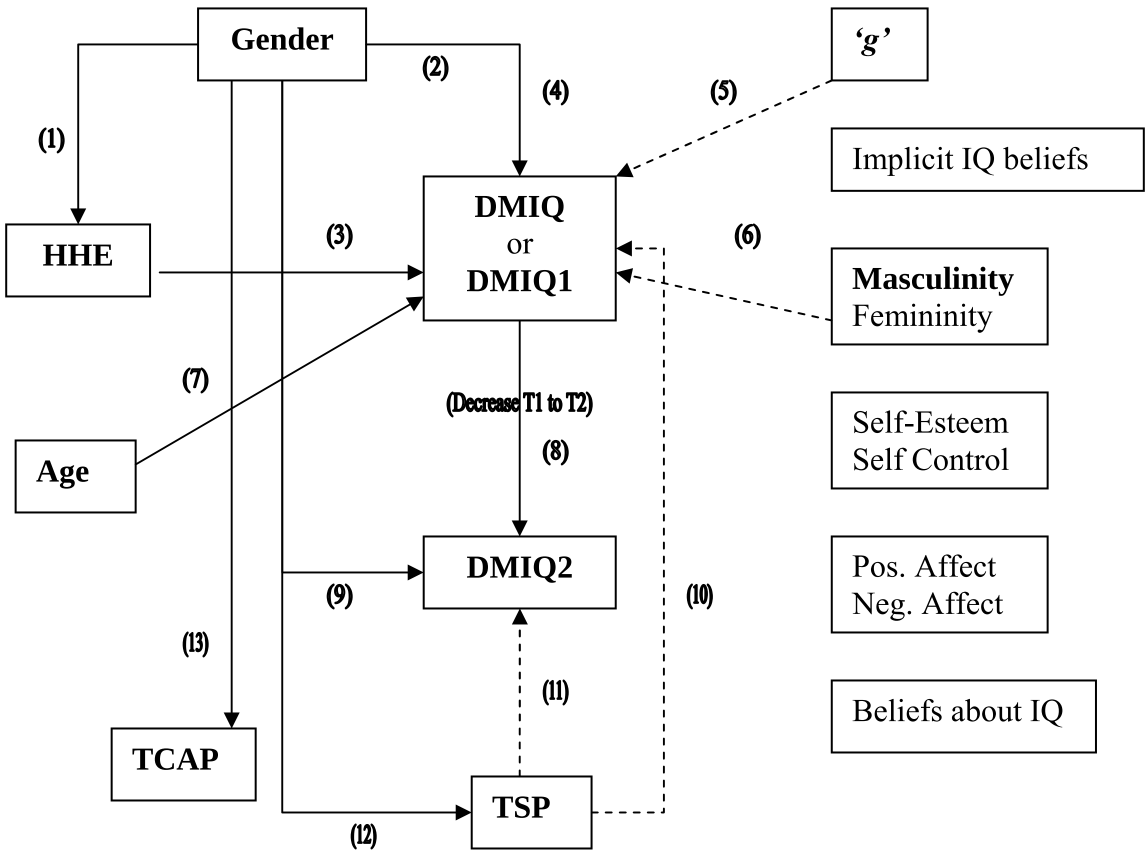
Furthermore, gender was the best predictor of the intelligence type (53%) over and above the various predictors, i.e. general intelligence (‘*g*’), implicit intelligence beliefs, beliefs about intelligence, gender identity and affect variables, self-constructs, task-confidence and the psychometric task. General intelligence (13%), masculinity (20%), and task-confidence (13% for the pre-estimation and 20% for the post-estimation condition) were the only variables to ‘*challenge*’ gender’s role as the best predictor of the domain-masculine intelligence type.

Moreover, the five experimental studies revealed that gender was the best predictor of the intelligence type in the pre-task estimation condition (20%), whereas task-confidence was the best predictor in the post-task estimation condition (13%). So, task-confidence in one’s ability to succeed on a gender stereotype-inducing task was a better determinant of the domain-masculine intelligence type than gender itself, but only after the actual task was completed.

Likewise, the psychometric task influenced the provided self-estimates in all experimental studies, with ‘*corrected*’ , i.e. significantly decreased, post-task estimates provided by both genders. Yet, despite the observed estimation ‘*correction*’, male estimates continued to be significantly higher than the female estimates, maintaining the existence of the hubris-humility effect in the post-task estimation condition. Hence, humility increased whilst hubris remained.

The results also affirmed the predicted male advantage in task-confidence (21%) and in the various psychometric tasks (21%), providing further support for the literature in the respective fields. Subsequently, the results also demonstrated that individuals are capable of accurate self-estimates that match their confidence levels. The findings were more complex in regards to the psychometric task as the supplied self-estimates did not accurately match the ability group of the provider. As the psychometric tasks were likely to activate cognitive gender-stereotypical biases, it was probable that those biases impacted on the accuracy of the provided self-estimates. Thus, while task-confidence positively influenced the accuracy of self-estimates, the psychometric tasks depressed the accuracy of self-estimates.

Figure 8.1.1: Determinants of gender differences in DMIQ Type and occurrence of HHE on DMIQ



Legend:
Arrow 1: Study 1, 2 ,3 ,4 ,5 ,6 ,7, 8, 9, 10, 11, 12, 13, 14, 15
Arrow 2: Study 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
Arrow 3: Study 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
Arrow 4: Study 1, 2, 5, 6, 10, 13, 14, 15
Arrow 5: Study 3, 4,
Arrow 6: Study 7, 8 ,9
Arrow 7: Study 3, 4, 7, 8, 9
Arrow 8: Study 11, 12, 13, 14, 15
Arrow 9: Study 13, 14
Arrow 10: Study11, 12
Arrow 11: Study 11, 12, 15
Arrow 12: Study 11, 12, 13
Arrow 13: Study 11, 12, 15

8.2. Limitations and future research

There are several limitations to the studies reported in this thesis. Firstly, most studies, with exception of Studies 10 and 11 were conducted with moderately small samples (N < 200). Yet, the sample size in the majority of studies was considered acceptable (N > 100). Only Studies 3, 9, 13, 15 used smaller samples (N <100) and those four studies were part of multiple study chapters (Chapters 3, 5, and 7) that aimed at validating or replicating the findings of all included studies. In order words, these four studies were similar to the remaining acceptable sample size studies in their respective chapters. However, the issue of moderately small sample sizes limited the possibility of using more sophisticated statistical techniques, such as SEM and LISREL.

To counter the sample size issue, the fifteen data sets were combined in a single data set (N = 2292) and the key hypotheses recomputed in order to validate the main objectives of this thesis but in a substantially larger data set. These findings are reported in the Appendix. The main objectives of this thesis were satisfactorily replicated with the combined total dataset, providing further support for the findings of the individual studies. Similarly, the majority of SEI programme studies have used comparable sample sizes (e.g. Furnham & Baguma, 1999; Furnham & Chamorro-Premuzic, 2005; Furnham & Fong, 2000; Furnham & Mottabu, 2004), making the observed results acceptable.

The second limitation of this research is the fact that some of the studies were conducted with undergraduate students from British universities (Studies 3, 4, 5), while several studies were a mixture of general public and undergraduates (Studies 1, 7, 12, 14), which may have impacted on the generalisability of the results. On the other hand, Studies 2, 6, 8, 9 and 10 were conducted with general public, foreign participants and in a precocious population. The experimental studies (Studies 11, 13, 15) were also conducted with a wide range of participants from the general public making these findings more robust and generalisable. Nonetheless, more heterogeneous samples could have yielded more varied self-estimates of ability, and in particular DMIQ, and possibly impact the observed results. It should be noted that numerous studies in the SEI research programmes have used university students as well as members of the general public and the SEI gender difference findings, and in particular the occurrence of HHE, was replicated in almost all studies and across geographies and cultures.

The third limitation was the age of the participants. Although many participants were recruited from the general public, a sizeable number of participants were in their early 20s and 30s, with few *'extreme'* cases of elderly participants as well as participants that were seventeen years old. Given that age has been shown to play a role in psychometric and self-estimated intelligence (e.g. Beier & Ackerman, 2001, 2003: Rammstedt & Rammsayer, 2002a, b) it is possible that it might have impacted upon the results. As a counter-preventive measure, age was included in the correlational and partial correlational analyses to explore its role in the intelligence type and the relevant variables. With exception of Studies 3, 4, 7, 8, and 9, age did not play a role in the prediction of DMIQ. The observed relationships between age and DMIQ indicated that older subjects provided higher DMIQ estimates, except in Study 7, where younger participants provided higher DMIQ estimates. These findings are in agreement with the literature in the field (Ackerman, 2000; Ackerman & Rolfhus, 1999; Beier & Ackerman, 2001, 2003; Deary, 2001; Deary et al., 2003; Rammstedt & Rammsayer, 2002a, b).

The fourth and main limitation of this study was the fact that DMIQ was assessed through a single estimate that could have been influenced by numerous factors, such as mood fluctuation, fatigue, fear, lack of concentration, socially desirable responding, and stress, at the time of estimation. As such it is possible that the acquired estimates were not only subjective but also unreliable. Still, DMIQ is an individualised score based on a combination of two scores: the mathematical/logical and spatial estimates. Similarly, numerous studies about the accuracy of 'subjective' assessments have shown that individuals are capable of accurate self-assessments of ability and that the current SEI measures are valid proxies of intellectual competence (Ackerman, 2002; Chamorro-Premuzic et al., 2010; Swim, 1994). Equally, the introduction of the experimental studies with multiple measurements of DMIQ estimates was intended to reduce the possible affects of 'subjective' measurement. The experimental findings replicated the earlier correlation results, providing further support for the observed results.

Based on the findings of this thesis that largely affirmed the main objectives, the main recommendation for future research is the employment of more sophisticated statistical analyses, such as SEM that allow for in-depth and simultaneous examination of multiple causal relationships and assumptions, which was not done in this thesis. In addition, as SEM allows for both confirmatory and exploratory modelling, and is thus suited for theory testing as well as theory development, it is an ideal contributor to the SEI research programme. Recent studies have demonstrated that the usage of sophisticated techniques and models, such as SEM yield more reliable data as well as expose faulty assumptions that were made using traditional statistical techniques (e.g. Chamorro-Premuzic et al., 2010; van der Sluis, 2010; von Stumm et al., 2009).

Likewise, studies with diverse and large study samples, preferably international, are recommended in order to produce more robust and generalisable results. In particular, future studies should focus on using non-student samples. Studies using British samples should try to include other cultures in order to ascertain that the magnitude of the observed gender differences in SEI is comparable and pan-cultural. Such studies should also try to determine what role does culture play in the observed gender differences in the self-estimated intelligence model, and in particular in the domain-masculine intelligence type. Similarly, it is recommended that studies concerned with gender differences in SEI should use the domain-masculine intelligence type as it has been shown to be the most sensitive predictor.

Equally, asking male and female participants whether they perceive the individual self-estimated intelligences as masculine or feminine could help the understanding of self-perceptions and the gender-stereotypical biases since they were shown to play a role in the domain-masculine intelligence type. Likewise, future studies should focus on the role cognitive gender-stereotypical biases play in the self-estimation process as well as their role in the observed gender differences in SEI, i.e. the hubris-humility effect.

It is also advisable to replicate the precocious population study, preferably internationally, to validate the herein reported results.

Lastly, based on the results of this thesis, it is recommended that the future studies about gender differences in SEI continue to employ experimental designs in order to explore what mechanisms play a role and/or influence the self-estimation of intelligence process.

References

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence, 22, 227-257.*

Ackerman, P. L. (2000). Domain-specific knowledge as the 'dark matter' of adult intelligence: gf/gc, personality and interest correlates. *Journal of Gerontology: Psychological sciences, 55B*, 69-84.

Ackerman, P. L. (2006). Cognitive sex differences and mathematics and science achievement. *American Psychologist, 61*, 722-723.

Ackerman, P. L., Beier, M. E., & Bowen, K. R. (2000). Explorations of crystallized intelligence: Completion tests, cloze tests, and knowledge. *Learning and Individual Differences, 12,* 105-121.

Ackerman, P. L., Beier, M. E., & Bowen, K. R. (2002). What we really know about our abiltieis and our knowledge. *Personality and Individual Differences, 33*, 587-605.

Ackerman, P. L., Bowen, K. R., Beier, M. E., & Kanfer, R. (2001). Determinants of individual differences and gender differences in knowledge. *Journal of Educational Psychology, 93*, 797-825.

Ackerman, P. L., & Rolfhus, E. L. (1999). The locus of adult intelligence: Knowledge, abilities, and non-ability traits. *Psychology and Aging, 14*, 314-330.

Ackerman, P. L., & Wolman, S. D. (2007). Determinants and validity of self-estimates of abilities and self-concept measures. *Journal of Applied Experimental Psychology, 13,* 57-78.

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation and the Better-Than-Average effect. *Journal of Personality and Social Psychology, 68*, 804-825.

Allik, J., Realo, A., Mottus, R., Borkenau, P., Kuppens, P., & Hrebickova, M. (2010). How people see others is different from how people see themselves: A replicable pattern across cultures. *Journal of Personality and Social Psychology, 99*, 870-882.

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science, 12*, 385-390.

Arden, R., & Plomin, R. (2006). Sex differences in variance of inteligence across childhood. *Personality and Individual Differences, 41,* 39-48.

Ariely, D. (2008*). Predictably irrational. The hidden forces that shape our decisions.* London: HarperCollinsPublishers.

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*, 29-46.

Aronson, J., & Steele, C. M. (2005). Stereotypes and the fragility of human competence, motivation, and self-concept. In C. Dweck & E. Elliot (Eds.), *Handbook of competence & motivation.* Guilford: New York.

Arrindell, W. A., Eisemann, M., Richter, J., Oei, T. P. S., Caballo, V. E., van der Ende, J., Sanavio, E., Bages, N., Feldman, L., Torres, B., Sica, C., Iwawaki, S., & Hatzichrisou, Ch. (2003). Masculinity-femininity as a national characteristic and its relationship with national agarophobic fear levels: Fodor's sex role hypothesis revitalized. *Behaviour Research and Therapy, 41*, 795-807.

Ayalon, H. (2003). Women and men go to university: Mathematical background and gender differences in choice of field in higher education. *Sex Roles, 48,* 277-290.

Baddeley (1968). A 3 minute reasoning test based on grammatical transformation. *Psychometric Science, 10,* 341-342.

Baenninger, M.A., & Newcombe, N. (1995). Environmental input to the development of sex-related differences in spatial and mathematical ability. *Learning and Individual Differences, 7*, 363-379.

Bailey, K.G., & Lazar, L. (1976). Accuracy of self-ratings of intelligence as a function of sex and level of ability in college students. *Journal of Genetic Psychology, 129*, 279-290.

Baker, D. P., Jones, D. P. (1993). Creating gender equality: Cross-national gender stratification and mathematical performance. *Sociology of Education, 66*, 91-103.

Baltes, P. B., & Schaie, K. W. (1976). On the plasticity of intelligence in adulthood and old age : Where Horn and Donaldson fail. *American Psychologist, 31*, 720-725.

Baltes, P. B. & Staundinger, U. M. (1993). The search for a psychology of wisdom. *Current Directions in Psychological Science, 2,* 75-80.

Baltes, P. B. & Staundinger, U. M. (2000). A metaheuristic (pragmatic) to orchestrate mind and virtue toward excellence. *American Psychologist, 55*, 122-136.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.

Barnett, S. M., & Ceci, S. J. (2002). When are where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128,* 612-637.

Bartsch, T. W., & Nesselroade, J. R. (1973). Test of the trait-state anxiety distinction using a manipulative, factor-analytic design. *Journal of Personality and Social Psychology, 27,* 58-64.

[Bates, T. C., & Shieles, A. (2003). Crystallized intelligence as a product of speed and drive for experience: the relationship of inspection time and openness to *g* and Gc. *Intelligence, 31,* 275-287.]

Bartlett, M. S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal statistical Society, 16* (Series B), 296-298.

Batey, M., Chamorro-Premuzic, T., & Furnham, A., (2009). Intelligence and personality as predictors of divergent thinking: The role of general, fluid and crystallised intelligence. *Thinking Skills and Creativity, 4,* 60-69.

Bedard, K., & Cho, I. (2010). Early gender test score gaps across OECD countries. *Economics of Education Review, 29,* 348-363.

Beier, M. E., & Ackerman, P. L. (2001). Current events knowledge in adults: An investigation of age, intelligence, and non-ability determinants. *Psychology and Aging, 16,* 615-628.

Beier, M. E., & Ackerman, P. L. (2003). Determinants of health knowledge: An investigation of age, gender, abilities, personality, and interests. *Journal of Personality and Social Psychology, 84 ,* 439-448.

Beloff, H. (1992). Mother, father and me: our intelligence. *The Psychologist*, *5,* 309-311.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42,* 155-162.

Bem, S., L. (1981a). Bem *Sex Role Inventory: Professional manual*. Palo Alto, CA: Consulting Psychologists Press.

Bem, S. L. (1981b). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88*, 354-364.

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioural and Brain Science, 11,* 169-232.

Benbow, C. P., Lubinski, D., Shea, D. L., & Eftekhari-Sanjani, H. (2000). Sex differences in mathematical reasoning ability at age 13: Their status 20-years later. *Psychological Science, 11*, 474-480.

Benbow, C. P., & Stanley, J. C. (1996). Inequity in equity: how "equity" can lead to inequity for high-potential students. *Psychology, Public Policy, and Law, 2,* 249-292.

Bennett, M. (1996). Men's and women's self-estimates of intelligence. *Journal of Social Psychology, 136,* 411-412.

Bennet, M. (1997). Self-estimates of ability in men and women. *Journal of Social Psychology, 137,* 530-541.

Betsworth, D. (1999). Accuracy of self-estimated abilities and the relationship between self-estimated abilities and realism for women. *Journal of Career Assessment, 7*, 35-43.

Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology, 59*, 960-970.

Beyer, S. (1998). Gender differences in self-perception and negative recall bias. *Sex Roles, 38*, 103-133.

Beyer, S. (1999). Gender differences in the accuracy of grade expectations and evaluations. *Sex Roles, 41*, 279-296.

Beyer, S., & Bowden, E. (1997). Gender differences in self-perceptions. *Personality and Social Psychology Bulletin, 23,* 157-172.

Biernat, M. (1991). Gender stereotypes and the relationship between masculinity and femininity: A developmental analysis. *Journal of Personality and Social Psychology, 61,* 351-365.

Binet, A., & Simon, T. (1905). New methods for the diagnosis of the intellectual level of subnormals. *L'Annee Psychologique, 11*, 191-244. Reprinted in J. J. Jenkins & D.G. Paterson (Eds.) *Studies of individual differences. The search for intelligence* (Elizabeth S. Kite, translator). New York: Appleton-Century –Crofts.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*, 246-263.

Bleeker, M. M., & Jacobs, J. E. (2004). Achievement in math and science: do mothers' beliefs matter 12 years later? *Journal of Educational Psychology, 96*, 97-109.

Bond, M. (1991). *Beyond the Chinese Face.* Hong Kong: Oxford University Press.

Bonnot, V., & Croizet, J-C. (2007). Stereotype internalization and women's math performance: The role of interference in working memory. *Journal of experimental Social Psychology, 43*, 857-866.

Borkenau, P., & Liebler, A. (1993). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality and Social Psychology, 65*, 546-553.

Bouchard, T. J. (1998). Genetic and environmental influences on adult intelligence and special mental abilities. *Human Biology, 70,* 257-279.

Bouchard, T. J., Lykken, D. T., McGue, M. Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: the Minnesota Study of Twins Reared Apart. *Science, 250*, 223-228.

Boyatzis, R., Goleman, D., & Rhee, K. (2000). Clustering competence in emotional intelligence: insights from the emotional competence inventory (ECI). In R. Bar-On & J.D.A. Parker (eds.): *Handbook of emotional intelligence* (pp. 343-362). San Francisco: Jossey-Bass.

Brody, N. (2003). What Sternberg should have concluded. *Intelligence, 31,* 339–342.

Brown, R. P., & Josephs, R. A. (1999). A burden of proof: Stereotype relevance and gender differences in math performance. *Journal of Personality and social Psychology, 76*, 246-257.

Brown, R. P., & Pinel, E. C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology, 39,* 626-633.

Brunner, M., Krauss, S., & Kunter, M. (2008). Gender differences in mathematics: Does the story need to be rewritten? *Intelligence, 36,* 403-421.

Bryon, M. (2006). *The Ultimate Psychometric Test Book*. London: Kogan Page Ltd.

Burson, K. A., Larrick, R. P., Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology, 90*, 60-77.

Byrd, M. & Stacey, B. (1993). Bias in IQ testing. *The Psychologist, 16,* 16-17.

Caramata, S., & Woodcock, R. (2006). Sex differences in processing speed: Developmental effects in males and females. *Intelligence, 34*, 231-252.

Carr, M., Hettinger-Steiner, H., Kyser, B., & Biddlecomb, B. (2008). A comparison of predictors of early emerging gender differences in mathematics competency. *Learning and Individual Differences, 18*, 61-75.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies.* New York: Cambridge University Press.

Casey, M. B. (1996). Gender, sex, and cognition: Considering the interrelationship between biological and environmental factors. *Learning and Individual Differences, 8,* 39-53.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 40,* 153-193.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence. A critical experiment. *Journal of Educational Psychology*, *54,* 1-22.

Ceci, S. J. (1990). *On Intelligence: more or less: A bio-ecological treatise on intellectual development*. Englewood Cliffs, NJ: Prentice-Hall Century Psychology Series.

[Chamorro-Premuzic, T. (2007). *Personality and Individual Differences.* Oxford, UK: Blackwell Publishing.]

Chamorro-Premuzic, T., & Arteche, A. (2008). Intellectual competence and academic performance: Preliminary validation of a model. *Intelligence, 36,* 564-573.

Chamorro-Premuzic, T., Arteche, A., Furnham, A., & Tricot, N. (2009). Assessing pupils' intelligence through self, parental, and teachers estimates. *Educational Psychology, 29,* 83-97.

Chamorro-Premuzic, T., & Furnham, A. (2005). *Personality and intellectual competence.* New Jersey: Laurence Erlbaum Associates.

Chamorro-Premuzic, T., Harlaar, N., Greven, C. U., & Plomin, R. (2010).

More than just IQ: A longitudinal examination of self-perceived abilities as predictors of academic performance in a large sample of UK twins. *Intelligence, 38*, 385-392.

Chamorro-Premuzic, T., Furnham, A., & Moutafi, J. (2004). The relationship between estimated and psychometric personality and intelligence score. *Journal of Research in Personality, 38*, 505-513.

Chamorro-Premuzic, T., Moutafi, J., & Furnham, A. (2005). The relationship between personality traits, subjectively-assessed and fluid intelligence. *Personality and Individual Differences, 38,* 1517-1528.

Chatard, A., Guimond, S., & Selimbegovic, L. (2007). "How good are you in math?" The effect of gender stereotypes on students' recollection of their school marks. *Journal of Experimental Social Psychology, 43*, 1017-1024.

Cohen, J. L. (1988). *Statistical power analysis for the behavioural sciences (2nd Ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cohen, J. L. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Cole, N. S. (1997). *The ETS gender study : How females and males perform in educational settings.* Princeton, NJ: Educational Testing Service.

Colom, R., & Garcia-Lopez, O. (2002). Sex differences in fluid intelligence among high school graduates. *Personality and Individual Differences*, *32,* 445-451.

[Colombia – Colombian Business Etiquette (n.d.). [WWW page].] URL http://www.cyborlink.com/besite/colombia.htm. Downloaded 15 October 2010.

Cooke-Simpson, A., & Voyer, D. (2007). Confidence and gender differences on the Mental Rotations Test. *Learning and Individual Differences, 17,* 181-186.

Costa, P. T., Jr. & McCrae, R.R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology, 38,* 668-678.

Crawford, M., Chaffin, R., & Fitton, L. (1995). Cognition in social context. *Learning and Individual Differences, 7,* 341-362.

Critcher, C. R., & Dunning, D. (2009). How chronic self-views influence (and mislead) self-assessments of task performance: Self-views shape bottom-up experiences with the task. *Journal of Personality and Social Psychology, 9*, 931-945.

Crombie, G., Sinclair, N., Silverthorn, N., Byrne, B. M., DuBois, D. L., & Trinneer, A. (2005). Predictors of young adolescents' math grades and course enrolment intentions: Gender similarities and differences. *Sex Roles, 52,* 351-367.

[Cronbach, L. J. (1949). *Essentials of Psychological Testing.* New York: Halpern.]

Cross, S.E., & Madson, :L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin, 122*, 5-37.

Csikszentmihalyi, M., Rathunde, K., & Whalen, S., (1997). *Talented Teenagers: The roots of success and failure.* Cambridge, UK: Cambridge University Press.

Dar-Nimrod, I. (2007). Math ability in women- nature versus nurture. *Nano Today, 2*, 56.

Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science, 314*, 435.

Davidson, J. E., & Downing, C. E. (2000). Contemporary Models of Intelligence. In R.J. Sternberg (Ed.), *Handbook of intelligence.* New York: Cambridge University Press.

Davies, A. P. C., & Shackelford, T. K. (2006). An evolutionary psychological perspective on gender similarities and differences. *American Psychologist, 61,* 640-641.

Deary, I. J. (2001). *Intelligence: A very short introduction*. Oxford, UK: Oxford University Press.

Deary, I. J., Thorpe, G., Wilson, V., Starr, J. M., & Whalley, L. J. (2003). Population sex differences in IQ at age 11: The Scottish Mental Survey, 1932. *Intelligence, 31*,533-542.

De Mooij, M., & Hofstede, G. (2010). The Hofstede model: Applications to global branding and advertising strategy and research. *International Journal of Advertising, 29*, 85-110.

DeNisi, A. S., & Shaw, J. B. (1977). Investigation of the uses of abilities. *Journal of Applied Psychology, 62*, 641-644.

Devilly, G.J. (2004). *The Effect Size Generator for Windows: Version 2.3* [Computer programme]. The Centre for Neuropsychology, Swinburne University, Australia. URL http://www.clintools.com/effectsize/effect_size_generator.html. Downloaded using gopher software, 1 April 2010.

de Vise, D. (2010). Report: More women than men in U.S. earned doctorates last year for first time. *Washington Post,* Downloaded 14 September 2010.

Diekman, A. B., Eagly, A. H., & Kulesa, P. (2002). Accuracy and bias in stereotypes about the social and political attitudes of women and men. *Journal of Experimental Social Psychology, 38,* 268-282.

Diener, C. I., & Dweck, C. S. (1980). An analysis of learned helplessness: II. The processing of success. *Journal of Personality and Social Psychology, 39*, 940-952.

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16,* 939-944.

Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98*, 198-208.

[Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology, 58,* 568-581.]

Dunning, D., Meyerowitz. J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessment of ability. *Journal of Personality and Social Psychology, 57*, 1082-1090.

Dweck, C.S. (1999). *Self-theories: Their role in motivation, personality and development.* Philadelphia, PA: Psychology Press.

Dykiert, D., Gale, C. R., & Deary, I. J. (2009). Are apparent sex differences in mean IQ scores created in part by sample restriction and increased male variance ? *Intelligence, 37,* 42-47.

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly, 11*, 135-172.

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychology of Women's Quarterly, 18,* 585-609.

Eccles-Parsons, J. E., Adler, T., & Meece J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology, 46*, 26-43.

Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. *Journal of Personality and Social Psychology, 84*, 5-17.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national pattern of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136,* 103-127.

Epley, N., & Dunning, D. (2000). Feeling "Holier than thou": Are self-serving assessments produced by errors in self- or social prediction? *Journal of Personality and Social Psychology, 79,* 861-875.

Eysenck, H. J. (1998). *Intelligence: A new look.* New Brunswick: Transaction Publishers.

Eysenck, H.J. (1990). *Know Your Own IQ*. Harmondsworth: Penguin.

Eysenck, H. J. (1994). *Test Your IQ*. London: Thorsons.

Facon, B. (2006). Does age moderate the effect of IQ on the differentiation of cognitive abilities during childhood? *Intelligence, 34,* 375-386.

[Fast, L. A., Lewis, J. L., Bryant, M. J., Bocian, K. A., Cardullo, R. A., Rettig, M., Hammond, K. A. (2010). Does math self-efficacy mediate the effect of the perceived classroom environment on standardized math test performance? *Journal of Educational Psychology, 102,* 739-740.]

Favia, L., & Fontane, A-M. (1997). Adolescents' personal conceptions of intelligence. *European Journal of Psychology of Education, 12*, 51-62.

Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist, 43*, 95-101.

Feingold, A. (1996). Cognitive gender differences: where are they, and why are they there? *Learning and Individual Differences, 8,* 25-32.

Ferriman, K., Lubinski, D., & Benbow, C. P. (2009). Work preferences, life values, and personal views of top math/science graduate students and the profoundly gifted: Developmental changes and gender differences during emerging adulthood and parenthood. *Journal of Personality and Social Psychology, 97*, 517-532.

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

Field, A. (2005). *Discovering Statistics Using SPSS* (2nd Eds.). London, UK: Sage.

Flynn, J. R. (1984). The mean IQ of Americans. Massive gains 1932 to 1978. *Psychological Bulletin, 95*, 29-51.

Flynn, J. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*, 171-191.

Forbes, C. E., & Schmader, T. (2010). Retraining attitudes and stereotypes to affect motivation and cognitive capacity under stereotype threat*. Journal of Personality and Social Psychology, 99*, 740-754.

Francis, L. J., & Wilcox, C. (1998). The relationship between Eysenck's personality dimensions and Bem's masculinity and femininity scales revisited. *Personality and Individual differences, 25*, 683-687.

Furnham, A. (1999). Sex differences in self-estimates of lay dimensions of intelligence. *Psychological Reports, 85*, 349-350.

Furnham, A. (2000). Parents' estimates of their own and their children's multiple intelligences. *British Journal of Developmental Psychology, 18,* 583-594.

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. (2002). Self-estimates of ten multiple intelligences: Sex differences and the perception of famous people. *European Psychologist, 7,* 245-254.

Furnham, A. (2003). *Meaning and Measurement of Intelligence Questionnaire.* Unpublished questionnaire, University College London, UK.

Furnham, A. (2004). Are lay people lumpers or splitters? The factor structure of, and sex differences related to, self-rated and other-rated abilities. *Learning and Individual Differences, 14,* 153-168.

[Furnham, A. (2005). Self-estimated intelligence, psychometric intelligence and personality. *Psychologia, 3,* 182-192.]

Furnham, A., & Akande, A., (2004). African parents' estimates of their own and their children's multiple intelligences. *Current Psychology: Developmental, Learning, Personality, Social, 22,* 281-294.

Furnham, A. & Baguma, P. (1999). Self-estimates of intelligence: A cross-cultural study from three continents. *North American Journal of Psychology, 1,* 69-78.

Furnham, A., & Budhani, S. (2002). Sex differences in the estimated intelligence of school children. *European Journal of Personality, 16,* 201-219.

Furnham, A., & Bunclark, K. (2006). Sex differences in parents' estimations of their own and their children's intelligence. *Intelligence, 34*, 1-14.

Furnham, A., Callahan, I., & Akande, D. (2004). Self-estimates of intelligence: A study in two African countries. *Journal of Psychology, 138*, 265-285.

Furnham, A., & Chamorro-Premuzic, T. (2005). Estimating one's own and one's relatives' multiple intelligence: A study from Argentina. *Spanish Journal of Psychology, 8,* 12-20.

Furnham, A., & Chamorro-Premuzic, T. (2006). Personality, intelligence and general knowledge. *Learning and Individual Differences, 16,* 79-90.

Furnham, A., Clark, K. & Bailey, K. (1999). Sex differences in estimates of multiple intelligences. *European Journal of Personality, 13*, 247-259.

Furnham, A., Crawshaw, J., & Rawles, R. (2006). Sex differences in self-estimates on two validated IQ test subscale scores. *Journal of Applied Social Psychology, 26,* 417-440.

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, Fong, G., & Martin, N. (1999). Sex and cross-cultural differences in the estimated multifaceted intelligence quotient score for self, parents and siblings. *Personality and Individual Differences, 26,* 1025-1034.

Furnham, A., & Fukumoto, S. (2008). Japanese parents' estimates of their own and their children's multiple intelligences: Cultural modesty and moderate differentiation. *Japanese Psychological Research, 50*, 63-76.

Furnham, A. & Gasson, L. (1998). Sex differences in parental estimates of their children's intelligence. *Sex Roles, 38*, 151-162.

Furnham, A., Hosoe, T. & Tang, T. (2003). Male hubris and female humility? A cross-cultural study of ratings of self, parental and sibling multiple intelligence in America, Britain, and Japan. *Intelligence*, *30,* 101-115.

[Furnham, A., Kidwai, A., & Thomas, C. (2001). Personality, psychometric intelligence and self-estimated intelligence, *Journal of Social Behaviour and Personality, 16,* 97-114. *]*

Furnham, A., & Mkhize, N. (2003). Zulu mothers' beliefs about their own and their childrens' intelligence. *Journal of Social Psychology, 143*, 83-94.

Furnham, A., Mkhize, N., & Mndaweni, T. (2004). Indian and isiZulu-speaking South African parents' estimates of their own and their children's intelligence. *South African Journal of Psychology, 34*, 364-385.

Furnham, A., & Mottabu, R. (2004). Sex and culture differences in the estimates of general and multiple intelligence: A study comparing British and Egyptian students. *Individual Differences Research, 2,* 82-94.

Furnham, A., Moutafi, J., & Chamorro-Premuzic, T. (2005). Personality and Intelligence: Gender, the Big Five, self-estimated and psychometric intelligence. *International Journal of Selection and Assessment,* 11-24.

Furnham, A., & Petrides, K. V. (2004). Parental estimates of five types of intelligence. *Australian Journal of Psychology, 56,* 10-17.

Furnham, A., Rakow, T. & Mak, T. (2002). The determinants of parents' beliefs about intelligence of their children: a study from Hong Kong. *International Journal of Psychology, 37*, 343-352.

Furnham, A., Rakow, T., Sarmany-Schiller, F., & De Fruyt, F. (1999). European differences in self-perceived multiple intelligences. *European Psychologist, 4,* 131-138.

Furnham, A. & Rawles, R. (1995). Sex differences in the estimation of intelligence. *Journal of Social Behaviour and Personality, 10*, *741-745.*

Furnham, A. & Rawles, R. (1999). Correlations between self-estimated and psychometrically measured IQ. *Journal of Social Psychology, 139*, *405-410*.

Furnham, A., Reeves, E., & Budhani, S. (2002). Parents think their sons are brighter than their daughters. *Journal of Genetic Psychology, 163,* 24-29.

Furnham, A., Shahidi, S. & Baluch, B. (2002). Sex and culture differences in perceptions of estimated multiple intelligences for self and family: A British-Iranian comparison. *Journal of Cross-cultural Psychology, 33*, 270-285.

Furnham, A. & Thomas, C. (2004). Parents' gender and personality and estimates of their own and their children's intelligence. *Personality and Individual Differences, 37,* 887-903.

Furnham, A., von Stumm, S., Makendrayogam, A., & Chamorro-Premuzic, T., (2009). A taxonomy of self-estimated human performance. The general factor í. *Journal of Individual Differences, 30*, 188-193.

Furnham, A. & Ward, C. (2001). Sex differences, test experience and the self-estimation of multiple intelligences. *New Zealand Journal of Psychology, 30,* 52-59.

Furnham, A., Wytykowska, A., & Petrides, K. (2005). Estimates of multiple intelligences : A study in Poland. *European Psychologist, 10,* 51-59.

Gallagher, A. M., & De Lisi, R. (1994). Gender differences in Scholastic Aptitude Test – Mathematics problem solving among high-ability students. *Journal of Educational Psychology, 86,* 204-211.

Garcia-Marques, L., Santos, A. S. C., & Mackie, D. M. (2006). Stereotypes : Static abstractions or dynamic knowledge structures*? Journal of Personality and Social Psychology, 91*, 814-831.

Gardner, H. (1983). *Frames of Mind: The Theory of Multiple Intelligences.* NY: Basic Books.

Gardner, H. (1993). *Multiple Intelligences: The Theory in Practice*. New York: Basic Books.

Gardner, H. (1999). *Intelligence Reframed: Multiple Intelligences for the 21st Century.* New York: Basic Books.

Gardner, H. (2006). On failing to grasp the core of MI theory: A response to Visser et al. *Intelligence, 34*, 503-505.

Gati, I., Fishman-Nadav, Y., & Shiloh, S. (2006). The relations between preferences for using abilities, self-estimated abilities, and measured abilities among career counselling clients. *Journal of Vocational Behaviour, 68*, 24-38.

Gibbs, B. G. (2010). Reversing fortunes or content change? Gender gaps in math-related skill throughout childhood. *Social Sciences Research, 39,* 540-569.

Goleman, D. (1995). *Emotional intelligence*. NY: Bantam.

Goleman, D. (1998). *Working with emotional intelligence*. NY: Bantam Books.

Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology, 29,* 17-28.

Gottfredson, L.S. (1997a). Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography. *Intelligence, 24*, 13-23.

Gottfredson, L. S. (1997b). Why *g* matters: The complexity of everyday life. *Intelligence, 24*, 79-132.

Gottfredson, L.S. (2000). Intelligence., In *Encyclopaedia of Sociology* (2nd Ed.), Borgatta, E.F. & Montgomery, J. V. (eds)., NY: Macmillan, 1359-1386.

Gottfredson, L.S. (2003a). Dissecting practical intelligence theory: Its claims and evidence. *Intelligence, 31,* 343–397.

Gottfredson, L.S. (2003b). On Sternberg's 'Reply to Gottfredson. *Intelligence, 31,* 415–424.

Gottfredson, L. S. (2003c). The challenge and promise of cognitive career assessment. *Journal of Career Assessment, 11*, 115-135.

Gottfredson, L. S. (2004a). Intelligence: is it the epidemiologists' elusive 'fundamental cause' of social class inequalities in health? *Journal of Personality and Social Psychology, 86,* 174-199.

Gottfredson, L. S. (2004b). Schools and the 'g' factor. *Wilson Quarterly*, Summer, 35-45.

[Gottfredson, L. S. (2005). What if the hereditarian hypothesis is true? *Psychology, Public Policy, and Law, 11*, 311-319.]

Griffin, D. W., Dunning, D., & Ross, L. (1990). The role of construal process in overconfident predictions about the self and others. *Journal of Personality and Social Psychology, 59*, 1128-1139.

Guenther, C. L., & Alicke, M. D. (2010). Deconstructing the Better-Than-Average effect. *Journal of Personality and Social Psychology, 99*, 755-770.

Guimond, S., Brunot, S., Chatard, A., Garcia, D. M., Martinot, D., Branscombe, N. R., Buunk, A. P., Desert, M., Haque, S., & Yzerbyt, V. (2007). Culture, gender and the self: Variations and impact of social comparison processes. *Journal of Personality and Social Psychology, 92*, 1118-1134.

Guimond, S., Martinot, D., Chatard, A., Crisp, R. J., & Redersdorff, S (2006). Social comparison, self-stereotyping, and gender differences in self-construals. *Journal of Personality and Social Psychology, 90*, 221-242.

Hall, J. A., & Carter, J. D. (1999). Gender-stereotype accuracy as an individual difference. *Journal of Personality and Social Psychology*, *77,* 350-359.

Halpern, D. F. (2000). *Sex difference is cognitive abilities* (3rd ed.). Mahwah, NJ: Erlbaum.

[Halpern, D. F. (2002). A cognitive-process taxonomy for sex differences in cognitive abilities. *Current Directions in Psychological Science, 13,* 135-139.]

Halpern, D.F., Benbow, C. P., Geary, D.C., Gur, R. C., Hyde, J.S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8,* 1-51.

Halpern, D. F., & Wright, T. M. (1996). A process-oriented model of cognitive sex differences. *Learning and Individual Differences, 8,* 3-24.

Harrington, T. F., & Schafer, W. D. (1996). A comparison of self-reported abilities and occupational ability patterns across occupations. *Measurement and Evaluation in Counselling and Development, 28,* 180-190.

Harrison, D. A., & Shaffer, M. A. (1994). Comparative examinations of self-reports and perceived absenteeism norms: Wading through Lake Wobegon*. Journal of Applied Psychology, 79*, 240-251.

Hartmann, P. (2006). Spearman's law of diminishing returns. A look at age differentiation. *Journal of Individual Differences, 27*, 199-207.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128.

Hedges, L. V., & Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. FL: Orlando: Academic Press.

Hedges, L.V., & Novell, A. (1995). Sex differences in mental scores, variability, and numbers of high-scoring individuals. *Science, 269,* 41-45.

Herrnstein, R.J., & Murray, C. (1994). *The Bell Curve: Intelligence and class structure in American life.* NY: Free Press.

Hirschy, A. J., & Morris, J. R. (2002). Individual differences in attributional style: The relational influence of self-efficacy, self-esteem, and sex role identity. *Personality and Individual Differences, 32*, 183-196.

Hoffman, C., & Hurst, N. (1990). Gender stereotypes: Perception or rationalization? *Journal of Personality and Social Psychology, 58*, 197-208.

Hofstede, G. (1998). *Masculinity and Femininity: The Taboo Dimension of National Cultures*. Thousand Oaks, CA: Sage Publications.

Hofstede, G. (2003). *Culture's Consequences**,** Comparing Values, Behaviours, Institutions, and Organizations across Nations*. (2nd Ed.)Thousand Oaks, CA: Sage Publications.

Hofstede, G. (n.d.). Cultural dimensions. [WWW page]. URL http://www.geert-hofstede.com. Downloaded 15 October 2010.

Hogan, H. (1978). IQ self-estimates of males and females. *Journal of Social Psychology, 106,* 137-138.

Holling, H., & Preckel, F. (2005). Self-estimates of intelligence – methodological approaches and gender differences. *Personality and Individual Differences, 38*, 503-517.

Hunt, E. (2000). Let's hear it for crystallized intelligence. *Learning and Individual Differences, 12,* 123-129.

Husain, M., & Millimet, D. L. (2009). The mythical 'boy crisis?' *Economics of Education Review, 28,* 38-48.

Hyde, J. S. (1996). Gender and cognition: A commentary on current research. *Learning and Individual Differences, 8*, 33-38.

[Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*, 581-592.]

[Hyde, J. S. (2006). Gender similarities still rule. *American Psychologist, 61*, 641-642.]

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990a). Gender differences in mathematics performance : A meta-analysis. *Psychological Bulletin, 107,* 139-155.

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990b). Gender comparisons of mathematics attitudes and affect. A meta-analysis. *Psychology of Women Quarterly, 14,* 299-324.

Hyde, J. S., & Linn, M.C. (1988). Gender differences in verbal ability : A meta-analysis. *Psychological Bulletin, 104*, 53-69.

Inzlicht, M., & Kang, S. K. (2010). Stereotype threat spillover: How coping with treats to social identity affects aggression, eating, decision making, and attention. *Journal of Personality and Social Psychology, 99*, 467-481.

Irwing, P., Cammock, T., & Lynn, R. (2001). Some evidence for the existence of a general factor of semantic memory and its components. *Personality and Individual Differences, 30*, 857-871.

Jacklin, C. N. (1989). Female and male: The issue of gender. *American Psychologist, 44*, 127-133.

Jackson, D. N., & Rushton, J. P. (2006). Males have greater *g*: sex differences in general mental ability from 100,000 17- to 18-year olds on the Scholastic Assessment Test. *Intelligence*, *34,* 479-486.

Jacobs, J. E., & Eccles, J. S. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and childrens' ability perceptions. *Journal of Personality and Social Psychology, 63*, 932-944.

Jensen, A. R. (1998). *The 'g' factor: The science of mental abilities*. London: Praeger.

Johnson, W., & Bouchard, T. J., Jr. (2007). Sex differences in mental abilities: *g* masks the dimensions on which they lie. *Intelligence, 35*, 23-39.

Jonsson, A.-C., & Allwood, C. M. (2003). Stability and variability in the realism of confidence judgments over time, content domain, and gender. *Personality and Individual Differences, 34,* 559-574.

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalised self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology, 83*, 693-710.

Kahneman, D., & Tversky, A. (2000*). Choices, values, and frames*. Cambridge: Cambridge University Press.

Kaiser, (1970). A second generation Little Jiffy. *Psychometrika, 35,* 401-415.

Kawakami, K., Steele, J. R., Cifa, C., Phills, C. E., & Dovidio, J. F. (2008). Approaching math increases math = me and math = pleasant. *Journal of Experimental Social Psychology, 44,* 818-825.

Keltikangas-Jäarvinen, L., Kivimäki, M., & Keskivaara, P. (2003). Parental practices, self-esteem and adult temperament: 17-year follow-up study of four population-based age cohorts. *Personality and Individual Differences, 34*, 431-447.

Kessel, C., & Linn, M. C. (1996). Grades or scores: Predicting future college mathematics performance. *Educational Measurement: Issues and Practice, 15,* 10–14.

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes and women’s math performance: How implicit gender-math stereotypes influence women’s susceptibility to stereotype threat. *Journal of Experimental Social Psychology, 43,* 825-832.

Kim, H. C. (2006). *A comparative study on gifted education for mathematics in Korea and foreign countries.* Unpublished master's thesis, Dankook University, South Korea.

Kim, Y.-H., Chiu, Ch.-y., & Zou, Z. (2010). Know thyself : Misperceptions of actual performance undermine achievement motivation, future performance, and subjective well-being. *Journal of Personality and Social Psychology, 99*, 395-409.

Klar, Y., & Giladi, E.E. (1999). Are most people happier than their peers, or are they just happy? *Personality and Social Psychology Bulletin, 25*, 585-594.

Klein, P. D. (1997) Multiplying the problems of intelligence by eight: A critique of Gardner's theory. *Canadian Journal of Education, 22*, 377-394.

Kling, K. C., Hyde, S. J., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin, 125*, 470-500.

Kornilova, T. V., Kornilov, S. A., & Chumakova, M. A. (2009). Subjective evaluations of intelligence and academic self-concept predict academic achievement: Evidence from a selective student population. *Learning and Individual Differences, 19,* 596-608.

Kray, L. J., Galinsky, A., Thompson, L. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology, 80,* 942-958.

Krueger, J., & Mueller, R. (2002). Unskilled, unaware of both? The Better-Than-Average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology, 82*, 180-188.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognising one’s own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*, 1124-1134.

Kruger, J., & Dunning, D. (2002). Unskilled or unaware – but why? A reply to Krueger and Mueller (2002). *Journal of Personality and Social Psychology, 82*, 189-192.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology, 86*, 148-161.

Kwan, V.S., John, O., Robin, R., & Kuang, L. (2008). Conceptualizing and assessing self-enhancement bias: A componential approach. *Journal of Personality and Social Psychology, 94,* 1062-1077.

Levine, S. C., Huttenlocher, J., Taylor, A., Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology, 35,* 940-949.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organisational Behaviour and Human Performance, 20,* 159-183.

Linn, M. C., & Hyde, J. (1989). Gender, mathematics, and science. *Educational Researcher, 18,* 17-27.

Linn, M. C., & Peterson, A. C., (1985). Emergence and characterisation of sex differences in spatial ability: A meta-analysis. *Child development, 56*, 1479-1498.

Linver, M. R., & Davis-Kean, P. E. (2005). The slippery slope: What predicts math grades in middle and high school? *New Directions for Child and Adolescent Development, 110,* 49-64.

Lippa, R. A. (2001). On deconstructing and reconstructing masculinity-femininity. *Journal of Research in Personality, 35*, 168-207.

Loehlin, J. C., Horn, J. M., & Willerman, L (1997). Heredity, environment, and IQ in the Texas Adoption Project. In Sternberg, R. J. & Grigorenko, E (eds), *Intelligence, Heredity and Environment*. Cambridge: Cambridge University Press.

[Lohman, D. F.(2001). Fluid intelligence, inductive reasoning, and working memory: Where the theory of Multiple Intelligences falls short. In N. Colangelo & S.

Assouline (Eds.), *Talent Development IV: Proceedings from the 1998*. Henry B. & Jocelyn Wallace National Research Symposium on talent development. Scottsdale, AZ: Gifted Psychology Press.]

Lubinski, D., & Benbow, C. P. (2000). States of excellence. *American Psychologist, 55*, 137-150.

Lubinski, D., & Humphreys, L. G. (1990). A broadly based analysis of mathematical giftedness. *Intelligence, 14*, 327-355.

Lynn, R. (1997). Geographic variations in intelligence. In H. Nyborg (Ed.). *The scientific study of human nature.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Lynn, R. (1999). Sex differences in intelligence and brain size: A developmental theory. *Intelligence, 27,* 1-12.

Lynn, R., Allik, J., & Irwing, P. (2004). Sex differences on three factors identified in Raven's Standard Progressive Matrices. *Intelligence, 32,* 411-424.

Lynn, R., & Irwing, P. (2002). Sex differences in general knowledge, semantic memory and reasoning ability. *British Journal of Psychology, 9*, 545-556.

Lynn, R., & Irwing, P. (2004). Sex differences on the progressive matrices: A meta-analysis. *Intelligence, 32*, 481-498.

Lynn, R., Irwing, P., & Cammock, T. (2002). Sex differences in general knowledge*. Intelligence, 30*, 27-39.

Lynn, R., Wilberg, S., & Margraf-Stiksrud, J. (2004). Sex differences in general knowledge in German high school students. *Personality and Individual Differences, 37,* 1249-1255.

Lytton, H., & Romney, D. M. (1991). Parents' differential socialization of boys and girls: A meta-analysis. *Psychological Bulletin, 109*, 267-296.

Mabe, P. A. III, & West, S. A. (1982). Validity of self-evaluations of ability: A review and meta-analysis. *Journal of Applied Psychology, 67*, 280-296.

Maccoby, E.E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford: CA: Stanford University Press.

Maddux, W. M., & Brewer, M. B. (2005). Gender differences in the relational and collective bases of trust. *Group Processes and Intergroup Relations, 8*, 159-171.

Marsh, H.W. (1990). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review, 2*, 77-172.

Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: gender differences in development of math and English constructs. *American Educational Research Journal, 35*, 705-738.

Marusic, I., & Bratko, D. (1998). Relations of masculinity and femininity with personality dimensions of the Five-factor model. *Sex Roles, 38,* 29-44.

Massa, L. J., Mayer, R. E., & Bohon, L. M. (2005). Individual differences in gender role beliefs influence spatial ability test performance. *Learning and Individual Differences, 15*, 99-111.

Matthews, G., Davies, R. D., Westerman, S. J., & Stammers, R. B. (2000). *Human performance*. East Sussex: Psychology Press.

Mayer, J.D., Roberts, R.D. & Barsade, S. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology, 59,* 507-536.

Mayr, E. (1982). *The growth of biological thought*. MA: Cambridge, Belknap Press.

McArdle, J. J., Hamagami, F., Meredith, W., Bradway, K. P. (2000). Modeling the dynamic hypotheses of Gf-Gc theory using longitudinal life-span data. *Learning and Individual Differences, 12*, 53-79.

McSweeney, B. (2002). Hofstede's 'Model of National Cultural Differences and Consequences: A Triumph of Faith - A Failure of Analysis, *Human Relations, 55,* 89-118.

Meece, J. L., Glienke, B. B., & Burg, S. (2006). Gender and motivation. *Journal of School Psychology, 44*, 351-373.

Meelissen, M., & Luyten, H. (2008). The Dutch gender gap in mathematics: Small for achievement, substantial for beliefs and attitudes. *Studies in Educational Evaluation, 34,* 82-93.

Mensa UK (n.d.). Mensa: The high IQ society. [WWW page]. URL http://www.mensa.org.uk/mensa/. Downloaded 30 July 2010.

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

Mischel, W. & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualising situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102,* 246-268.

Moore, D. A., & Small, D. A. (2007). Error and bias in comparative judgment: On being both better and worse than we think we are. *Journal of Personality and Social Psychology, 92,* 972-989.

Nasser, R., & Singhal, S. (2006). How youth in India and Lebanon rate their intelligence. *Journal of Social Sciences, 2*, 93-99.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and Unknowns. *American Psychologist, 51,* 77-101.

Neto, F., Furnham, A., & Paz, R. (2007). Sex and culture differences in perceptions of estimated multiple intelligence for self and family: A Macanese-Portuguese comparison. *International Journal of Psychology, 42*, 124-133.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, Me = female, therefore math ≠ me. *Journal of Personality and Social Psychology, 83*, 44-59.

Novell, A., & Hedges, L.V.(1998). Trends in gender differences in academic achievement from 1960-1994: An analysis of differences in mean, variance and extreme scores. *Sex Roles, 39,* 21-43.

Ogle, L., Sen, A., Pahlke, E., Jocelyn, L., Kostberg, D., Roye, S., & Williams, T. (2003). *International comparisons in fourth-grade reading literacy: Findings from the Progress in international Literacy Study (PIRLS) of 2001* (NCES 2003-073). Washington, DC: U.S. Government Printing Office.

Pallant, J. (2007). *SPSS Survival Manual: A step by step guide to data analysis using SPSS for Windows* (3rd edition). New York: Open University Press.

Pallier, G. (2003). Gender differences in the self-assessment of accuracy on cognitive tasks. *Sex Roles, 48*, 265-276.

Paulhus, D. L. (1991). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*, 598-609.

Paulhus, D. L., Lysy, D., & Yik, M. (1998). Self-report measures of intelligence: are they useful as proxy IQ tests? *Journal of Personality, 66*, 523-555.

Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement : Exploring the early grades and the extremes. *Social Science Research, 37,* 239-253.

Peterson, E. R., & Whiteman, M. C. (2007). “I think I can, I think I am…”: The interrelationships among self-assessed intelligence, self-concept, self-efficacy and the personality trait intellect in university students in Scotland and New Zealand. *Personality and Individual Differences, 43*, 959-968.

Petiprin, G. L., & Johnson, M. E. (1991). Effects of gender, attributional style, and item difficulty on academic performance. *Journal of Personality, 125*, 45-50.

Petrides, K.V., Furnham, A., & Martin, G. N. (2004). Estimates of emotional and psychometric intelligence: evidence for gender-based stereotypes. *The Journal of Social Psychology, 144,* 149-162.

Petrides, K.V., Pita, R., Kokkinaki, F. (2007). The location of trait emotional intelligence in personality factor space. *British Journal of Psychology, 98*, 273-289.

Pinker, S. (2002). *The blank slate: The modern denial of human nature.* New York: Viking.

Plomin, R. (1999). Genetics and general cognitive ability. *Nature, 402* (Suppl.), C25-29.

Plomin, R., & Spinath, F. M. (2004). Intelligence: genetics, genes, and genomics. *Journal of Personality and Social Psychology, 86*, 112-129.

Pomerantz, E. M., & Ruble, D. N. (1997). Distinguishing multiple dimensions of conceptions of ability: implications for self-evaluation. *Child Development, 68,* 1165-1680.

Poropat, A. (2010). How smart do you think you are? A meta-analysis on the validity of self-estimated intelligence scores. Unpublished manuscript. *Personality and Individual Differences.*

Preckel, F., Zeidner, M., Goetz, T., Schleyer, E. J. (2008). Female 'big fish' swimming against the tide: The 'big-fish-little-pond effect' and gender –ratio in special gifted classes. *Contemporary Educational Psychology, 33*, 78-96.

Prediger, D., (1999). Basic structure of work-relevant abilities. *Journal of Counselling Psychology, 46,* 173-184.

Pullman, H., & Allik, J. (2000). The Rosenberg Self-Esteem Scale: its dimensionality, stability and personality correlates in Estonian. *Personality and Individual Differences, 28*, 701-715.

Rammstedt, B., & Rammsayer, T. H (2000a). Sex-related differences in time estimation: the role of personality. *Personality and Individual Differences, 29*, 301-312.

Rammstedt, B., & Rammsayer, T. H. (2000b). Sex differences in self-estimates of different aspects of intelligence. *Personality and Individual Differences, 29*, 869-880.

Rammstedt, B., & Rammsayer, T. H. (2001). Gender differences in self-estimated intelligence in children and early adolescents. *German Journal of Educational Psychology, 15,* 207-217.

Rammstedt, B., & Rammsayer, T. H. (2002a). Gender differences in self-estimated intelligence and their relation to gender-role orientation. *European Journal of Personality, 16,* 382-396.

Rammstedt, B., & Rammsayer, T. H. (2002b). Self-estimated intelligence. Gender differences, relationship to psychometric intelligence and moderating effects of level of education. *European Psychologist, 7*, 275-284.

Reilly, J., & Mulhern, G. (1995). Gender differences in self-estimated IQ: The need for care in interpreting group data. *Journal of Personality, 66,* 523-555.

Reis, S. M., & Renzulli, J.S. (2010). Is there still a need for gifted education ? An examination of current research. *Learning and individual differences, 20,* 308-317.

Roberts, T.-A. (1991). Gender and the influence of evaluations on self-assessments in achievement settings. *Psychological Bulletin, 109*, 297-308.

Rojahn, J., & Naglieri, J. A. (2006). Developmental gender differences on the Naglieri nonverbal ability test in a nationally normed sample of 5-17 year olds. *Intelligence, 34*, 253-260.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Roznowski, M., Reith, J., & Hong, S. (2000). A further look at youth intellectual giftedness and its correlates: Values, interests, performance and behaviour. *Intelligence, 28*, 87-113.

Rudman, L. A., & Phelan, J. E. (2010). The effect of priming gender roles on women's implicit gender beliefs and career aspirations. *Social Psychology, 41*, 192-202.

Rushton, J. P., & Jensen, A. R. (2005a). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law, 11*, 235-294.

Rushton, J. P., & Jensen, A. R. (2005b). Wanted more race realism, less moralistic fallacy. *Psychology, Public Policy, and Law*, *11,* 328-336.

Ryckman, R. M., Gold, J. A., Rodda, W. C. (1971). Confidence rating shifts and performance as a function of locus of control, self-esteem, and initial task performance. *Journal of Personality and Social Psychology, 18,* 305-310.

Rydell, R. J., Rydell, M. T., & Boucher, K, L. (2010). The effect of negative performance stereotypes on learning. *Journal of Personality and Social Psychology, 99*, 883-896.

[Saklofske, D.H., Yang, Z., Zhu, J., & Austin, E. J. (2008). Spearman's law of diminishing returns in normative sample for the WISC-IV and WAIS-III. *Journal of Individual Differences, 29*, 57-69.]

Sax, L. J. (1994). Mathematical self-concept: How college reinforces the gender gap. *Research in Higher Education, 35*, 141-166.

Sax, L. J., & Harper, C. E. (2007). Origins of the gender gap: Pre-college and college influences on differences between men and women. *Research in Higher Education, 48,* 669-694.

Segall, M., Dasen, P., Berry, J, & Poortinga, Y. (1999). *Human behaviour in global perspective: An introduction to cross-cultural psychology.* New York: Allyn & Bacon.

Seibt, B., & Forster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology, 87*, 38-56.

Scarr, S. (1985) An authors frame of mind: Review of Frames of mind: The theory of multiple intelligences. *New Ideas in Psychology, 3,* 95-100.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85,* 440-452.

Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg Self-Esteem scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology, 89,* 623-642.

Schroeder-Davis, S. (1993). Coercive egalitarianism: Subverting achievement through neglect and hostility. *Gifted Education Press Quarterly, 7*, 2-9.

[Schwartzer, R., (1993). *Measurement of perceived self-efficacy: Psychometric scales for cross-cultural research.* Berlin: Frei Universiteit Press.]

[Schwartzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs.* Windsor, UK: NFER-NELSON: 35-37.]

Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology, 93*, 604-614.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10,* 80-83.

Sleeper, L. A., & Nigro, G. N. (1987). It's not who you are but who you're with: Self-confidence in achievement settings. *Sex Roles, 16*, 57-69.

Smith, L. L., & Reise, S. P. (1998). Gender differences on negative affectivity: An IRT study of differential item functioning on the Multidimensional Personality Questionnaire Stress Reduction Scale. *Journal of Personality and Social Psychology, 75*, 1350-1362.

Spearman, C. (1904). 'General intelligence', objectively determined and measured. *American Journal of Psychology, 15,* 201–293.

Spelke, E.S. (2005). Sex differences in intrinsic aptitude for mathematics and science? A critical review. *American Psychologist, 60,* 950-958.

Stankov, L. (2000). The theory of fluid and crystallized intelligence. New findings and recent developments. *Learning and Individual Differences, 12*, 1-3.

Stankov, L., Boyle, G., & Cattell, R. B. (1995). Models and paradigms in intelligence research. In D. Saklofske and M. Zeidner (Eds.), *International Handbook of Personality and Intelligence.* New York: Plenum.

Steele, C. M. (1997). A threat in the air: How stereotypes shape the intellectual identities and performance. *American Psychologist, 52*, 613-629.

Steele, J. R., & Ambady, N. (2006). "Math is hard!". The effect of gender priming on women's attitudes. *Journal of Experimental Social Psychology, 42*, 428-436.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797-811.

Steinmayr, R., Beauducel, A., & Spinath, B. (2010). Do sex differences in a faceted model of fluid and crystallised intelligence depend on the method applied? *Intelligence, 38,* 101-110.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. NY: Cambridge University Press.

Sternberg, R. J. (1990). *Metaphors of mind: Conceptions of the nature of intelligence.* New York: Cambridge University Press.

Sternberg, R. J. (1991). Death, taxes, and bad intelligence tests. *Intelligence, 15,* 257-270.

Sternberg, R. J., Conway, B.E., Ketron, J., & Bornstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology, 41*, 37-55.

Sternberg, R. J., & Davidson, J. E. (2005). *Conception of giftedness.* 2nd edition Boston, MA: Cambridge University Press.

Sternberg, R. J., Forsythe, G. G., Hedlund, J., Horvath, J. A., Wagner, R. K., Williams, W. M., et al. (2000). *Practical intelligence in everyday life.* New York: Cambridge University Press.

[Sternberg, R. J., & Grigorenko, E. L. (1997). *Intelligence, heredity, and environment.* New York: Cambridge University Press.]

Sternberg, R. J., & Grigorenko, E. L. (Eds.) (2002). *The general intelligence factor: How general is it?* Mahwah, NJ: Erlbaum.

Sternberg, R. J., & Grigorenko, E. L. (2006). Cultural intelligence and successful intelligence. *Group and Organisation Management, 31,* 27-39.

Stieger, S., Kastner, C., Voracek, M., von Stumm, S., Chamorro-Premuzic, T., & Furnham, A. (2010). Independent effects of personality and sex on self-estimated intelligence: Evidence from Austria. *Psychological Reports, 107,* 553-563.

[Stipek, D., & Gralinski, H. (1999). Gender differences in children's achievement-related beliefs and emotional responses to success and failure in mathematics. *Journal of Educational Psychology, 83,* 361-371.]

Storek, J. (2007). *Task Success Probability Estimation Measure.* Unpublished measure, University College London, UK.

Stumpf, H., & Stanley, J. C. (1996). Gender-related differences on the College Board's advanced placement and achievement tests, 1982-1992. *Journal of Educational Psychology, 88,* 353-364.

Swami, V., & Furnham, A. (2010). Self-assessed intelligence: Inter-ethnic, rural-urban, and sex differences in Malaysia. *Learning and Individual Differences, 20,* 51-55.

Swami, V., Furnham, A., & Kannan, K. (2006). Estimating self, parental, and partner multiple intelligences: A replication in Malaysia. *Journal of Social Psychology, 146*, 645-655.

Swami, V., Furnham, A., Maakip, I., Ahmad, M. S., Nawi, N., H. M., Voo, P. S. K., Christopher, A. N. , & Garwood, J. (2008). Beliefs about the meaning and measurement of intelligence: A cross-cultural comparison of American, British, and Malaysian undergraduates. *Applied Cognitive Psychology, 22*, 235-246.

Swim, J. K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *Journal of Personality and Social Psychology, 66,* 21-36.

Szymanowicz, A., Chamorro-Premuzic, T., & Furnham, A. (2011). *Gender differences in self-estimates of general, mathematical, spatial, and verbal intelligence: Four meta-analyses*. Unpublished manuscript. University College London, UK.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 85,* 710-718.

Terman, L. M. (1925). *Genetic studies of genius: Vol. 1. Mental and physical traits of a thousand gifted children*. Stanford, CA: Stanford University Press.

Terman, L. M., & Merrill, M. A. (1937). *Measuring intelligence.* Boston: Houghton Mifflin.

The College Board, Office of Research and Development (1998). *Research Summary: SAT and Gender Differences*. NY: New York.

Thurstone, L. L. (1919). Mental tests for college entrance. *Journal of Educational Psychology, 10*, 129-142.

Turner, J. (1999). Some current issues in research on social identity and self-categorization theories. In N. Ellemers, R. Spears, B. Doosje (Eds.), *Social Identity,* (pp. 6-34). Oxford, England: Blackwell.

University of Kent (2007). *Numerical reasoning, spatial and crystallised intelligence aptitude problems. Careers Services*. [WWW page]. URL http://www.kent.ac.uk/careers/ test.htm. Downloaded May, 2007.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotation, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*, 599-604.

Van der Sluis, S., Derom, C., Thiery, E., Bartels, M., Polderman, T. J. C., Verhulst, F. C., Jacobs, N., van Gestel, S., de Geus, E. J. C., Dolan, C. V., Boomsma, D. I., & Posthuma, D. (2008). Sex differences on the WISC-R in Belgium and Netherlands. *Intelligence, 36,* 48-67.

Van der Sluis, S., Vinkhuyzen, A. A. E., Boomsma, D. I., & Posthuma, D. (2010). Sex differences in adults' motivation to achieve. *Intelligence, 38,* 433-446.

Vispoel, W. P., & Forte Fast, E. E. (2000). Response Biases and Their Relation to Sex Differences in Multiple Domains of Self-Concept. *Applied Measurement in Education, 13,* 79-97.

Visser, B.A., Ashton, M.C., & Vernon, P. A. (2006a). Beyond *g*: Putting multiple intelligence theory to the test. *Intelligence, 34*, 487-502.

Visser, B. A., Ashton, M. C., & Vernon, P. A. (2006b). *g* and the measurement of multiple intelligences: a response to Gardner. *Intelligence, 34*, 507-510.

Visser, B. A., Ashton, M. C., & Vernon, P. A. (2008). What makes you think you're so smart?: Measured abilities, personality, and sex differences in relation to self-estimates of multiple intelligences*. Journal of Individual Differences,* 29, 35-44.

Vlachos, F., Andreou, G., & Andreou, E. (2003). Biological and environmental influences in visuospatial abilities. *Learning and Individual Differences, 13,* 339-347.

von Stumm, S., Chamorro-Premuzic, T., & Furnham, A. (2009). Decomposing self-estimates of intelligence: Structure and sex differences across 12 nations. *British Journal of Psychology, 100*, 429-442.

Voyer, D., & Doyle, R. A. (2010). Item type and gender differences on the Mental Rotation Test. *Learning and Individual Differences, 20,* 469-472.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological Bulletin, 117,* 250-270.

Waterhouse, L. (2006). Inadequate evidence for multiple intelligences, Mozart effect, and emotional intelligence theories. *Educational Psychologist, 41,* 247-255.

Watkins, M. W. (2000). *Monte Carlo PCA for Parallel Analysis*. Software.

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

Watson, D., & Tellegen, A., (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219-235.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology, 76*, 820-838.

Wechsler, D., (1944). *Measurement of adult intelligence* (3rd ed.). Baltimore: William and Wilkins.

Weinberg, R. A. (1989). Intelligence and IQ: Landmark issues and great debates. *American Psychologist, 44*, 98-104.

Weisgram, E. S., & Bigler, R. S. (2006). Girls and science careers: The role of altruistic values and attitudes about scientific tasks. *Journal of Applied Developmental Psychology, 27,* 326-348.

Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behaviour: A review of possible mechanisms. *Psychological Bulletin, 127*, 797-826.

White, J. (2005). *Howard Gardner: The myth of Multiple Intelligence.* View Point, 16. Institute of Education, University of London.

[Wikipedia, the free encyclopedia (n.d.). Affect Psychology. [WWW page]. URL http://en.wikipedia.org/wiki/Affect_(psychology). Dowloaded 30 June 2010.]

Wikipedia, the free encyclopedia (n.d.). Effect size. [WWW page]. URL http://en.wikipedia.org/wiki/Effect_size. Dowloaded 5 November 2010.

Wikipedia, the free encyclopedia (n.d.). Sex and gender distinction. [WWW page]. URL http://en.wikipedia.org/wiki/Sex_and_gender_distinction. Downloaded 17 November 2010.

Wilcox, C., & Francis, L. J. (1997). Beyond gender stereotyping: Examining the validity of the Bem Sex-Role Inventory among 16- to 19-yearold females in England. *Personality and Individual Differences, 23,* 9-13.

Wober, M. (1973). East African undergraduates' attitudes concerning the concept of intelligence. *British Journal of Social and Clinical Development, 12*, 431-432.

Wonderlic Personnel Test, Inc. (1992). *Wonderlic Personnel Test and Scholastic Level Exam: User's Manual*. Libertyville, IL. Author.

Xie, Y., & Shauman, K. A. (2003). *Women in science: Career processes and outcomes.* Cambridge: Harvard University Press.

Yang, S., & Sternberg, R. J. (1997). Conceptions of intelligence in ancient Chinese philosophy. *Journal of Theoretical and Philosophical Psychology, 17*, 101-119.

Yuen, M., & Furnham, A. (2005). Sex differences in self-estimation of multiple intelligences among Hong Kong Chinese adolescents. *High Ability Studies, 16,* 187-199.

Zoomerang (n.d.). [Online survey engine]. URL http://www.zoomerang.com. Used between 2007-2010.

Appendix

9.1. Study 14A

Gender, TCAP, Seven Task Success Probability Probes as Predictors of the Domain-Masculine Intelligence Type

9.1.1. Introduction

Study 14A contains the results of the first experimental condition of Study 14. The analyses reported here concern the seven TSP probe condition. Study 14A uses an independent population sample. The aim of Study 14A is to validate the results of Studies 11 to 14, while examining the impact of the increased number of TSP probes.

Thus, HHE is expected to occur on DMIQ1 and DMIQ2 (H1). Significant decrease in DMIQ estimates is predicted from T1 to T2 (H2). Gender differences, i.e. male advantage, are expected to be observed on the psychometric task (TCAP) (H3). Likewise, gender differences are also expected in task-success probability estimation (TSP), with males being more confident about their abilities than females (H4). As in all previous studies, gender is expected as the best determinant of DMIQ1 (H5) and DMIQ2 (H6).

9.1.2. Method

Participants

Forty-eight participants took part in this study. There were 28 females (58%) and 20 males. Their age ranged from 21 to 60 (M = 24.43, SD = 7.35) years. 67% of participants completed A-levels or similar level of education, 2% achieved non-university level of education, 15% achieved BA/BSc level, and 10% achieved MA/MSc/MBA or equivalent level of education and 4% had earned a Doctorate or a PhD degree.

Measures

Repeated Measure of Domain-masculine Intelligence Type (DMIQ)

See Study 11 (section 7.2.2). Alpha for DMIQ1 was .85 and DMIQ2 .84.

Psychometric Aptitude Task

Total Correct Aptitude Problems (TCAP) Numerical and Spatial Psychometric Aptitude Problems (University of Kent, Career Services, 2009; http://www.kent.ac.uk/careers/test.htm)

See Study 14 (section 7.5.2.).The alpha for the fifteen numerical reasoning items was .82 and for the nine spatial items .61 (the inter-item correlation was r =.16). Alpha for TCAP (al problems combined) was .83.

Task Success Probability (TSP)

Task Success Probability Estimation Measure (Storek, 2007)

See Study 11 (section 7.2.2). The alpha for the seven-item measure was .92 and the inter-item correlation was = .61.

Procedure

Participants were recruited from the general public. They were recruited through an email campaign by the main researcher and eight second-year students who participated in a mini-research study group in spring 2009 that the main researcher was leading. An email invitation, with an URL link (www.zoomerang.com) to the study and a background explanation of the study was sent to all participants. The snow-balling technique of participant recruitment was used, i.e. participants were asked to forward the study invitation and the URL link to as many acquaintances as possible. In total, 128 individuals logged onto the site during February and May 2009. The data was gathered through an online survey engine www.Zoomerang.com and participation was voluntary.

Detailed scoring instructions were given at the beginning of each measure, including timing instructions for the numerical and reasoning problems. Participants were aware that the study was approved by UCL Ethics Committee, meeting confidentiality and Data Protection requirements. Debrief feedback, correct answers for the psychometric task and a feedback box, were provided at the end. All participants were fluent in English and no language or other problems were reported.

9.1.3. Results

9.1.3.1. Domain-masculine intelligence and the Hubris and Humility Effect in T1 and T2

In order to test hypothesis one, two independent samples t-tests were computed to assess whether significant gender differences on DMIQ1 and DMIQ2. Results that are presented in Table 9.1.1. confirmed the existence of significant gender differences in both estimating conditions, with males providing higher self-estimates than females. Hypothesis 1 was confirmed.

Table 9.1.1: Overview of Independent Samples t-Tests and Effect Sizes for DMIQ1 and DMIQ2

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
DMIQ/ DMIQ1	119.78 (14.88) 20	108.77 (16.71) 28	2.35(46)*	11.01	1.59	20.43	.11	.70
DMIQ2	114.68 (11.25) 20	99.55 (19.15) 28	3.16(46)**	15.12	5.48	27.76	**.18**	**.96**

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Hedge's Adjustment *d* is adjusted for sample size and used in both tests. Large effect sizes are in bold.

To test whether significant decrease in DMIQ estimates from T1 to T2 occurred after the task, a paired-samples t-test was conducted. There was a statistically significant decrease in DMIQ1 (M = 113.35, SD = 16.74) to DMIQ2 (M = 105.85, SD = 17.85), t(47) = 4.33, p = .00, two-tailed, r = .76, p =.00, N=48. The mean decrease in DMIQ estimates was 7.50 (SD = 12.02) with 95% confidence interval ranging from 4.01 to 10.99. Cohen's d statistic (.62) indicated a medium effect size. Hypothesis 2 was confirmed.

9.1.3.2. Gender Differences in Psychometric Aptitude Problems (TCAP) and the Task Success Probability Estimation (TSP)

Table 9.1.2. gives an overview of the 2x2 χ^2 tests and effect sizes for the fifteen numerical reasoning and nine spatial problems. No significant gender differences were observed on any of the 24 problems.

Table 9.1.2: 2 x 2 Chi Square Tests and Effect Sizes for 15 Numerical, Reasoning and 9 Spatial Problems – Per Gender and % Correct Answer

		Correct Answer Wrong	Correct Answer Right	Total	Yates Continuity Correction Value for 2x2	Asymp. Sig. (2-sided)	Phi (φ) Coefficient
Block 1 Numerical Reasoning Q6	Male N	0	20	20	2.30	.13	-.29*
	% Within Gender	0%	100%	100%			
	% Within Correct Answer	0%	47%	42%			
	% of Total	0%	42%	42%			
	Female N	5	23	28			
	% Within Gender	18%	82%	100%			
	% Within Correct Answer	100%	54%	58%			
	% of Total	10%	48%	58%			
	Total N	5	43	48			
	% Within Gender	10%	90%	100%			
Q7	Male N	2	18	20	.00	1.00	-.01
	% Within Gender	10%	90%	100%			
	% Within Correct Answer	40%	42%	42%			
	% of Total	4%	38%	42%			
	Female N	3	25	28			
	% Within Gender	11%	89%	100%			
	% Within Correct Answer	60%	58%	58%			
	% of Total	6%	52%	58%			

	Total N	5	43	48			
	% Within Gender	10%	90%	100%			
Q8	Male N	9	11	20	.00	1.00	-.05
	% Within Gender	45%	55%	100%			
	% Within Correct Answer	39%	44%	42%			
	% of Total	19%	23%	42%			
	Female N	14	14	28			
	% Within Gender	50%	50%	100%			
	% Within Correct Answer	61%	56%	58%			
	% of Total	29%	29%	58%			
	Total N	23	25	48			
	% Within Gender	48%	52%	100%			
Block 2 Q10	Male N	2	18	20	.43	.51	-.15
	% Within Gender	10%	90%	100%			
	% Within Correct Answer	25%	45%	42%			
	% of Total	4%	37%	42%			
	Female N	6	22	28			
	% Within Gender	21%	77%	100%			
	% Within Correct Answer	75%	55%	58%			
	% of Total	13%	46%	58%			
	Total N	8	40	48			
	% Within Gender	17%	83%	100%			
Q11	Male N	6	14	20	.00	1.00	-.02
	% Within Gender	30%	70%	100%			
	% Within Correct Answer	40%	42%	42%			
	% of Total	13%	29%	42%			
	Female N	9	19	28			
	% Within Gender	32%	68%	100%			
	% Within Correct Answer	60%	58%	58%			
	% of Total	19%	40%	58%			
	Total N	15	33	48			
	% Within Gender	31%	69%	100%			
Q12	Male N	2	18	20	.00	1.00	-.06
	% Within Gender	10%	90%	100%			
	% Within Correct Answer	33%	43%	42%			
	% of Total	4%	38%	42%			
	Female N	4	24	28			
	% Within Gender	14%	86%	100%			
	% Within Correct Answer	67%	57%	58%			
	% of Total	8%	50%	58%			
	Total N	6	42	48			
	% Within Gender	12%	88%	100%			
Block 3 Q14	Male N	11	2	20	.00	1.00	-.02
	% Within Gender	55%	5%	100%			
	% Within Correct Answer	41%	40%	42%			
	% of Total	23%	3%	42%			
	Female N	16	12	28			
	% Within Gender	57%	43%	100%			
	% Within Correct Answer	59%	57%	58%			
	% of Total	33%	25%	58%			
	Total N	27	21	48			
	% Within Gender	56%	44%	100%			
Q15	Male N	8	12	20	.00	1.00	-.03
	% Within Gender	40%	60%	100%			
	% Within Correct Answer	40%	43%	42%			
	% of Total	17%	25%	42%			

	Female N	12	16	28			
	% Within Gender	43%	57%	100%			
	% Within Correct Answer	60%	57%	58%			
	% of Total	25%	33%	58%			
	Total N	20	28	48			
	% Within Gender	42%	58%	100%			
Q16	Male N	13	7	20	.54	.46	.15
	% Within Gender	65%	35%	100%			
	% Within Correct Answer	48%	33%	42%			
	% of Total	22%	15%	42			
	Female N	14	14	28			
	% Within Gender	50%	50%	100%			
	% Within Correct Answer	52%	67%	58%			
	% of Total	29%	29%	58%			
	Total N	27	21	48			
	% Within Gender	56%	44%	100%			
Block 4	Male N	2	18	20	.00	1.00	.05
Q18	% Within Gender	10%	90%	100%			
	% Within Correct Answer	50%	41%	42%			
	% of Total	4%	38%	42%			
	Female N	2	26	28			
	% Within Gender	7%	93%	100%			
	% Within Correct Answer	50%	59%	58%			
	% of Total	4%	54%	58%			
	Total N	4	44	48			
	% Within Gender	8%	92%	100%			
Q19	Male N	1	19	20	.31	.58	-.15
	% Within Gender	5%	95%	100%			
	% Within Correct Answer	20%	44%	42%			
	% of Total	2%	40%	42%			
	Female N	4	24	28			
	% Within Gender	14%	86%	100%			
	% Within Correct Answer	80%	56%	58%			
	% of Total	8%	50%	58%			
	Total N	5	43	48%			
	% Within Gender	10%	90%	100%			
Q20	Male N	5	15	20	.05	.83	-.08
	% Within Gender	25%	75%	100%			
	% Within Correct Answer	36%	44%	42%			
	% of Total	10%	31%	42%			
	Female N	9	19	28			
	% Within Gender	32%	68%	100%			
	% Within Correct Answer	64%	56%	58%			
	% of Total	19%	40%	58%			
	Total N	14	34	48			
	% Within Gender	29%	71%	100%			
Q21	Male N	1	19	20	.31	.58	-.15
	% Within Gender	5%	95%	100%			
	% Within Correct Answer	20%	44%	42%			
	% of Total	2%	40%	42%			
	Female N	4	24	28			
	% Within Gender	14%	86%	100%			
	% Within Correct Answer	80%	56%	58%			
	% of Total	8%	50%	58%			
	Total N	5	43	48			
	% Within Gender	10%	90%	100%			

Q22	Male N	15	5	20	.22	.64	.11
	% Within Gender	75%	25%	100%			
	% Within Correct Answer	46%	33%	42%			
	% of Total	31%	10%	42%			
	Female N	18	10	28			
	% Within Gender	64%	36%	100%			
	% Within Correct Answer	55%	67%	58%			
	% of Total	38%	21%	58%			
	Total N	33	15	48			
	% Within Gender	69%	31%	100%			
Q23	Male N	18	2	20	3.69	.06	.33*
	% Within Gender	90%	10%	100%			
	% Within Correct Answer	51%	15%	42%			
	% of Total	38%	4%	42%			
	Female N	17	11	28			
	% Within Gender	61%	39%	100%			
	% Within Correct Answer	49%	85%	58%			
	% of Total	35%	23%	58%			
	Total N	35	13	48			
	% Within Gender	73%	27%	100%			
Block 5	Male N	2	18	20	.00	1.00	-.01
Spatial	% Within Gender	10%	90%	100%			
Q25	% Within Correct Answer	40%	42%	42%			
	% of Total	4%	38%	42%			
	Female N	3	25	28			
	% Within Gender	11%	89%	100%			
	% Within Correct Answer	60%	58%	58%			
	% of Total	6%	52%	58%			
	Total N	5	43	48			
	% Within Gender	10%	90%	100%			
Q26	Male N	11	9	20	.00	1.00	-.02
	% Within Gender	55%	45%	100%			
	% Within Correct Answer	41%	43%	42%			
	% of Total	23%	19%	42%			
	Female N	16	12	28			
	% Within Gender	57%	43%	100%			
	% Within Correct Answer	59%	57%	58%			
	% of Total	33%	25%	58%			
	Total N	27	21	48			
	% Within Gender	56%	44%	100%			
Q27	Male N	13	7	20	.00	1.00	.04
	% Within Gender	65%	35%	100%			
	% Within Correct Answer	43%	39%	42%			
	% of Total	27%	15%	42%			
	Female N	17	11	28			
	% Within Gender	61%	39%	100%			
	% Within Correct Answer	57%	61%	58%			
	% of Total	35%	23%	58%			
	Total N	30	18	48			
	% Within Gender	62%	38%	100%			
Block 6	Male N	6	14	20	.00	1.00	.02
Q29	% Within Gender	30%	70%	100%			
	% Within Correct Answer	43%	41%	42%			
	% of Total	13%	29%	42%			
	Female N	8	20	28			
	% Within Gender	29%	71%	100%			
	% Within Correct Answer	57%	59%	58%			
	% of Total	17%	42%	58%			

	Total N	14	34	48			
	% Within Gender	29%	71%	100%			
Q30	Male N	8	12	20	.27	.1	.12
	% Within Gender	40%	60%	100%			
	% Within Correct Answer	50%	38%	42%			
	% of Total	17%	25%	42%			
	Female N	8	20	29			
	% Within Gender	29%	71%	100%			
	% Within Correct Answer	50%	63%	58%			
	% of Total	17%	42%	58%			
	Total N	16	32	48			
	% Within Gender	33%	67%	100%			
Q31	Male N	4	16	20	.00	1.00	-.02
	% Within Gender	20%	80%	100%			
	% Within Correct Answer	40%	42%	42%			
	% of Total	8%	33%	42%			
	Female N	6	22	28			
	% Within Gender	21%	79%	100%			
	% Within Correct Answer	60%	58%	58%			
	% of Total	13%	46%	58%			
	Total N	10	38	48			
	% Within Gender	21%	79%	100%			
Block 7	Male N	15	5	40	.00	1.00	-.04
Q33	% Within Gender	75%	25%	100%			
	% Within Correct Answer	41%	46%	42%			
	% of Total	31%	10%	42%			
	Female N	22	6	28			
	% Within Gender	79%	21%	100%			
	% Within Correct Answer	60%	55%	58%			
	% of Total	46%	13%	58%			
	Total N	37	11	48			
	% Within Gender	77%	23%	100%			
Q34	Male N	10	10	20	.48	.49	.14
	% Within Gender	50%	50%	100%			
	% Within Correct Answer	50%	36%	42%			
	% of Total	21%	21%	42%			
	Female N	10	18	28			
	% Within Gender	36%	64%	100%			
	% Within Correct Answer	50%	64%	58%			
	% of Total	21%	38%	58%			
	Total N	20	28	48			
	% Within Gender	42%	58%	100%			
Q35	Male N	3	17	20	1.03	.31	-.20
	% Within Gender	15%	85%	100%			
	% Within Correct Answer	25%	47%	42%			
	% of Total	6%	35%	42%			
	Female N	9	19	28			
	% Within Gender	32%	68%	100%			
	% Within Correct Answer	75%	53%	58%			
	% of Total	19%	40%	58%			
	Total N	12	36	48			
	% Within Gender	25%	75%	100%			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

A medium sized negative significant effect size was observed on Q6, Phi coefficient =-.29, p <.05 and a medium sized positive effect size was observed on Q23, Phi coefficient =.33, p <.05. An independent t-test for Total Correctly Solved Aptitude Problems (TCAP) failed to reach significance and confirmed that males and females did not differ in their performance on the aptitude problems, t(46) .06, p =.95; M_{Male} = 15.65, SD_{Male} = 3.78; M_{Female} = 15.57, SD_{Female} = 5.37; MD = .08, 95% CI: -2.66 to -2.82; η^2 = .01; Hedge's Adjustment d =.02. Thus, hypothesis 3 was not confirmed.

Independent samples t-tests were run for the Total Task Success Probability measure and the seven TSP probes to investigate whether significant gender differences occurred on these variables. Results revealed no significant results on the total measure and only on one out of seven TSP probes. Significant gender differences were observed on TSP5, with males (M = 3.05, SD = 1.00) providing higher estimates than females (M = 2.25, SD = 1.11); t(46) = 2.57, p < .05, two-tailed. The magnitude of the differences in the means (mean difference = .80, 95% CI: .17 to 1.43) was medium (η^2 = .13, Cohen's d =.76). Hypothesis 4 was partially confirmed.

9.1.3.3. Gender, TSP, TCAP as Predictors of DMIQ1 and DMIQ2

Firstly, the relationship between DMIQ1, DMIQ2, gender, TSP and TCAP was assessed. Results are presented in Table 9.1.3. As in earlier studies, DMIQ1 and DMIQ2 were strongly intercorrelated (r = .76, p =.00). Equally, gender correlated negatively with DMIQ1 (r = -.33, p<.05) and DMIQ2 (r = -.42, p <.01), with females providing lower scores than males. As in earlier studies, positive correlations were observed between TSP and DMIQ1 (r = .49, p =.00) and DMIQ2 (r =.61, p =.00) and between TCAP and DMIQ1 (r = .43, p <.01) and DMIQ2 (r =.39, p <.01). Likewise, a medium strength positive correlation was observed between TSP and TCAP (r =. 39, p<.01).

Table 9.1.3: Correlations, Means and Standard Deviations between DMIQ1 and DMIQ2, Gender, TSP, TCAP, and Age

	DMIQ1	DMIQ2	G	TSP	TCAP	A
	113.35	105.85	1.58	2.84	15.60	24.43
	(16.74)	(17.85)	(.50)	(.94)	(4.60)	(7.35)
DMIQ1						
DMIQ2	.76***					
Gender	-.33*	-.42**				
TSP	.49***	.61***	-.25			
TCAP	.43**	.39**	-.01	.39**		
Age	.24	.22	-.11	.27	.13	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). *N = between 46 and 48.*

As in previous studies, age was included in the analysis to examine whether it had impact on the DMIQ estimates. The age range was 39 years. No significant relationships between age and the remaining variables were observed.

In order to test whether gender was the best predictor DMIQ1 and DMIQ2, two simultaneous multiple regressions were performed. Results are shown in Table 9.1.4. The dependent variables were DMIQ1 and DMIQ2. Gender, TSP, and TCAP were the independent variables. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity.

The first model was significant F (3,44) = 8.45, p =.00, Adjusted R^2 =.32, f^2=.59), with the overall model explaining 37% of total variance. TSP (β = .31, p <.05, r_{part} = .27) and TCAP (β = .31, p <.05, r_{part} = .29) were significant predictors of

DMIQ1, explaining 7% and 8% of the variance, respectively. TCAP was the best predictor of DMIQ1, followed by TSP. Hypothesis 5 was not confirmed.

Table 9.1.4: Beta coefficients for Simultaneous Multiple Regressions of Gender, TSP and TCAP on DMIQ1 and DMIQ2

Dependent Variable N = 48	***DMIQ1*** β	*t*	***DMIQ2*** β	*t*
Gender	-.25	-2.01	**-.31**	**-2.77****
TSP	**.31**	**2.25***	**.45**	**3.69****
TCAP	**.31**	**2.38***	.21	1.81
Regression Model	**F(3, 44) = 8.45*****		**F(3,44) = 13.93*****	
R^2	.37		.49	
R^2 Change	.37		.49	
Adj. R	.32		.45	
f^2	.59		.96	

* *p* <.05, ** *p* <.01, *** *p* <.001 *Note:* Significant values are in bold.

The second model was also significant F (3,44) = 13.93, p =.00, Adjusted R^2 = .45, f^2=.96), with the overall model explaining 49% of total variance. TSP (β = .45, p <.01, r_{part} = .40) and gender (β = -.31, p <.01, r_{part} = -.30) were significant predictors of DMIQ2, explaining 16% and 9% of the variance, respectively. Contrary to prediction, TSP was the best predictor of DMIQ2, followed by gender. Hypothesis 6 was not confirmed.

Thus, hypotheses 1 and 2 were confirmed and hypothesis 4 was partially confirmed. Hypotheses 3, 5, and 6 were not confirmed.

9.1.4. Discussion

The results of the first experimental condition that included seven TSP probes revealed diverse outcomes. The existence of HHE on DMIQ1 and DMIQ2 was confirmed as was decrease in DMIQ estimates from T1 to T2. However, gender differences, i.e. male advantage, in the psychometric task, were not observed and the male advantage in task confidence was only partially confirmed. In addition, gender was not the best predictor of DMIQ1 and DMIQ2. TCAP was the predictor of DMIQ1 and TSP of DMIQ2.

9.2. Study 14B

Gender, TCAP, Four Task Success Probability Estimation Probes as Predictors of the Domain-Masculine Intelligence Type

9.2.1. Introduction

Study 14B replicates Study 14A. The design and execution is identical to Study 14A. However, only four TSP probes are used. Thus, the focus of Study 14B is whether the reduction in TSP probes impacts the predicted relationships or results in dissimilar results.

Therefore, HHE is predicted to occur on DMIQ1 and DMIQ2 (H1). A significant decrease is expected from DMIQ1 to DMIQ2, following the psychometric and confidence tasks (H2). Male advantage is predicted on TCAP (H3) as well as on the task-success probability estimation probes, with males being more confident about their abilities than females (H4). As in previous studies, gender is expected to be the best predictor of DMIQ1 (H5) and DMIQ2 (H6).

9.2.2. Method

Participants

Sixty-one participants took part in this study. There were 32 males (53%) and 29 females. Their age ranged from 17 to 50 (M = 24.08, SD = 8.61) years. 85% of participants completed A-levels or similar level of education, 7% achieved GSCE or similar level of education and 8% achieved BA/BSc level.

Measures

Repeated Measure of Domain-masculine Intelligence Type (DMIQ)

See Study 11 (section 7.2.2). Alpha for DMIQ1 was .47 and the inter-item correlation was = .32. For DMIQ2, alpha was .65 and the inter-item correlation was = .52.

Psychometric Aptitude Task

Total Correct Aptitude Problems (TCAP); Numerical and Spatial Psychometric Aptitude Problems (University of Kent, Career Services, 2009; http://www.kent.ac.uk/careers/test.htm)

See Study 14 (section 7.5.2). The alpha for the fifteen numerical reasoning items was .73 and for the nine spatial items .42 (the inter-item correlation was r =.07). Alpha for the TCAP was .74.

Task Success Probability (TSP)

Task Success Probability Estimation Measure (Storek, 2007)

See Study 11 (section 7.2.2). The four individual task success probability estimation probes made up the Task Success Probability (TSP) measure, with individual scores computed for all participants. The alpha for the four-item measure was .81 and the inter-item correlation was = .51.

Procedure

Participants were recruited from the general public through an email campaign by the main researcher and eight second year students, who participated in a mini-research study group in spring 2009 that the main researcher was leading. An email invitation, with an URL link (www.zoomerang.com) to the study and a background explanation of the study was sent to all participants. The snow-balling technique of participant recruitment was used, i.e. participants were asked to forward the study invitation and the URL link to as many acquaintances as possible. In total, 136 individuals logged on the site during February and May 2009. The data was gathered through an online survey engine www.Zoomerang.com and participation was voluntary. Detailed scoring instructions were given at the beginning of each measure, including timing instructions for the numerical and reasoning problems.

Participants were aware that the study was approved by UCL Ethics Committee, meeting confidentiality and Data Protection requirements. Debrief feedback, correct answers for the psychometric problems, together with the feedback box, were available at the end of the survey. All participants were fluent in English and no language or other problems were reported.

9.2.3. Results

9.2.3.1. HHE and DMIQ1 and DMIQ2

Two independent samples t-tests were computed to assess whether significant gender differences occurred on DMIQ1 and DMIQ2. Results are presented in Table 9.2.1., revealing that significant gender differences occurred in both DMIQ1 and DMIQ2, with males providing higher self-estimates than females, further affirming the existence of the hubris-humility effect on the DMIQ type. Hypothesis 1 was confirmed.

Table 9.2.1: Overview of Independent Samples t-Tests and Effect Sizes for DMIQ1 and DMIQ2

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
DMIQ/ DMIQ1	114.86 (8.76) 32	101.00 (9.22) 29	6.02(59)***	13.86	9.25	18.47	**.38**	**1.54**
DMIQ2	111.33 (10.87) 32	89.62 (11.48) 29	7.59(59)***	21.71	15.98	27.43	**.50**	**1.94**

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Hedge's Adjustment *d* is adjusted for sample size and used in both tests. Large effect sizes are in bold.

A paired t-test was conduced to test whether significant decrease in DMIQ occurred from T1 to T2. There was a statistically significant decrease in DMIQ estimates from T1 (*M* = 108.27, *SD* = 11.31) to T2 (*M* = 101.01, *SD* = 15.55), t(60) = 5.55, *p* = .00, two-tailed, *r* = .75, *p* =.00, N=61. The mean decrease in DMIQ estimates was 7.26 (*SD* = 10.23) with 95% confidence interval ranging from 4.64 to 9.88. Cohen's *d* statistic (.71) indicated a medium effect size. Hypothesis 2 was confirmed.

9.2.3.2. Gender Differences in TCAP and TSP

Table 9.2.2. gives an overview of the 2x2 χ^2 tests and effect sizes for the fifteen numerical reasoning and nine spatial problems. Out of the 24 problems, significant gender differences were observed on seven problems: Q8, Q10, Q13, Q18, Q20, Q25, and Q27. Males gave significantly more right answers to six problems Q8, Q10, Q13, Q18, Q20 and Q27. Only on Q25 did females significantly outperformed males. Using Cohen's (1988) effect size criteria, medium sized negative significant effect sizes were observed on the seven problems, i.e. Q25 (φ= -.47, *p*=.00), Q20 (φ = -.45, *p* =.00), Q13 (φ = -.39, *p* =.00), Q18 (φ = -.35 *p* <.01), Q8 (φ = -.34, *p* = .01), Q10 (φ = -.34, *p* =.01), and Q27 (φ = -.20, *p* =.01).

Table 9.2.2: 2 x 2 Chi Square Tests and Effect Sizes for 15 Numerical, Reasoning and 9 Spatial Problems – Per Gender and % Correct Answer

		Correct Answer		Total	Yates Continuity Correction Value for 2x2	Asymp. Sig. (2-sided)	Phi (φ) Coefficient
		Wrong	Right				
Block 1 Numerical Reasoning Q6	Male N	2	30	32	.31	.58	-.13
	% Within Gender	6%	94%	100%			
	% Within Correct Answer	33%	55%	52%			
	% of Total	3%	49%	52%			
	Female N	4	25	29			
	% Within Gender	14%	86%	100%			
	% Within Correct Answer	67%	46%	48%			

	% of Total	6%	41%	48%			
	Total N	6	55	61			
	% Within Gender	10%	90%	100%			
Q7	Male N	2	30	32	.2.58	.11	-.25*
	% Within Gender	6%	94%	100%			
	% Within Correct Answer	22%	58%	52%			
	% of Total	3%	49%	52%			
	Female N	7	22	29			
	% Within Gender	24%	78%	100%			
	% Within Correct Answer	78%	58%	48%			
	% of Total	12%	36%	48%			
	Total N	9	52	61			
	% Within Gender	15%	85%	100%			
Q8	Male N	9	23	32	.5.80	.02*	-.34**
	% Within Gender	28%	72%	100%			
	% Within Correct Answer	33%	68%	52%			
	% of Total	15%	38%	52%			
	Female N	18	11	29			
	% Within Gender	62%	38%	100%			
	% Within Correct Answer	67%	32%	48%			
	% of Total	30%	18%	48%			
	Total N	27	34	61			
	% Within Gender	44%	56%	100%			
Q9	Male N	8	24	32	.66	.42	-.14
	% Within Gender	25%	75%	100%			
	% Within Correct Answer	42%	57%	52%			
	% of Total	13%	40%	52%			
	Female N	11	18	29			
	% Within Gender	38%	62%	100%			
	% Within Correct Answer	58%	43%	48%			
	% of Total	18%	30%	48%			
	Total N	19	42	61			
	% Within Gender	31%	69%	100%			
Q10	Male N	10	22	32	5.86	.02*	-.34**
	% Within Gender	31%	69%	100%			
	% Within Correct Answer	35%	69%	52%			
	% of Total	16%	36%	52%			
	Female N	19	10	29			
	% Within Gender	66%	35%	100%			
	% Within Correct Answer	66%	31%	48%			
	% of Total	31%	16%	48%			
	Total N	29	32	61			
	% Within Gender	48%	52%	100%			
Q11	Male N	3	29	32	.78	.38	-.16
	% Within Gender	9%	91%	100%			
	% Within Correct Answer	33%	56%	52%			
	% of Total	5%	48%	52%			
	Female N	6	23	29			
	% Within Gender	21%	79%	100%			
	% Within Correct Answer	67%	44%	48%			
	% of Total	10%	38%	48%			
	Total N	9	52	61			
	% Within Gender	15%	85%	100%			
Q12	Male N	20	12	32	1.34	.25	-.18
	% Within Gender	63%	38%	100%			
	% Within Correct Answer	47%	68%	52%			

	% of Total	33%	20%	52%			
	Female N	23	6	29			
	% Within Gender	79%	21%	100%			
	% Within Correct Answer	54%	33%	48%			
	% of Total	38%	10%	48%			
	Total N	43	18	61			
	% Within Gender	70%	30%	100%			
Q13	Male N	16	16	32	7.48	.01*	-.39***
	% Within Gender	50%	50%	100%			
	% Within Correct Answer	40%	80%	52%			
	% of Total	26%	26%	52%			
	Female N	25	4	29			
	% Within Gender	86%	14%	100%			
	% Within Correct Answer	61%	20%	48%			
	% of Total	41%	6%	48%			
	Total N	41	20	61			
	% Within Gender	67%	33%	100%			
Q14	Male N	17	15	32	2.50	.11	-.24
	% Within Gender	53%	47%	100%			
	% Within Correct Answer	44%	68%	52%			
	% of Total	28%	25%	52%			
	Female N	22	7	29			
	% Within Gender	76%	24%	100%			
	% Within Correct Answer	56%	32%	48%			
	% of Total	36%	12%	48%			
	Total N	39	22	61			
	% Within Gender	64%	36%	100%			
Block 2	Male N	5	27	32	.68	.41	-.15
Q16	% Within Gender	16%	84%	100%			
	% Within Correct Answer	39%	56%	52%			
	% of Total	8%	44%	52%			
	Female N	8	21	29			
	% Within Gender	28%	72%	100%			
	% Within Correct Answer	62%	44%	48%			
	% of Total	13%	34%	48%			
	Total N	13	48	61			
	% Within Gender	21%	79%	100%			
Q17	Male N	5	27	32	.26	.61	-.11
	% Within Gender	16%	84%	100%			
	% Within Correct Answer	42%	55%	52%			
	% of Total	8%	44%	52%			
	Female N	7	22	29			
	% Within Gender	24%	76%	100%			
	% Within Correct Answer	58%	45%	48%			
	% of Total	12%	36%	48%			
	Total N	12	49	61%			
	% Within Gender	20%	80%	100%			
Q18	Male N	12	20	32	6.13	.01*	-.35**
	% Within Gender	38%	63%	100%			
	% Within Correct Answer	36%	71%	52%			
	% of Total	20%	33%	52%			
	Female N	21	8	29			
	% Within Gender	72%	28%	100%			
	% Within Correct Answer	64%	29%	48%			
	% of Total	34%	13%	48%			

	Total N	33	28	61			
	% Within Gender	54%	46%	100%			
Q19	Male N	6	26	32	.66	.42	-.14
	% Within Gender	19%	81%	100%			
	% Within Correct Answer	40%	57%	52%			
	% of Total	10%	43%	52%			
	Female N	9	20	29			
	% Within Gender	31%	69%	100%			
	% Within Correct Answer	60%	44%	48%			
	% of Total	15%	33%	48%			
	Total N	15	46	61			
	% Within Gender	25%	75%	100%			
Q20	Male N	11	21	32	10.70	.00***	-.45***
	% Within Gender	34%	66%	100%			
	% Within Correct Answer	32%	78%	52%			
	% of Total	18%	34%	52%			
	Female N	23	6	29			
	% Within Gender	79%	21%	100%			
	% Within Correct Answer	68%	22%	48%			
	% of Total	38%	10%	48%			
	Total N	34	27	61			
	% Within Gender	56%	44%	100%			
Q21	Male N	28	4	32	.00	1.00	-.03
	% Within Gender	88%	12%	100%			
	% Within Correct Answer	52%	57%	52%			
	% of Total	46%	7%	52%			
	Female N	26	3	29			
	% Within Gender	90%	10%	100%			
	% Within Correct Answer	48%	43%	48%			
	% of Total	42%	5%	48%			
	Total N	54	7	61			
	% Within Gender	89%	11%	100%			
Block 3	Male N	3	29	32	2.29	.13	-.24
Spatial	% Within Gender	9%	91%	100%			
Q23	% Within Correct Answer	27%	58%	52%			
	% of Total	5%	48%	52%			
	Female N	8	21	29			
	% Within Gender	28%	72%	100%			
	% Within Correct Answer	73%	42%	48%			
	% of Total	13%	34%	48%			
	Total N	11	50	61			
	% Within Gender	18%	82%	100%			
Q24	Male N	27	5	32	1.26	.26	.18
	% Within Gender	84%	16%	100%			
	% Within Correct Answer	57%	36%	52%			
	% of Total	44%	8%	52%			
	Female N	20	9	29			
	% Within Gender	69%	31%	100%			
	% Within Correct Answer	43%	64%	48%			
	% of Total	33%	15%	48%			
	Total N	47	14	61			
	% Within Gender	77%	23%	100%			
Q25	Male N	18	14	32	11.24	.00***	-.47***
	% Within Gender	56%	44%	100%			
	% Within Correct Answer	39%	93%	52%			
	% of Total	30%	23%	52%			
	Female N	28	1	29			
	% Within Gender	97%	3%	100%			

	% Within Correct Answer	61%	7%	48%			
	% of Total	46%	2%	48%			
	Total N	46	15	61			
	% Within Gender	75%	25%	100%			
Q26	Male N	12	20	32	2.76	.10	-.25
	% Within Gender	38%	63%	100%			
	% Within Correct Answer	40%	65%	52%			
	% of Total	20%	33%	52%			
	Female N	18	11	29			
	% Within Gender	62%	38%	100%			
	% Within Correct Answer	60%	36%	48%			
	% of Total	30%	18%	48%			
	Total N	30	31	61			
	% Within Gender	49%	51%	100%			
Q27	Male N	6	26	32	.4.75	.03*	-.31*
	% Within Gender	19%	81%	100%			
	% Within Correct Answer	30%	63%	52%			
	% of Total	10%	43%	52%			
	Female N	14	15	29			
	% Within Gender	48%	52%	100%			
	% Within Correct Answer	70%	37%	48%			
	% of Total	23%	25%	48%			
	Total N	20	41	61			
	% Within Gender	33%	67%	100%			
Block 4	Male N	3	29	32	1.46	.23	-.20
Q29	% Within Gender	9%	91%	100%			
	% Within Correct Answer	30%	57%	52%			
	% of Total	5%	48%	52%			
	Female N	7	22	29			
	% Within Gender	24%	76%	100%			
	% Within Correct Answer	70%	43%	48%			
	% of Total	12%	36%	48%			
	Total N	10	51	61			
	% Within Gender	16%	84%	100%			
Q30	Male N	26	6	32	2.16	.14	-.24
	% Within Gender	81%	19%	100%			
	% Within Correct Answer	48%	86%	52%			
	% of Total	43%	10%	52%			
	Female N	28	1	29			
	% Within Gender	97%	3%	100%			
	% Within Correct Answer	52%	14%	48%			
	% of Total	46%	2%	48%			
	Total N	54	7	61			
	% Within Gender	89%	11%	100%			
Q31	Male N	14	18	32	.00	1.00	-.01
	% Within Gender	44%	56%	100%			
	% Within Correct Answer	52%	53%	52%			
	% of Total	23%	30%	52%			
	Female N	13	16	29			
	% Within Gender	45%	55%	100%			
	% Within Correct Answer	48%	47%	48%			
	% of Total	21%	26%	48%			
	Total N	27	34	61			
	% Within Gender	44%	56%	100%			
Q32	Male N	5	27	32	.68	.41	-.15
	% Within Gender	16%	84%	100%			
	% Within Correct Answer	39%	56%	52%			
	% of Total	8%	44%	52%			

Female N	8	21	29
% Within Gender	28%	72%	100%
% Within Correct Answer	62%	44%	48%
% of Total	13%	34%	48%
Total N	13	48	61
% Within Gender	21%	79%	100%

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

To further ascertain whether gender differences occurred on TCAP and on the Task-Success Probability (TSP) a number of independent samples t-tests were computed. Results are presented in Table 9.2.3.

Table 9.2.3: Independent Samples t-tests and Effect Sizes for TCAP and Total TSP and 4 Individual TSP Probes

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Difference	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
TCAP	15.63 (3.37) 32	11.14 (3.10) 29	5.40(59)***	4.49	2.82	6.15	**.33**	**1.39**
Total TSP	3.03 (.83) 32	2.19 (.84) 29	3.95(59)***	.84	.42	1.27	**.21**	**1.01**
TSP1	3.00 (1.14) 32	1.83 (.93) 29	4.39(59)***	1.17	.64	1.71	**.25**	**1.12**
TSP2	2.69 (1.03) 32	1.76 (.91) 29	3.71(59)***	.93	.43	1.43	**.19**	**.96**
TSP 3	3.25 (1.30) 32	2.62 (1.08) 29	2.05(59)*	.63	.01	1.24	.07	.53
TSP 4	3.19 (1.03) 32	2.55 (1.21) 29	2.21(59)*	.64	.06	1.21	.08	.57

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: *d* = Cohen's d. Large effect sizes are in bold.

The results further confirmed that males correctly solved significantly more psychometric problems than females. The observed effect size was very large. Thus, hypothesis 3 was confirmed. Similarly, significant gender differences were observed on the total TSP measure and the four individual probes, with males providing higher task confidence ratings than females. The effect sizes ranged from medium to very large. Hypothesis 4 was also confirmed.

9.2.3.3. TSP, TCAP and Gender as Predictors of DMIQ1 and DMIQ2

Firstly, the relationships between DMIQ1, DMIQ2, gender, TSP and TCAP were investigated. The results are presented in Table 9.2.4. Validating previous findings, a strong inter-correlation was observed between DMIQ1 and DMIQ2, (r = .75, p =.00). Equally, gender correlated strongly and negatively with DMIQ1 (r = -.62, p=.00) and DMIQ2 (r = -.70, p =.00), with females providing lower self-estimates. Gender correlated negatively with TCAP (r = -.58, p =.00). TCAP correlated positively with DMIQ1 (r = .43, p <.01) and DMIQ2 (r = .55, p =.00) as well as with TSP (r = .45, p =.00). A strong negative relationship was observed between gender and TSP (r = -.46, p =.00). TSP also correlated strongly and positively with DMIQ1 (r = .57, p =.00) and DMIQ2 (r = .58, p =.00).

Table 9.2.4: Correlations, Means and Standard Deviations between DMIQ, DMIQ2, Gender, TSP, TCAP and Age

	DMIQ1	DMIQ2	G	TSP	TCAP	A
	108.27	101.01	1.48	2.63	13.49	24.08
	(11.31)	(15.55)	(.50)	(.93)	(3.93)	(8.61)
DMIQ1						
DMIQ2	.75***					
Gender	-.62***	-.70***				
TSP	.57***	.58***	-.46***			
TCAP	.43**	.55***	-.58***	.45***		
Age	-.18	-.17	-.03	-.19	-.07	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). *N = between 57 and 61.*

As in earlier studies, age was included to examine its impact on DMIQ. The age range was 33 years. No significant relationships were observed between age and the remaining variables. Thus, age played no role in the DMIQ type.

In order to ascertain whether gender is the best predictor of DMIQ1 and DMIQ2, two simultaneous hierarchical regressions were conducted.

Results are shown in Table 9.2.5. The dependent variables were DMIQ1 and DMIQ2. Gender, TSP, and TCAP were the independent variables. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity.

The first model was significant F (3,57) = 17.86, p =.00, Adjusted R^2 =.46, f^2=.92), with the overall model explaining 48% of total variance. Gender (β = -.45, p =.00, r_{part} = -.35) and TSP (β = .36, p <.01, r_{part} = .31) were significant predictors of DMIQ1, explaining 12% and 10% of the variance, respectively. In line with the hypothesis, gender was the best predictor of DMIQ1. Hypothesis 5 was confirmed.

Table 9.2.5. Beta coefficients for Simultaneous Multiple Regressions of Gender, TSP and TCAP on DMIQ1 and DMIQ2

Dependent Variable N = 61	***DMIQ1*** β	*t*	***DMIQ2*** β	*t*
Gender	**-.45**	**-3.72***	**-.49**	**-4.53***
TSP	**.36**	**3.27**	**.29**	**2.94**
TCAP	.00	.04	.14	1.33
Regression Model	**F(3, 57) = 17.86***		**F(3 57) = 27.29***	
R^2	.48		.59	
R^2 Change	.48		.59	
Adj. R	.46		.57	
f^2	.92		1.44	

* p <.05, ** p <.01, *** p <.001 *Note:* Significant values are in bold.

The second model was also significant F (3,57) = 27.29, p =.00, Adjusted R^2 = .57, f^2= 1.44), with the overall model explaining 59% of total variance. Gender (β = -.49, p =.00, r_{part} = -.38) and TSP (β = .29, p <.01, r_{part} = .25) were significant predictors of DMIQ2, explaining 14% and 6% of the variance, respectively. Thus, gender was also the best predictor of DMIQ2, confirming hypothesis 6.

Thus, hypotheses 1, 3, 3, 4, 5 and 6 were confirmed.

9.2.4. Discussion

The results of the second experimental condition that included the reduced number of TSP probes were different from Study 14A. Contrary to the results of the study with seven TSP probes, this study that included four TSP probes, confirmed all hypotheses. The existence of HHE on DMIQ1 and DMIQ2 was confirmed as was the decrease in DMIQ estimates from T1 to T2. Equally, male advantage in the psychometric task was observed. Likewise, significant gender differences were observed in task confidence, with higher values reported by males than females. The observed effect sizes varied from medium to very large. In addition, gender was confirmed as the best predictor of DMIQ1 and DMIQ2. Thus, it seems that the reduction in TSP probes had a positive effect on the observed results.

9.3. Study 14C

Gender and TCAP predictors of the Domain-Masculine Intelligence Type

9.3.1. Introduction

The third experimental condition includes no task-success probability probes and is intended to function as a control condition. The overall design and execution of this study is identical to Studies 14A and 14B, with the exception of no task-success probes. Study 14C aims to corroborate whether the exclusion of the TSP probes will impact the observed results, as compared to Studies 14A and 14B.

Thus, HHE is predicted to occur on DMIQ1 and DMIQ2 (H1). Significant decrease in DMIQ estimates is expected to occur from T1 to T2 (H2). Equally, significant gender differences are to be observed in the psychometric task, with males providing more correct answers than females (H3). Finally, as in Studies 14A and 14C, gender is expected as the best predictor of DMIQ1 (H4) and DMIQ2 (H5) over and above TCAP.

9.3.2. Method

Participants

Forty-eight participants took part in this study. There were 24 males (50%) and 24 females. Their age ranged from 17 to 50 (M = 24.02, SD = 8.36) years. 4% achieved GSCE or similar level of education, 58% of participants completed A-levels or similar level of education, 10% achieved non-university higher education, 15% achieved BA/BSc level and 13% achieved MA/MSc or similar level of education.

Measures

Repeated Measure of Domain-masculine Intelligence Type (DMIQ)

See Study 11 (section 7.2.2). Alpha for DMIQ1 was .80 and the inter-item correlation was = .69. For DMIQ2, alpha was .92 and the inter-item correlation was = .85.

Psychometric Aptitude Task

Total Correct Aptitude Problems (TCAP); Numerical and Spatial Psychometric Aptitude Problems (University of Kent, Career Services, 2009; http://www.kent.ac.uk/careers/test.htm)

See Study 14 (section 7.5.2).The alpha for the fifteen numerical reasoning items was .75 and for the nine spatial items .64 (the inter-item correlation was r =.16). Alpha for TCAP was .79.

Procedure

Participants were recruited from the general public through an email campaign by the main researcher and eight second year students, who participated in a mini-research study group in spring 2009 that the main researcher was leading. An email invitation, with an URL link (www.zoomerang.com) to the study and a background explanation of the study was sent to all participants. The snow-balling technique of participant recruitment was used, i.e. participants were asked to forward the study invitation and the URL link to as many acquaintances as possible. In total, 173 individuals logged onto the site during February and May 2009. The data was gathered through an online survey engine www.Zoomerang.com and participation was voluntary. Detailed scoring instructions were given at the beginning of each measure, including timing instructions for the numerical and reasoning problems.

Participants were aware that the study was approved by UCL Ethics Committee, meeting confidentiality and Data Protection requirements. Debrief feedback, correct answers to the numerical and spatial problems, together with the feedback box, were available at the end of the survey. All participants were fluent in English and no language or other problems were reported.

9.3.3. Results

9.3.3.1. HHE and DMIQ1 and DMIQ2

To test hypothesis one, two independent samples t-tests were computed. Results are presented in Table 9.3.1. Significant gender differences, with males providing higher self-estimates than females, were found on DMIQ1 and DMIQ2. Hypothesis 1 was confirmed.

Table 9.3.1: Overview of Independent Samples t-Tests and Effect Sizes for DMIQ1 and DMIQ2

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI L	95% CI U	Effect Size η^2	Effect Size *d*
DMIQ/ DMIQ1	124.88 (14.27) 24	103.69 (13.78) 24	5.23(46)***	21.19	13.04	29.34	**.37**	**1.51**
DMIQ2	120.44 (22.01) 24	97.00 (13.98) 24	4.40(46)***	23.44	12.72	34.15	**.30**	**1.27**

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Hedge's Adjustment *d* is adjusted for sample size and used in both tests. Large effect sizes are in bold.

To test whether a significant decrease in DMIQ estimates occurred from T1 to T2, a paired-samples t-test was conducted. There was a statistically significant decrease in DMIQ from T1 (*M* = 114.28, *SD* = 17.53) to T2 (*M* = 108.72, *SD* = 21.75), t(47) = 4.80, $p = .00$, two-tailed, $r = .94$, $p = .00$, N=48. The mean decrease in DMIQ estimates was 5.56 (*SD* = 8.05) with 95% confidence interval ranging from 3.23 to 7.90. Cohen's *d* statistic (.69) indicated a medium effect size. Hypothesis 2 was confirmed.

9.3.3.2. Gender Differences in Numerical and Reasoning and Spatial Psychometric Aptitude Problems (TCAP)

Table 9.3.2. gives an overview of the 2x2 χ^2 tests and effect sizes for the fifteen numerical reasoning and nine spatial problems. Out of the 24 problems, significant gender differences were observed only on one problem, Q24, where males gave significantly more right answers than did females. Using Cohen's (1988) effect size criteria, medium sized negative significant effect sizes were observed on two problems, i.e. Q20 (φ = -.32, p<.05) and Q24 (φ = -.37, p <.05). Independent samples t-test for TCAP was significant, with males correctly solving more problems than females, t(46) 2.02, p <.05; Mean Differences = 2.42, 95% CI:.01 to 4.83; M_{Male} = 16.67, SD_{Male} = 4.01, M_{Female} = 14.25, SD_{Female} = 4.29; η^2 =.08, Cohen's d =.58. Hypothesis 3 was confirmed.

Table 9.3.2: 2 x 2 Chi Square Tests and Effect Sizes for 15 Numerical, Reasoning and 9 Spatial Problems – Per Gender and % Correct Answer

		Correct Answer Wrong	Correct Answer Right	Total	Yates Continuity Correction Value for 2x2	Asymp. Sig. (2-sided)	Phi (φ) Coefficient
Block 1 Numerical Reasoning Q6	Male N	2	22	24	.00	1.00	-.07
	% Within Gender	8%	92%	100%			
	% Within Correct Answer	40%	51%	50%			
	% of Total	4%	46%	50%			
	Female N	3	21	24			
	% Within Gender	12%	88%	100%			
	% Within Correct Answer	60%	49%	50%			
	% of Total	6%	44%	50%			
	Total N	5	43	48			
	% Within Gender	10%	90%	100%			
Q7	Male N	2	22	24	.00	1.00	-.07
	% Within Gender	8%	92%	100%			
	% Within Correct Answer	40%	51%	50%			
	% of Total	4%	46%	50%			
	Female N	3	21	24			
	% Within Gender	12%	88%	100%			
	% Within Correct Answer	60%	49%	50%			
	% of Total	62%	44%	50%			
	Total N	5	43	48			
	% Within Gender	10%	90%	100%			

Q8	Male N	8	16	24	.00	1.00	-.04
	% Within Gender	33%	67%	100%			
	% Within Correct Answer	47%	52%	50%			
	% of Total	17%	33%	50%			
	Female N	9	15	24			
	% Within Gender	38%	62%	100%			
	% Within Correct Answer	53%	48%	50%			
	% of Total	19%	31%	50%			
	Total N	17	31	48			
	% Within Gender	35%	65%	100%			
Q9	Male N	3	21	24	1.89	.17	-.25
	% Within Gender	12%	88%	100%			
	% Within Correct Answer	27%	57%	50%			
	% of Total	6%	44%	50%			
	Female N	8	16	24			
	% Within Gender	33%	67%	100%			
	% Within Correct Answer	73%	43%	50%			
	% of Total	17%	33%	50%			
	Total N	11	37	48			
	% Within Gender	23%	77%	100%			
Q10	Male N	5	19	24	.11	.74	-.10
	% Within Gender	21%	79%	100%			
	% Within Correct Answer	42%	53%	50%			
	% of Total	11%	40%	50%			
	Female N	7	17	24			
	% Within Gender	29%	71%	100%			
	% Within Correct Answer	58%	47%	50%			
	% of Total	15%	35%	50%			
	Total N	12	36	48			
	% Within Gender	25%	75%	100%			
Q11	Male N	3	21	24	.27	.60	.15
	% Within Gender	12%	88%	100%			
	% Within Correct Answer	75%	48%	50%			
	% of Total	6%	44%	50%			
	Female N	1	23	24			
	% Within Gender	4%	96%	100%			
	% Within Correct Answer	25%	52%	50%			
	% of Total	2%	48%	50%			
	Total N	4	44	48			
	% Within Gender	8%	92%	100%			
Q12	Male N	10	14	24	.08	.77	-.08
	% Within Gender	42%	58%	100%			
	% Within Correct Answer	46%	54%	50%			
	% of Total	21%	29%	50%			
	Female N	12	12	24			
	% Within Gender	50%	50%	100%			
	% Within Correct Answer	55%	46%	50%			
	% of Total	25%	25%	50%			
	Total N	22	26	48			
	% Within Gender	46%	54%	100%			
Q13	Male N	9	15	24	2.08	.15	-.25
	% Within Gender	37%	63%	100%			
	% Within Correct Answer	37%	63%	50%			
	% of Total	19%	31%	50%			
	Female N	15	9	24			
	% Within Gender	63%	37%	100%			
	% Within Correct Answer	63%	37%	50%			
	% of Total	31%	19%	50%			

	Total N	24	24	48			
	% Within Gender	50%	50%	100%			
Q14	Male N	11	13	24	.00	1.00	.04
	% Within Gender	46%	54%	100%			
	% Within Correct Answer	52%	48%	50%			
	% of Total	23%	27%	50%			
	Female N	10	14	24			
	% Within Gender	42%	58%	100%			
	% Within Correct Answer	48%	52%	50%			
	% of Total	21%	29%	50%			
	Total N	21	27	48			
	% Within Gender	44%	56%	100%			
Q15	Male N	6	18	24	.00	1.00	.00
	% Within Gender	25%	75%	100%			
	% Within Correct Answer	50%	50%	50%			
	% of Total	12%	38%	50%			
	Female N	6	18	24			
	% Within Gender	25%	75%	100%			
	% Within Correct Answer	50%	50%	50%			
	% of Total	12%	38%	50%			
	Total N	12	36	48			
	% Within Gender	25%	75%	100%			
Q16	Male N	5	19	24	1.55	.21	-.23
	% Within Gender	21%	79%	100%			
	% Within Correct Answer	33%	58%	50%			
	% of Total	10%	40%	50%			
	Female N	10	14	24			
	% Within Gender	42%	58%	100%			
	% Within Correct Answer	67%	42%	50%			
	% of Total	21%	29%	50%			
	Total N	15	33	48			
	% Within Gender	31%	69%	100%			
Q17	Male N	9	15	24	.34	.56	-.13
	% Within Gender	37%	63%	100%			
	% Within Correct Answer	43%	57%	50%			
	% of Total	19%	31%	50%			
	Female N	12	12	24			
	% Within Gender	50%	50%	100%			
	% Within Correct Answer	57%	44%	50%			
	% of Total	25%	25%	50%			
	Total N	21	27	48			
	% Within Gender	44%	56%	100%			
Q18	Male N	3	21	24	.00	1.00	.07
	% Within Gender	12%	88%	100%			
	% Within Correct Answer	60%	49%	50%			
	% of Total	6%	44%	50%			
	Female N	2	22	24			
	% Within Gender	8%	92%	100%			
	% Within Correct Answer	40%	51%	50%			
	% of Total	4%	46%	50%			
	Total N	5	43	48			
	% Within Gender	10%	90%	100%			
Q19	Male N	11	13	24	1.35	.24	-.21
	% Within Gender	46%	54%	100%			
	% Within Correct Answer	41%	62%	50%			
	% of Total	23%	27%	50%			

	Female N	16	8	24			
	% Within Gender	67%	33%	100%			
	% Within Correct Answer	59%	38%	50%			
	% of Total	33%	17%	50%			
	Total N	27	21	48			
	% Within Gender	56%	44%	100%			
Q20	Male N	13	11	24	3.49	.06	-.32*
	% Within Gender	54%	46%	100%			
	% Within Correct Answer	39%	73%	50%			
	% of Total	27%	23%	50%			
	Female N	20	4	24			
	% Within Gender	83%	17%	100%			
	% Within Correct Answer	61%	27%	50%			
	% of Total	42%	8%	50%			
	Total N	33	15	48			
	% Within Gender	69%	31%	100%			
Block 2	Male N	3	21	24	.15	.70	-.11
Spatial	% Within Gender	12	88%	100%			
Q21	% Within Correct Answer	37%	53%	50%			
	% of Total	6%	44%	50%			
	Female N	5	19	24			
	% Within Gender	21%	79%	100%			
	% Within Correct Answer	62%	48%	50%			
	% of Total	10%	40%	50%			
	Total N	8	40	48			
	% Within Gender	17%	83%	100%			
Q22	Male N	16	8	24	.00	1.00	-.05
	% Within Gender	67%	33%	100%			
	% Within Correct Answer	49%	53%	50%			
	% of Total	33%	17%	50%			
	Female N	17	7	24			
	% Within Gender	71%	29%	100%			
	% Within Correct Answer	51%	47%	50%			
	% of Total	35%	15%	50%			
	Total N	33	15	48			
	% Within Gender	69%	31%	100%			
Q23	Male N	17	7	24	.00	1.00	.00
	% Within Gender	71%	29%	100%			
	% Within Correct Answer	50%	50%	50%			
	% of Total	35%	15%	50%			
	Female N	17	7	24			
	% Within Gender	71%	29%	100%			
	% Within Correct Answer	50%	50%	50%			
	% of Total	35%	15%	50%			
	Total N	34	14	48			
	% Within Gender	71%	29%	100%			
Q24	Male N	3	21	24	4.94	.03*	-.37*
	% Within Gender	12%	88%	100%			
	% Within Correct Answer	21%	62%	50%			
	% of Total	6%	44%	50%			
	Female N	11	13	24			
	% Within Gender	46%	54%	100%			
	% Within Correct Answer	79%	38%	50%			
	% of Total	23%	27%	50%			
	Total N	14	34	48			
	% Within Gender	29%	71%	100%			
Q25	Male N	4	20	24	1.00	.32	-.19
	% Within Gender	17%	83%	100%			

	% Within Correct Answer	33%	56%	50%			
	% of Total	80%	42%	50%			
	Female N	8	16	24			
	% Within Gender	33%	67%	100%			
	% Within Correct Answer	67%	44%	50%			
	% of Total	17%	33%	50%			
	Total N	12	36	48			
	% Within Gender	25%	75%	100%			
Q26	Male N	3	21	24	.15	.70	-.11
	% Within Gender	12%	88%	100%			
	% Within Correct Answer	37%	53%	50%			
	% of Total	6%	44%	50%			
	Female N	5	19	24			
	% Within Gender	21%	79%	100%			
	% Within Correct Answer	62%	48%	50%			
	% of Total	10%	40%	50%			
	Total N	8	40	48			
	% Within Gender	17%	84%	100%			
Q27	Male N	20	4	24	.89	.35	-.21
	% Within Gender	83%	17%	100%			
	% Within Correct Answer	47%	80%	50%			
	% of Total	42%	8%	50%			
	Female N	23	1	24			
	% Within Gender	96%	4%	100%			
	% Within Correct Answer	54%	20%	50%			
	% of Total	48%	2%	50%			
	Total N	43	5	48			
	% Within Gender	90%	10%	100%			
Q28	Male N	6	18	24	.84	.36	-.18
	% Within Gender	25%	75%	100%			
	% Within Correct Answer	38%	56%	50%			
	% of Total	12%	38%	50%			
	Female N	10	14	24			
	% Within Gender	42%	58%	100%			
	% Within Correct Answer	63%	44%	50%			
	% of Total	21%	29%	50%			
	Total N	16	32	48			
	% Within Gender	33%	67%	100%			
Q29	Male N	4	20	24	.00	1.00	.00
	% Within Gender	17%	83%	100%			
	% Within Correct Answer	50%	50%	50%			
	% of Total	8%	42%	50%			
	Female N	4	20	24			
	% Within Gender	17%	83%	100%			
	% Within Correct Answer	50%	50%	50%			
	% of Total	8%	42%	50%			
	Total N	8	40	48			
	% Within Gender	17%	83%	100%			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

9.3.3.3. Gender and TCAP as Predictors of DMIQ1 and DMIQ2

Firstly, the relationships between DMIQ1, DMIQ2, gender and TCAP were explored. Results of the correlational analysis are presented in Table 9.3.3.

Table 9.3.3: Correlations, Means and Standard Deviations between DMIQ1, DMIQ2, Gender, TCAP and Age

	DMIQ1	DMIQ2	G	TCAP	A
	114.28	108.72	1.50	15.56	24.02
	(17.53)	(21.75)	(.51)	(4.28)	(8.36)
DMIQ1					
DMIQ2	.94***				
Gender	-.61***	-.55***			
TCAP	.13	.12	-.29*		
Age	-.01	-.10	-.02	.04	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). *N = between 47 and 48.*

As in previous studies, DMIQ1 and DMIQ2 were strongly intercorrelated (r = .94, p =.00). Similarly, gender correlated strongly and negatively with DMIQ1 (r = -.61, p=.00) and DMIQ2 (r = -.55, p =.00), with females providing lower scores than males. Gender also correlated negatively with TCAP (r = -.29, p <.05).

Age was included to examine whether it plays a role in the DMIQ type. The age range was 33 years. As in Studies 14A and 14B, no significant relationships were observed between age and the remaining variables.

In order to ascertain whether gender is the best predictor of DMIQ1 and DMIQ2, two simultaneous hierarchical regressions were conducted. Results are shown in Table 9.3.4. The dependent variables were DMIQ1 and DMIQ2. Gender and TCAP were the independent variables. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity.

The first model was significant F (2,45) = 13.55, p =.00, Adjusted R^2 =.35, f^2=.61), with the overall model explaining 38% of total variance. Gender (β = -.63, p =.00, r_{part} = -.60) was the only and best predictor of DMIQ1, explaining 36% of variance. Thus, gender was the best and only predictor of DMIQ1. Hypothesis 4 was confirmed.

Table 9.3.4: Beta coefficients for Simultaneous Multiple Regressions of Gender, TCAP on DMIQ1 and DMIQ2

Dependent Variable N = 48	***DMIQ1*** β	*t*	***DMIQ2*** β	*t*
Gender	**-.63**	**-5.10*** **	**-.56**	**-4.26*** **
TCAP	-.05	-.44	-.04	-.29
Regression Model	**F(2, 45) = 13.55*** **		**F(2 45) = 9.55*** **	
R^2	.38		.30	
R^2 Change	.38		.30	
Adj. R	.35		.27	
f^2	.61		.43	

* p <.05, ** p <.01, *** p <.001 *Note:* Significant values are in bold.

The second model was also significant F (2,45) = 9.55, p =.00, Adjusted R^2 = .27, f^2=.43), with the overall model explaining 30% of total variance. Gender (β = -.56, p =.00, r_{part} = -.53) was the only and the best predictor of DMIQ2, explaining 28 % of variance. Hypothesis 5 was confirmed.

Thus, hypotheses 1, 2, 3, 4 and 5 were confirmed.

9.3.4. Discussion

The results of the third experimental or the control condition differed from the results of Study 14A. However, the results of Study 14C were similar to the results of Study 14B in that all hypotheses were confirmed. The existence of HHE on DMIQ1 and DMIQ2 was confirmed as was the decrease in DMIQ estimates from T1 to T2. Equally, gender differences, i.e. male advantage, in the psychometric task, were observed. The observed effect size was medium. In addition, gender was validated as the best predictor of DMIQ1 and DMIQ2 over TCAP. The results provided further evidence for the assertion that reduction or exclusion of the TSP probes had a positive effect on the observed results.

9.4. Psychometric Problems and TSP Probes Used in Chapter 7

Study 11

Numerical and Reasoning Problems (Bryon, 2006)

1.

? x 12 = 132 ____ 15.02 ÷1,000 = *?* __ 1,200 x *?* =____

Using the scale, indicate how likely you are to succeed on the same task, but with increased difficulty:

1 2 3 4 5

Very unlikely Very likely

****

2.

381 355 329 303 277 *?* ____ 3:4 6:8 9:12 12:16 *?* ____ 67 24 8 2 9 *?* 24 18 3 ____

Using the scale, indicate how likely you are to succeed on the same task, but with increased difficulty:

1 2 3 4 5

Very unlikely Very likely

****

3.

After a discount of 8% a computer is advertised for sale at £552; what was the original price of the computer?

a) £550 b) 600 c) 654 d) 656 _________

On a street map, ¾ of a centimetre represents one kilometre. What distance, in kilometres, is represented by 1 ¾ centimetres?

a) 1 ½ b) 2 c) 2 ⅓ d) 2 ½ e) 2 ⅝ _________

A box contains two coins. One coin is heads on both sides and the other is heads on one side and tails on the other. One coin is selected from the box at random and the face of one side is observed. If the face is heads what is the per cent change that the other side is heads?

a) 25% b) 33% c) 50% d) 66% e) 88% _________

Using the scale, indicate how likely you are to succeed on the same task, but with increased difficulty:

1 2 3 4 5

Very unlikely Very likely

x ***

4.

DIFFERENT is to CORRESPONDING as SUPERIOR is to __________

a) elder b) junior c) manager

“Mip mop mup” means “you are ready”.
“Map mip mep” means “better be ready”.
“Myp map mop” means “tourists are better”.
What words would you use to say: “ Better be tourists?” The order that you place the words in is unimportant – you only need to find the correct words to use. ______

If all Gannucks are Dorks and most Gannuks are Xorgs, the statement that some Dorks are Xorgs is:

a) True b) False c) Indeterminable from data

Using the scale, indicate how likely you are to succeed on the same task, but with increased difficulty:

1 2 3 4 5

Very unlikely Very likely

****

○●●	●○●	●●○	?	○●●	●●○	○●●	○●●
●○●	●●○	○●●		●○●	○●○	○○●	○●○
●●○	○●●	○●●		●●○	●●●	●●○	○○●
				A	B	C	D

213	134	729	?	497	137	243	246
358	628	516		968	685	378	178
246	336	235		751	362	266	369
				A	B	C	D

■□□	□■□	□□■		□□■	■□□	■□□	■□□
□■□	□□■	■□□		□■□	□□■	□■□	□■□
■□□	□■□	□□■		□□■	□■□	■□□	□□■
				A	B	C	D

Using the scale, indicate how likely you are to succeed on the same task, but with increased difficulty:
1 2 3 4 5
Very unlikely ... Very likely

Study 12

Numerical and Reasoning Problems (Bryon, 2006)

1.
? x 12 = 132 ____ 15.02 ÷1,000 = *?* __ 1,200 x *?* =____

Using the scale, indicate how likely you are to succeed on the same task, but with increased difficulty:
1 2 3 4 5
Very unlikely ... Very likely

****

2.
381 355 329 303 277 *?* ____ 3:4 6:8 9:12 12:16 *?* ____

Using the scale, indicate how likely you are to succeed on the same task, but with increased difficulty:
1 2 3 4 5
Very unlikely ... Very likely

****

3.
After a discount of 8% a computer is advertised for sale at £552; what was the original price of the computer?
a) £550 b) 600 c) 654 d) 656 __________

On a street map, ¾ of a centimetre represents one kilometre. What distance, in kilometres, is represented by 1 ¾ centimetres?
a) 1 ½ b) 2 c) 2 ⅓ d) 2 ½ e) 2 ⅝ __________

Using the scale, indicate how likely you are to succeed on the same task, but with increased difficulty:
1 2 3 4 5
Very unlikely ... Very likely

x

****

4.
DIFFERENT is to CORRESPONDING as SUPERIOR is to __________
a) elder b) junior c) manager

"Mip mop mup" means "you are ready".
"Map mip mep" means "better be ready".
"Myp map mop" means "tourists are better".
What words would you use to say: " Better be tourists?" The order that you place the words in is unimportant – you only need to find the correct words to use. ______

If all Gannucks are Dorks and most Gannuks are Xorgs, the statement that some Dorks are Xorgs is:
a) True b) False c) Indeterminable from data

Using the scale, indicate how likely you are to succeed on the same task, but with increased difficulty:
1 2 3 4 5
Very unlikely Very likely

****

○●●	●○●	●●○	?	○●●	●●○	○●●	○●●
●○●	●●○	○●●		●○●	○●○	○○●	○●○
●●○	○●●	○●●		●●○	●●●	●●○	○○●
				A	B	C	D

213	134	729	?	497	137	243	246
358	628	516		968	685	378	178
246	336	235		751	362	266	369
				A	B	C	D

■□□	□■□	□□■		□□■	■□□	■□□	■□□
□■□	□□■	■□□		□■□	□□■	□■□	□■□
■□□	□■□	□□■		□□■	□■□	■□□	□□■
				A	B	C	D

Using the scale, indicate how likely you are to succeed on the same task, but with increased difficulty:
1 2 3 4 5
Very unlikely Very likely

Study 13

Numerical and Reasoning Problems (Bryon, 2006)

The following 6 items are from a well-known IQ test. The maximum allocated time for these 6 problems is 3 minutes (180 seconds). Use your watch to time your performance and do NOT, under any circumstances go over the allocated time or your results will be excluded from the survey! Leave any unfinished problems blank.

12. Please complete the missing part (?):
1,200 x ? = 48,000

13. On a street map, ¾ of a centimetre represents one kilometre. What distance, in kilometres, is represented by 1 ¾ centimetres?
   a) 1 ½
   b) 2
   c) 2 1/3
   d) 2 ½
   e) 2 5/8

14. After a discount of 8% a computer is advertised for sale at £552; what was the original price of the computer?
   a) £550
   b) £654

c) £600
d) £656

16. DIFFERENT is to CORRESPONDING as SUPERIOR is to ?
a) elder
b) junior
c) manager

17. Mip mop mup" means "you are ready" "Map mip mep" means "better be ready" "Myp map mop" means "tourists are better" What words would you use to say: " Better be tourists?"

18. If all Gannucks are Dorks and most Gannuks are Xorgs, the statement that some Dorks are Xorgs is:
a) true
b) false
c) indeterminable from data

Crystallised Intelligence (GKT, Irwing, Cammock, & Lynn, 2001)

Please answer the following 10 questions. You have maximum 2 minutes (120 seconds) to write your answers in the designated fields. If you go over the allowed time, your answers will be excluded.

20. What is the longest river in Europe?
21. What is the capital of Mongolia?
22. What is the hardest substance known to man?
23. Who composed the Goldberg variations?
24. What metal is liquid at normal room temperature?
25. Who wrote the novel Anna Karenina?
26. Which American president was assassinated in 1865?
27. What is the largest planet in the solar system?
28. Who directed the movie Saving Private Ryan?
29. What is the largest mammal?

This is the end. Thank you so much for your participation!
Correct answers:
12) 40
13) 2 1/3
14) £600
16) Junior
17) Map mip myp
18) True
20) Volga
21) Ulaanbaater; Ulanbater
22) Diamond
23) J. S. Bach
24) Mercury
25) L. Tolstoy
26) A. Lincoln
27) Jupiter
28) S. Spielberg
29) (The Blue) Whale

Studies 14A, 14B and14C

TCAP (Numerical and Spatial Aptitude Problems, University of Kent, Careers Services (2009)

Here you will be asked to solve ***15 numerical tasks****. The total allowed time for this section is* ***8 minutes.*** *The use of calculators is NOT permitted during this test. Do NOT go over the time limit - if you do, your answers will be automatically excluded from the survey results. Leave any answered questions blank. If you want to check how well you did, write your answers down. The correct answers will be given at the end of this survey. Thank you.*

6. Please complete the missing number by selecting one answer: **83 - 17 = 56 + ?**
a) 6
b) 10
c) 16
d) 20
e) 30

7. Please complete the missing number by selecting one answer: **56 / 7 = ? - 5**
a) 11
b) 13
c) 14
d) 15
e) 16

8. Please complete the missing number by selecting one answer: **20/0.8 = ?**
a) 14
b) 15
c) 16
d) 24
e) 25

9. Indicate how likely are you to succeed on the same task, but with increased difficulty:
a) Very unlikely
b) Unlikely
c) Neither unlikely or likely/ Do not know
d) Likely
e) Very likely

10. Which is the largest fraction? **3/4 7/8 4/5 7/9 7/10**
a) 3/4
b) 7/8
c) 4/5
d) 7/9
e) 7/10

11. If oranges cost 5 for 75p, how many can you buy for £2.70 (assuming they can be bought per piece)?
a) 15
b) 16
c) 17
d) 18
e) 19

12. You are paid £250 per week. You get an increase of 4%, plus an extra £5.00 per week. What will your new weekly pay be?
a) £260
b) £265
c) £270
d) £275
e) £280

13. Indicate how likely are you to succeed on the same task, but with increased difficulty:

a) Very unlikely
b) Unlikely
c) Neither unlikely or likely/ Do not know
d) Likely
e) Very likely

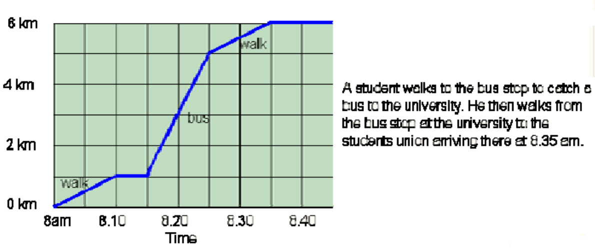
14. A car left Canterbury at 7:12am and arrived in Birmingham, 180 miles distant, at 10:57am. What was the average speed in miles per hour?
a) 42
b) 44
c) 46
d) 48
e) 50

15. Carla driver drivers 8 km South then 6 km West and 2 km South again. She then drives 3 km east to avoid a traffic jam before driving 6 km North. How many kilometres is she from her starting point?
a) 4
b) 5
c) 6
d) 7
e) 8

16. An aircraft flies 930 miles in 75 minutes. How many miles does it fly in 4 hours 45 minutes, assuming a constant speed?
a) 3112
b) 3477
c) 3512
d) 3522
e) 3534

17. Indicate how likely are you to succeed on the same task, but with increased difficulty:
a) Very unlikely
b) Unlikely
c) Neither unlikely or likely/ Do not know
d) Likely
e) Very likely

Student walks to the bus stop to catch a bus to the university. He then walks from the bus stop at the university to the students' union, arriving there at 8:35am.



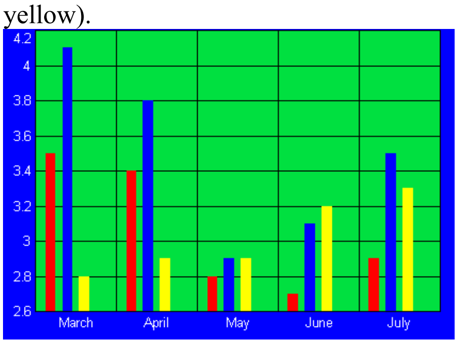
18. How far does the student walk in total?
a) 1 km
b) 2 km
c) 3km
4) 4 km
e) 5 km

19. How far is he from the university students' union t 8:20am?
a) 1 km
b) 2 km

c) 3 km
d) 4 km
e) 5 km

20. What is the average speed of the bus?
a) 14 km/h
b) 24 km/h
c) 32 km/h
d) 40 km/h
e) 48 km/h

The graph to the left gives the number of computers sold each month (in thousands) by three different computer manufactures. Manufacturer 1 (in red), Manufacturer 2 (in blue), and Manufacturer 3 (in yellow).



The graph to the left gives the number of computers sold each month (in thousands) by three different computer manufacturers. Manufacturer 1 (in red), Manufacturer 2 (in blue) and Manufacturer 3 (in yellow).

21. Which month showed the largest total decrease in PC sales over the previous months?
a) March
b) April
c) May
d) June
e) July

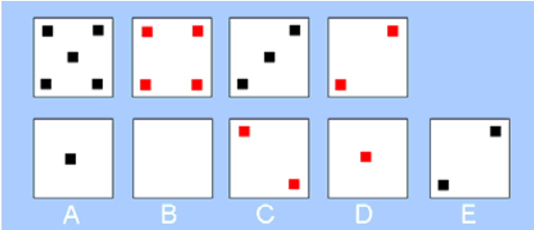
22. What % of 2nd manufacturer's sales were made in April (to the nearest %)?
a) 16
b) 22
c) 27
d) 27
e) 33

23. If the average profit made on each PC sold by manufacturer 3 over all 5 months was £78, what was the total profit on all sales in this period by that manufacturer?
a) £650,000
b) £820, 000
c) £1,095, 600
d) £1,777,800

24. Indicate how likely are you to succeed on the same task, but with increased difficulty:
a) Very unlikely
b) Unlikely
c) Neither unlikely or likely/ Do not know
d) Likely
e) Very likely

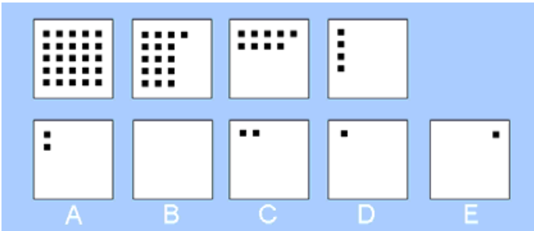
*Here you will be asked to solve **9 spatial tasks**. The total allowed time for this section is **5 minutes**. Do NOT go over the time limit - if you do, your answers will be automatically excluded from the survey results. Leave unanswered questions blank. If you want to check how well you did, write your answers down. The correct answers will be given at the end of* *this survey. Thank you.*

25. Which letter provides the most logical solution?



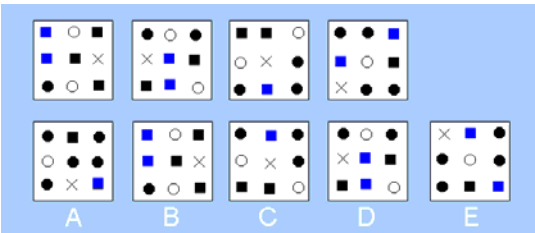
a) A
b) B
c) C
d) D
e) E

26. Select the most logical solution



a) A
b) B
c) C
d) D
e) E

27. Complete the sequence by choosing the most logical solution

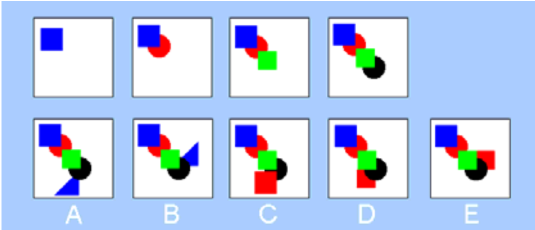


a) A
b) B
c) C
d) D
e) E

28. Indicate how likely are you to succeed on the same task, but with increased difficulty:
a) Very unlikely
b) Unlikely
c) Neither unlikely or likely/ Do not know

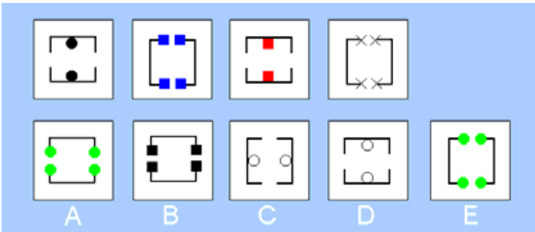
d) Likely
e) Very likely

29. Which picture/letter completes the sequence?



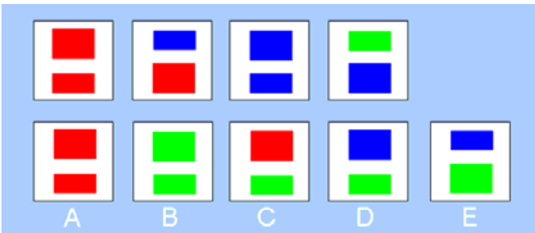
a) A
b) B
c) C
d) D
e) E

30. Which picture completes the sequence?



a) A
b) B
c) C
d) D
e) E

31. Choose the most logical picture to complete the sequence

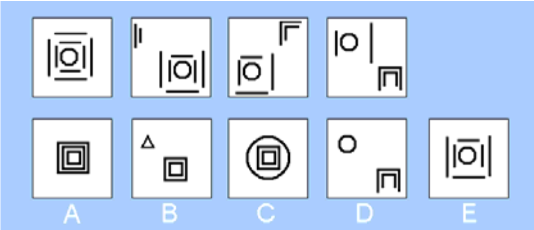


a) A
b) B
c) C
d) D
e) E

32. Indicate how likely are you to succeed on the same task, but with increased difficulty:
a) Very unlikely
b) Unlikely
c) Neither unlikely or likely/ Do not know

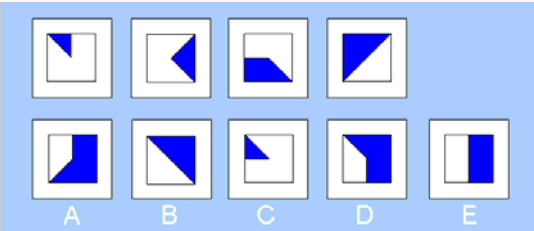
d) Likely
e) Very likely

33. Complete the sequence



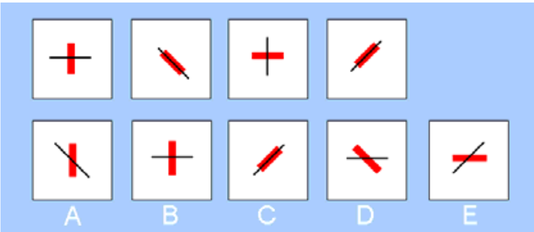
a) A
b) B
c) C
d) D
e) E

34. Complete the sequence by choosing the most logical sequence



a) A
b) B
c) C
d) D
e) E

35. Complete the sequence



a) A
b) B
c) C
d) D
e) E

36. Indicate how likely are you to succeed on the same task, but with increased difficulty:
a) Very unlikely
b) Unlikely
c) Neither unlikely or likely/ Do not know

d) Likely
e) Very likely

The correct answers are as follows:

Q6: B = 10, Q7: B = 13, Q8: E = 25, Q10: B = 7/8, Q11: D = 18, Q12: B = £265, Q14: D = 48, Q15: B = 5, Q16: E = 3534, Q18: B = 2km, Q19: C = 3 km, Q20: B = 24 km/h, Q21: C = May, Q22: B = 22%, Q23: E = 1,777,800, Q25: A, Q26: D, Q27: A, Q29: C, Q30: D, Q31: B, Q33: C, Q34: D, Q35: B

Study 15

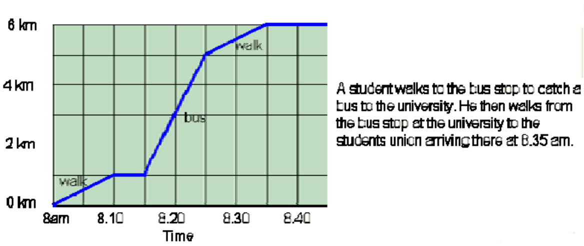
*You are now asked to solve **5 tasks**. You will have maximum. **5 minutes** for this task. **Do NOT go over** the time limit or you will be automatically disqualified.. You can **NOT** use of calculators or any other tools/help. Leave the unanswered questions blank. If you want to know how well you did, write down your answers. The correct answers are given at the end of this survey!*

8. Please complete the missing number by selecting one answer:
56 / 7 = ? - 5
  a) 11
  b) 13
  c) 14
  d) 15
  e) 16

9 Which is the largest fraction? 3/4   7/8   4/5   7/9   7/10
  a) ¾
  b) 7/8
  c) 4/5
  d) 7/9
  e) 7/10

10. Carla driver drivers 8 km South then 6 km West and 2 km South again. Shen then drives 3 km east to avoid a traffic jam before driving 6 km North. How many kilometers is she from her starting point?
  a) 4
  b) 5
  c) 6
  d) 7
  e) 8

Student walks to the bus stop to catch a bus to the university. He then walks from the bus stop at the university to the students' union, arriving there at 8:35am.



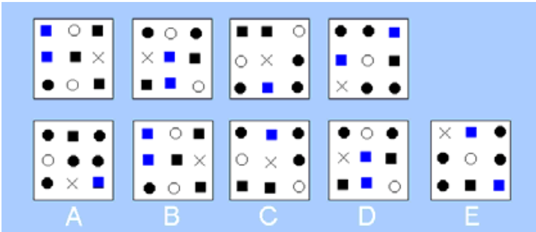
11. How far does the student walk in total?

a) 1 km
b) 2 km
c) 3 km
d) 4 km
e) 5 km

12. What is the average speed of the bus?
a) 14 km p/h
b) 24 km p/h
c) 32 km p/h
d) 40 km p/h
e) 48 km p/h

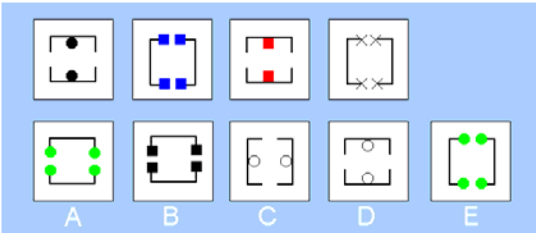
*Here you will be asked to solve **5 spatial tasks**. You will have maximum **4 minutes** to complete this task.. **Do NOT go over** the time limit or you will be automatically disqualified. Leave any unanswered questions blank. If you want to check how well you did, write your answers down as the correct answers will be given at the end.*

14. Complete the sequence by choosing the most logical solution



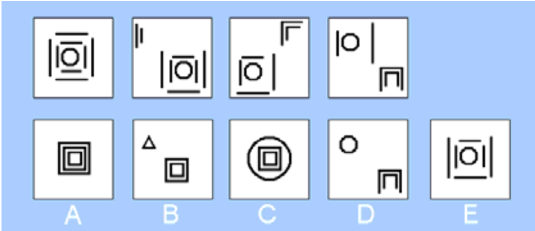
A
B
C
D
E

15. Which picture completes the sequence?



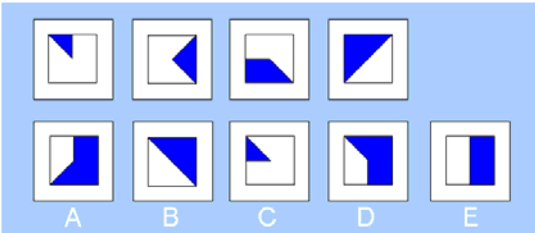
A
B
C
D
E

16. Complete the sequence



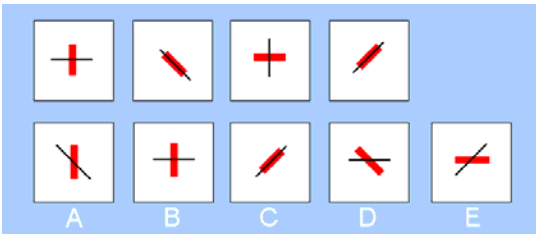
A
B
C
D
E

17. Complete the sequence by choosing the most logical sequence



A
B
C
D
E

18. Complete the sequence



A
B
C
D
E

*Now answer **5 questions from one of the most known & used IQ test.** The total allowed time for this section is **2 minutes**. **Do NOT go over** the time limit or your answers will be automatically excluded. Leave any unanswered questions blank. As with the preceding two sections, answers will be given at the end, so note them down if you want to know how well you did!*

20. What is the longest river in Europe?
21. Who composed the Goldberg Variations?
22. What metal is liquid at room temperature?
23. Who wrote the novel Anna Karenina?

24. What is the largest planet in the Solar System?

Correct answers: Q8: B = 13, Q9: B = 7/8, Q10: B = 5, Q11: B = 2km, Q12: B = 24 km/h, Q14: A, Q15: D, Q16: C, Q17: D, Q18: B, Q20: Volga, Q21: J. S. Bach, Q22: Mercury, Q23: L. Tolstoy, Q24: Jupiter

9.5. Study 16

Overview of the Combined Results of All Studies

9.5.1. Introduction

The fifteen datasets were combined to create one large database with 2292 subjects. The following variables that were present in all fifteen studies were used: gender, age, and domain-masculine intelligence type estimates. Furthermore, DMIQ1 estimates were available for 2137 subjects and DMIQ2 estimates for 694 subjects and were taken from the experimental studies. In addition, Task-Success Probability (TSP) for 670 subjects and Total Correct Aptitude Problems (TCAP) entries for 970 subjects were also extracted from the five experimental studies and included. In order to validate the main objectives of this thesis as well as corroborate the previous findings, the following hypotheses were tested:

HHE will be observed on DMIQ, i.e. on DMIQ1 and DMIQ2, when applicable (H1). A significant decrease in DMIQ from T1 to T2 is expected to occur (H2). Gender is expected to be the best predictor of DMIQ, i.e. DMIQ1 and DMIQ2 when applicable, over and above age, TSP and TCAP (H3). Age is not expected to correlate with DMIQ, i.e. DMIQ1 and DMIQ2 (H4). Gender differences are expected in TSP, with males being more confident about their abilities than females (H5). Gender differences are also predicted in TCAP, with males successfully resolving more problems than females (H6). Gender is expected to moderate the relationship between TSP and DMIQ1 (H7) and DMIQ2 (H8). Gender is also expected to moderate the relationship between TCAP and DMIQ1 (H9) and DMIQ2 (H10).

9.5.2. Method

Participants

The fifteen databases were combined, totalling 2292 participants. There were 1380 (60%) females and 912 males. Their age ranged from 17 to 80 (M = 26.35, SD = 12.04) years.

Measures

Repeated Measure of Domain-masculine Intelligence Type (DMIQ)

See Study 11 (section 7.2.2).

Psychometric Aptitude Task

Total Correct Aptitude Problems (TCAP); Numerical Reasoning, Spatial and Crystallised Intelligence Aptitude Problems (University of Kent, Career Services, 2009; http://www.kent.ac.uk/careers/test.htm)

The psychometric aptitude problem scores from the five online experimental design studies were combined, ranging from 13 to 24 problems per study. All five studies had numerical reasoning and spatial problems and two studies also included general knowledge (Gc) problems. The numerical reasoning and spatial problems that were in public domain for online psychometric training purposes by the Career Services of University of Kent were adopted and used. The crystallised intelligence (Gc) questions were adopted from the General Knowledge Test (Irwing, Cammock and Lynn, 2001), covering general knowledge, science, literature, geography. Total number of correctly solved aptitude problems, or Total TCAP score was computed per individual.

Task Success Probability (TSP)

Task Success Probability Estimation Measure (Storek, 2007).

See Study 11 (section 7.2.2.)

9.5.3. Results

9.5.3.1. Domain-Masculine Intelligence Type and the Hubris and Humility Effect in T1 and T2

Two independent t-tests were computed to assess whether significant gender differences in DMIQ were observed in the pre- and post-task estimation conditions, with medium and very large effect sizes. Results revealed significant gender differences in DMIQ/DMIQ1 and DMIQ2, with males providing higher domain-masculine intelligence type estimates than did females. Results are presented in Table 9.5.1. This finding provided further support for the existence of HHE on DMIQ type. Hypothesis 1 was confirmed.

Table 9.5.1: Overview of Independent Samples t-Tests and Effect Sizes for DMIQ1 and DMIQ2

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI L	U	Effect Size η^2	*d*
DMIQ/ DMIQ1	120.96 (17.19) 869	107.83 (16.04) 1268	17.83(1779)***	13.14	11.69	14.58	.13	.79
DMIQ2	115.61 (16.70) 318	98.77 (17.18) 376	13.03(692)***	16.84	14.31	19.38	**.20**	**.99**

* *p* < .05 ** *p* < .01 *** *p* < .001 (2-tailed). Note: *d* = Hedge's Adjustment *d* is Cohen's *d* adjusted for sample size. Large effect sizes are in bold.

To test whether a significant decrease occurred in DMIQ from pre- to post-task estimation condition, a paired-samples t-test was conducted. There was a statistically significant decrease in the domain-masculine intelligence type self-estimates from T1 (*M* = 112.14, *SD* = 16.94) to T2 (*M* = 106.54, *SD* = 18.85), t(689) = 15.34, *p* = .00, two-tailed, *r* = .86, *p* =.00, N=690. The mean decrease in domain-masculine intelligence self-estimates was 5.60 (*SD* = 9.59) with 95% confidence interval ranging from 4.88 to 6.31. Cohen's *d* (.58) indicated a medium effect size. Hypothesis 2 was confirmed.

9.5.3.2. Gender as the best Predictor of DMIQ/DMIQ1 and DMIQ2

Firstly, the relationships between DMIQ/DMIQ1 and DMIQ2, gender, TSP, TCAP and age were explored. Table 9.5.2. shows the results of the correlational and partial correlational analyses.

Table 9.5.2: Correlations and Partial Correlations, Means and Standard Deviations between DMIQ/DMIQ1, DMIQ2, Gender, TSP, TCAP and Age

	DMIQ/1	DMIQ2	G	TSP	TCAP	A
	113.17	106.49	1.60	3.21	7.33	26.35
	(17.73)	(18.92)	(.49)	(2.14)	(5.42)	(12.04)
DMIQ/DMIQ1						
DMIQ2	.86***					
Gender	-.36***	-.44***				
TSP	.10*	.17***	-.11**			
TCAP	.16***	.26***	-.22***	-.11**		
Age	.34***	-.07	-.12***	.29***	.13***	
Controlled for Age						
DMIQ/DMIQ1						
DMIQ2	. 94***					
Gender	-.35***	-.46***				
TSP	.00	.20***	-.08*			
Total Correct Aptitude Prbs	.13***	.27***	-.21***	-.15***		

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).   *N between 654 and 2287.*

As in the individual studies, DMIQ1 and DMIQ2 were strongly intercorrelated (r = .86, p =.00). In accordance with the previous findings, gender correlated negatively with DMIQ/DMIQ1 (r = -.36, p=.00) and DMIQ2 (r = -.44, p =.00), with females providing lower scores than males on both occasions. As in previous studies, positive correlations were observed between TSP and DMIQ/DMIQ1 (r = .10, p <.05) and DMIQ2 (r =.17, p =.00). Positive correlations were also observed between TCAP and DMIQ2 (r =.16, p=.00) and TCAP and TSP (r =.26, p =.00). Negative correlations were observed between TSP and gender (r =-.11, p <.01), TCAP and gender (r =-.22, p =.00) as well as TSP and TCAP (r =-.11, p <.01).

Next, the impact of gender on DMIQ/DMIQ1 and DMIQ2, gender, TSP, TCAP and age was explored. The data was split per gender and the correlations recomputed. Results are presented in Table 9.5.3. There were notable differences between male and female results. Particularly, the relationships between TSP and the remaining variables and TCAP and the remaining variables revealed divergent correlational patterns for each gender. It is likely that the varying number of probes and problems from the five experimental studies influenced the newly computed Total TSP and Total TCAP variables as well as the individualised scores.

For males, no significant relationship was observed between TSP and DMIQ1 (r = .00, p =.95) or DMIQ T2 (r =.03, p =.60). However, in the female subsample, TSP correlated with DMIQ1 (r = .12, p <.05) and DMIQ2 (r=.23, p = .00), although the correlations were small to medium sized. For males, small positive relationships were observed between TCAP and DMIQ1 (r =.13, p <.05) and DMIQ2 (r =.18, p <.01) and a negative relationship between TCAP and TSP (r = -.18, p <.01). The first two relationships are in line with the findings of previous studies. Yet, for females, the relationship between TCAP and DMIQ1 as well as DMIQ2 was not significant and a small positive correlation was observed between TCAP and DMIQ2 (r =.21, p =.00).

Table 9.5.3: Correlations, Means and Standard Deviations between DMIQ/DMIQ1, DMIQ2, Gender, TSP, TCAP and Age – Per Gender

Males	DMIQ1	DMIQ2	TSP	TCAP	A
n between 306 and 909	120.96	115.61	3.46	8.75	28.06
	(17.19)	(16.70)	(2.19)	(5.37)	(12.99)
DMIQ/DMIQ					
DMIQ2	.86***				
TSP	.00	.03			
TCAP	.13*	.18**	-.18**		
Age	.32***	-.12*	.28***	.05	
Females	DMIQ1	DMIQ2	TSP	TCAP	A
n between 364 and 1378	107.83	98.77	3.00	6.32	25.22
	(16.04)	(17.18)	(2.08)	(5.23)	(11.23)
DMIQ/DMIQ1					
DMIQ2	.80***				
TSP	.12*	.23***			
TCAP	.07	.21***	-.08		
Age	.33***	-.10	.29***	.15***	

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).

As in earlier studies, age was included to examine its role in DMIQ. The age range of participants was 63 years. Significant relationships were observed between age and the remaining variables, with the exception of DMIQ2. In order to further investigate age's role, the correlation analysis was re-computed and age was partialled out. The results are presented in Tables 9.5.2 and 9.5.3. Preliminary analyses were performed to ensure no violation of the assumptions of normality, linearity and homoscedasticity. The inspection of the zero order correlations suggested that controlling for age had a limited impact on the strength of the observed relationships, with the exception of the correlation between TSP and DMIQ/DMIQ1, which ceased to be significant (r =.00, p =.99), suggesting that age impacted on the relationship between the intelligence type and the task-success probes. An independent samples t-test for age was significant, t(1744)5.40, p<.00; $M_{Males} = 28.06$, $SD_{Males} = 12.99$,

$M_{Females}$ = 25.22, $SD_{Females}$ = 11.23; Mean Difference = 2.85, 95% CI 1.81 to 3.88; η^2 =.02, Hedge's Adjustment *d* =.23.

Equally, the gender-specific correlational results revealed that for both males and females, age correlated positively with DMIQ1 (r =.32, p =.00) and (r =.33, p =.00), respectively. For DMIQ2, only one significant relationship was observed between age and gender ($r = -.12, p<.05$) in the male subsample. The results imply that older participants of both genders provided higher DMIQ/DMIQ1 estimates but only younger males provided higher DMIQ2 estimates.

In addition, for males and females, a positive significant relationship was observed between age and TSP (r =.28, p =.00) and (r =.29, p =.00), revealing that older subjects of both genders had higher task confidence. For TCAP, a positive relationship was observed with age but only for females, (r = .15, p=.00), indicating that older female participants were more successful in solving the psychometric problems. These findings are in line with the existing literature on gender and sex differences in cognitive abilities. Hypothesis 4 was not confirmed.

In order to investigate whether gender was the best predictor of DMIQ/DMIQ1 and DMIQ2, two simultaneous multiple regressions were performed. The dependent variables were DMIQ/DMIQ1 and DMIQ2, and TSP, TCAP, age and gender were the independent variables. Results are reported in Table 9.5.4. Preliminary analyses were conducted to ensure no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. As Mahalanobis distance values were violated in both analyses, more stringent criteria (20.52) were set, the analyses recomputed, with the number of participants considerably reduced.

The first model that used DMIQ/DMIQ1 as a dependent variable was significant ($F(4,653) = 41.18$, p =.00, Adjusted R^2 =.20, f^2=.25), with the overall model explaining 20% of total variance. All four variables were significant predictors of the intelligence type. TSP (β = .11, p <.01, r_{part} = .10), TCAP (β = .21, p =.00, r_{part} = .20), age (β = -.10, p <.01, r_{part} = -.09) and gender (β = -.34, p =.00, r_{part} = -.33), accounting for 1%, 4%, 1%, and 11% of variance respectively. Gender was the best predictor of DMIQ/DMIQ1, followed by TCAP, TSP and age.

The second model was also significant (F(4,646) = 58.17, p =.00, Adjusted R^2 = .26, f^2=.37), with the overall model explaining 27% of total variance. Again, all four variables were significant predictors of DMIQ2. TSP (β = .19, p =.00, r_{part} = .18), TCAP (β = .21, p =.00, r_{part} = .21), age (β = -.16, p =.00, r_{part} = -.15) and gender (β = -.38, p =.00, r_{part} = -.37), accounting for 3%, 5%, 2%, and 14% of variance respectively. Gender was the best predictor of DMIQ2, followed by TCAP, TSP and age. Hypothesis 3 was confirmed.

Table 9.5.4: Beta coefficients for Simultaneous Multiple Regressions of TSP, TCAP, Age and Gender onto DMIQ/DMIQ1 and DMIQ2

Dependent Variable N = 664	***DMIQ1*** β	*t*	***DMIQ2*** β	*t*
TSP	**.11**	2.88**	**.19**	5.33***
TCAP	**.21**	5.76***	**.21**	6.14***
Age	**-.10**	-2.60**	**-.16**	-4.48***
Gender	**-.34**	-9.54***	**-.38**	-10.95***
Regression Model	**F(4, 653) = 41.18*** **		**F(4, 646) = 58.17*** **	
R^2	.20		.27	
R^2 Change	.20		.27	
Adj. R^2	.20		.26	
f^2	.25		.37	

p <.05, ** *p* <.01, *** *p* <.001 *Note:* Significant values are in bold.

9.5.3.3. Gender Differences in the Task Success Probability Estimation (TSP) and the Psychometric Aptitude Problems (TCAP)

To test hypotheses 5 and 6 two independent samples t-tests were computed. Results are presented in Table 9.5.5. The test for TSP revealed significant gender differences between males and females; with males being more confident than females about their ability to successfully solve similar, yet more difficult psychometric aptitude tasks. The independent samples t-test for TCAP was also significant, with males correctly solving significantly more psychometric problems than did females. Thus, hypotheses 5 and 6 were confirmed.

Table 9.5.5: Overview of Independent Samples t-Tests and Effect Sizes for TSP and TCAP

	Males M (SD) n	Females M (SD) n	*t(df)*	Mean Diff.	95% CI		Effect Size	
					L	U	η^2	*d*
TSP	3.46 (2.19) 306	3.00 (2.08) 364	2.78(668)**	.46	.14	.78	.01	.22
TCAP	8.75 (5.37) 397	6.32 (5.23) 564	7.01(959)***	2.43	1.75	3.11	.07	.46

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed). Note: *d* = Hedge's Adjustment *d* is adjusted for sample size and used in both tests.

9.5.3.4. Impact of Gender on the Relationship between TSP on DMIQ/DMIQ1 and DMIQ2

TSP was collapsed into a categorical variable, Group 1 containing individuals who provided the lowest task-success estimates, Group 2 individuals that provided average task-success estimates and Group 3 individuals who were the most confident about their ability to succeed. Results are presented in Table 9.5.6.

Table 9.5.6: Overview of TSP Banded

	TSP	N
Group 1	<=3	238
Group 2	3-4	235
Group 3	4+	197

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender affects the relationship between TSP and DMIQ/DMIQ1 and DMIQ2. Results are presented in Table 9.5.7.

For DMIQ/DMIQ1, homogeneity of variance assumption was violated (Levene Statistic $p <.05$), indicating the groups variances were not equal. An alternative check for comparing variances was used. Firstly, the largest and the smallest standard deviations were squared. The largest squared SD was divided by the smallest squared SD, with resulting value of 1.26, which is smaller than the recommended value of 2, suggesting that the group variances, albeit not equal, were tolerable. Subsequently, the significance level was adjusted to $p <.01$.

The interaction effect between TSP and gender estimation conditions was not significant, $F(2,658) = .14$, $p = .87$, $\eta_p^2 = .00$. There was a statistically significant main effect for TSP, $F(2,658) = 21.35$, $p =.00$, $\eta_p^2 = .06$, with medium effect size. The main effect for gender, $F(1,658) = 87.48$, $p =.00$, $\eta_p^2 = .12$, was significant, with medium effect size.

Planned contrasts revealed significant differences between Group 1 and Group 2 (Contrast Estimate -7.54, $p =.00$) and between Group 2 and Group 3 (Contrast Estimate -4.94, $p <.01$). Post-hoc comparisons using the Games-Howell test indicated that the mean scores for Group 1 (<=3) were significantly different from Group 2 (3-4 as well as from Group 3 (4+). Group 2 mean scores also significantly differed from Group 3 mean scores.

Table 9.5.7: Two 2-way ANOVAs (TSP and gender) on DMIQ/DMIQ1 and DMIQ2

Variable	TSP Groups	Mean Score (SD)			*F*-score		
		Total	Males	Females	TSP	Gender	TSP x Gender
DMIQ1	G1 (L)	105.74 (16.33)	113.48 (17.43)	101.84 (14.29)	21.35***	87.48***	.14
	G2 (M)	112.29 (15.98)	118.77 (16.42)	106.68 (13.30)			
	G3 (H)	118.72 (16.40)	122.94 (14.43)	112.39 (17.23)			
DMIQ2	G1 (L)	97.25 (17.86)	108.48 (17.48)	91.68 (15.28)	40.50***	112.05***	1.69
	G2 (M)	107.74 (16.53)	114.85 (15.83)	101.61 (14.61)			
	G3 (H)	115.62 (17.79)	120.03 (16.00)	109.09 (18.40)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: DMIQ1 = Domain-masculine intelligence type at pre-task estimation condition; DMIQ2 = Domain-masculine intelligence type at post-task estimation condition. TSP = Task-success probability estimation condition.

Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Males provided higher DMIQ/DMIQ1 estimates in across the three task-success groups, providing additional support for higher male self-confidence in the DMIQ estimation process. Hypothesis 7 was partially confirmed.

For DMIQ2, the interaction effect between TSP estimation condition and gender was not significant, F(2,651) = 1.69, $p = .19$, $\eta_p^2 = .01$. There was a statistically significant main effect for TSP, F(2,651) = 40.50, $p =.00$, $\eta_p^2 = .11$, with medium effect size. The main effect for gender, F(1,651) = 112.05, $p =.00$, $\eta_p^2 = .15$ was also significant, with large effect size. Planned contrasts revealed significant differences between Group 1 and Group 2 (Contrast Estimate -11.32, $p =.00$) and between Group 2 and Group 3 (Contrast Estimate -6.33, $p =.00$). Post-hoc comparisons using the Tukey HSD and Bonferroni tests revealed that the mean scores for Group 1 (<=3) were significantly different from Group 2 (3-4) as well as from Group 3 (4+). Group 2 mean scores also significantly differed from Group 3 mean scores. Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. As in DMIQ1, males provided higher DMIQ2 estimates in all three task-success groups. Hypothesis 8 was partially confirmed.

9.5.3.5. Impact of Gender on the Relationship between TCAP and DMIQ/DMIQ1 and DMIQ2

TCAP was collapsed into a categorical variable with three groups, with Group 1 containing individuals who solved the lowest numbers of psychometric problems, Group 2 containing individuals who solved an average number of problems, and Group 3 individuals that correctly solved the most problems. Results are presented in Table 9.5.8.

Table 9.5.8: Overview of TCAP Banded

	TCAP	N
Group 1	<=3	347
Group 2	7-9	295
Group 3	10+	319

Note: Computed using Visual Bander technique (SPPS 13.0)

Two 2-way between-groups analyses of variance were conducted to explore whether gender influences the relationship between TCAP and DMIQ/DMIQ1 and DMIQ2. Results are presented in Table 9.5.9.

For DMIQ/DMIQ1, the homogeneity of variance assumption was violated (Levene Statistic $p < .05$), indicating the groups variances were not equal. An alternative check for comparing variances was used. Firstly, the largest and the smallest standard deviations were squared. The largest squared SD was divided by the smallest squared SD, with resulting value of 1.17, which is smaller than the recommended value of 2, suggesting that the group variances, albeit not equal, were tolerable. Subsequently, the significance level was adjusted to $p < .01$.

Table 9.5.9: Two 2-way ANOVAs (TCAP and gender) on DMIQ/DMIQ1 and DMIQ2

Variable	TCAP Groups	Mean Score (SD) Total	Males	Females	*F*-score TCAP	Gender	TCAP x Gender
DMIQ1	G1 (L)	107.76 (17.95)	114.46 (18.31)	104.94 (17.10)	16.54***	98.59***	.62
	G2 (M)	109.67 (16.83)	116.55 (15.59)	105.14 (16.10)			
	G3 (H)	117.77 (15.35)	123.19 (14.59)	110.54 (13.25)			
DMIQ2	G1 (L)	94.39 (16.47)	102.10 (14.98)	91.60 (16.18)	30.29***	76.25***	.69
	G2 (M)	103.28 (19.10)	112.27 (17.16)	97.39 (18.02)			
	G3 (H)	112.94 (16.88)	119.55 (15.21)	104.14 (14.89)			

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed).
Note: DMIQ1 = Domain-masculine intelligence type at pre-task estimation condition; DMIQ2 = Domain-masculine intelligence type at post-task estimation condition. TCAP = Total correct aptitude problems.

The interaction effect between TCAP and gender was not significant, F(2,854) = .62, p =.54, η_p^2 = .00. The main effect for TCAP, F(2 854) = 16.54, p =.00, η_p^2 = .04 was significant, with small effect size. The main effect for gender, F(1,854) = 98.59, p =.00, η_p^2 = .10 was also significant, with medium effect size. Planned contrasts revealed significant differences between Group 1 and Group 3 (Contrast Estimate -4.15, p <.01) and between Group 2 and Group 3 (Contrast Estimate -6.02, p =.00). Post-hoc comparisons using the Games-Howell test indicated that the mean scores for Group 1 (<=6) were significantly different from Group 3 (10+). The mean scores for Group 2 (7-9) were significantly different from Group 3. No other significant mean score differences between the groups were observed. This was confirmed by the Ryan-Einot-Gabriel-Welch Range test of homogenous subsets. Males provided significantly higher DMIQ/DMIQ1 estimates across all three TCAP groups than did females, providing further support for the existence of male hubris in the estimation process. Hypothesis 9 was partially confirmed.

For DMIQ2, the interaction effect between TCAP and gender was not significant, $F(2,688) = .69$, $p > .05$, $\eta_p^2 = .00$. The main effect for TCAP, $F(2,688) = 30.29$, $p = .00$, $\eta_p^2 = .08$ was significant, with medium effect size. The main effect for gender, $F(1,688) = 76.25$, $p = .00$, $\eta_p^2 = .10$, was also significant, with medium effect size. Planned contrasts revealed significant differences between Group 1 and Group 2 (Contrast Estimate -11.49, $p = .00$) and between Group 2 and Group 3 (Contrast Estimate -7.02, $p = .00$).Post-hoc comparisons using the Tukey HSD and Bonferroni tests indicated that the mean scores for Group 1 (<=6) were significantly different from Group 2 (7-9) as well as from Group 3 (10+). Group 2 mean scores also significantly differed from Group 3 mean scores. Results were confirmed by the Ryan-Einot-Gabriel-Welch Range test. As in previous analysis, males provided higher DMIQ2 estimates than did females in all three TCAP groups. Hypothesis 10 was partially confirmed.

9.5.4. Discussion

The results of the combined samples provided further support for the results of the individual studies. Overall, the occurrence of HHE on DMIQ/DMIQ1 and DMIQ2 was confirmed with medium and very large effect sizes. Equally, a significant decrease in DMIQ estimates from DMIQ/DMIQ1 to DMIQ2 was observed (d =.58), with a medium effect sizes was observed. Equally, gender was found to be the best predictor of DMIQ/DMIQ1 and DMIQ2, accounting for 11% and 14% of explained variance respectively. Male advantage was confirmed on both TSP and TCAP, with males providing higher task-success estimate probes and solving correctly more psychometric problems than did female participants.

The only unpredicted but observed result was the relationship between age and the DMIQ/DMIQ1 and DMI2 that revealed that age played a role in the intelligence type, with older subjects providing higher estimates. When the data was split per gender, the correlational analysis revealed identical results for both genders, but only for the DMIQ/pre-task estimates, with older male and female participants providing higher estimates. However, for the post-task estimates, only younger males provided higher estimates.

The relationship between TSP and DMIQ type and the role of gender therein was investigated next. The findings confirmed the results of the five experimental studies, with males providing higher DMIQ/DMIQ1 and DMIQ2 estimates than females in all three task-success groups, providing further support for higher task confidence in males. Likewise, the provided DMIQ/DMIQ1 estimates reflected accurately the three TSP groups, with lowest DMIQ/DMIQ1 estimates provided by the lowest TSP group, average estimates by the average TPS group and the highest DMIQ/DMIQ1 estimates by the group with highest task-success probability estimates. These results were observed for the total sample as well as for each gender, providing additional support for the assertion that individuals are capable of accurate self-estimates of ability. Results for DMIQ2 were identical.

Lastly, the role of gender in the relationship between TCAP and the intelligence type was examined. As with TSP, the results of the combined sample affirmed the earlier results of the individual studies. For DMIQ/DMIQ1, males provided higher DMIQ/DMIQ1 estimates across the three TCAP groups, providing further support for the existence of male hubris in the self-estimation process. Equally, the intelligence type estimates accurately reflected the three TCAP groups, with the lowest DMIQ/DMIQ1 estimates provided by the group that solved the fewest psychometric problems, average estimate by the group that solved an average number of problems and the highest DMIQ/DMIQ1 estimates by the group that solved the most psychometric problems. The results were identical for DMIQ2.

Thus, the results of the overall combined sample further affirmed the findings of the individual studies. Equally, the observed results uphold the earlier findings that were made with smaller sample sizes.