Learning and Individual Differences 50 (2016) 23–29

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

The role of gender, task success probability estimation and scores as predictors of the domain-masculine intelligence type (DMIQ)

Josephine Storek [a], Adrian Furnham [a,b,*]

[a] Research Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom
[b] Norwegian Business School, Oslo University College London, United Kingdom

ARTICLE INFO

Article history:
Received 21 April 2015
Received in revised form 1 June 2016
Accepted 20 June 2016
Available online xxxx

ABSTRACT

This paper reports a study aimed at understanding correlates of self-estimated intelligence. Participants twice estimated their mathematical and spatial intelligence (called domain-masculine intelligence type: DMIQ) on a normal distribution, before and after taking ability tests. They completed a number of short numerical and logical ability tests after which they estimated their performance at a similar, more difficult task. Males gave higher estimates than females and did better on the tests. As predicted their estimates of their DMIQ reduced on the second occasion after testing. Gender, task score and estimated performance were all significant predictors of both DMIQ scores. Task confidence was the best predictor of both before and after test estimates, over and above gender and test score, explaining 17% and 23% of variance, respectively. This is explained in terms of Dweck's (2007) mindset theory and Eccles and Wigfield's (2002) motivation theory. Results are discussed in terms of the literature on self-estimated intelligence and stereotype threat.

1. Introduction

While an extensive body of self-estimated intelligence (SEI) research is available, only few SEI studies have used psychometric measures to compare the accuracy and validity of SEI estimates (Batey, Chamorro-Premuzic, & Furnham, 2009; Chamorro-Premuzic, Moutafi, & Furnham, 2005; Holling & Preckel, 2005). To the best of our knowledge, this research is the first experimental design in SEI that focuses on assessing gender differences in self-estimated intelligence using ability tests, repeated measurement as well as investigating the role of task confidence.

Evidence from more than thirty studies shows that stable and consistent universal gender differences in SEI exist in the general population (Furnham, 2001; Furnham & Shagabutdinova, 2012; Stieger et al., 2010; von Stumm, Chamorro-Premuzic, & Furnham, 2009). The strongest gender differences observed is on mathematical/logical and spatial intelligences, followed by overall (g) and verbal intelligence, with significantly higher self-estimates provided by males than females (Furnham, 2001; Rammstedt & Rammsayer, 2002a, 2002b). The magnitude of gender differences in mathematical/logical, spatial, overall and verbal self-assessed intelligences was further revealed in meta-analytical study (Szymanowicz & Furnham, 2011), with the biggest weighted mean effect sizes for mathematical/logical ($d = 0.44$), followed by spatial ($d = 0.43$), overall ($d = 0.37$) and verbal ($d = 0.07$) intelligences, with males providing higher estimates in all but verbal intelligence.

This phenomenon is known as the *hubris-humility effect* (HHE) (Beloff, 1992; Storek & Furnham, 2012, 2013a, 2013b, 2014). 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 unclear whether 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. Indeed, female self-estimates were shown to be significantly more accurate than were those of males.

Male self-estimates have been shown to be significantly inflated compared to their actual psychometric scores (Reilly & Mulhern, 1995). These findings were further substantiated by Carr, Hettinger-Steiner, Kyser, and Biddlecomb (2008), who reported that girls were more accurate in assessing their mathematical skills and knowledge, despite low math ability confidence. Unsurprisingly, boys were overconfident, with poorer performance.

One possible explanation is that in these studies females experience stereotype threat which increases their performance anxiety and hence outcomes both on actual personality tests as well as those examining self-estimates. It seems possible that people develop, often inaccurate, general perceptions of their overall ability (“ability self”) over time particularly as a results of schooling, which would impact on their self-rated ability and even test performance. There is a great deal of interest in the concept of stereotype threat as well as critique of its importance

* Corresponding author at: Research Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom.
E-mail address: a.furnham@ucl.ac.uk (A. Furnham).

http://dx.doi.org/10.1016/j.lindif.2016.06.011
1041-6080/

(Flore & Wicherts, 2015). Nevertheless, gender differences in intelligence may be mediated by various social and cultural factors that impact on consistent and stable gender-based stereotypic threats when it comes to anything concerning intelligence and its measurement. Age has been shown to be related to self-estimated intelligence for various reasons: younger people are often better educated than older people, but older people are often more self-confident and have received more feedback about their intelligence. Hence we will control for age in this study.

To further explore the *male-normative* perception of intelligence (Furnham, 2000), the *domain-masculine intelligence type* (DMIQ), which is a composite of mathematical/logical and spatial intelligences (Storek & Furnham, 2012, 2013a, 2013b, 2014), was introduced.

This study was designed to ascertain the determinants of gender differences in the *domain-masculine intelligence* by introducing a number of timed psychometric tasks (TCAP) and confidence assessments (TSP). As in previous research (Storek & Furnham, 2012, 2013a, 2013b, 2014), 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 task confidence as well as actual scores. 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.

The issue of task confidence is important in this literature and relatively unexplored. Storek and Furnham (2013a) examined the relationship between fixed vs growth mindset ideas derived from Dweck (2007). The suggestion was that those with a growth/incremental/malleable mindset would have greater task confidence over time because of their belief in their ability to learn. However, they found little evidence of a significant relationship between SEI and mindset. However, this may have been because the mindset measure was too general, and neither task nor ability specific. Further, it took not account of specific feedback from task success or failure. In this study we asked participants to estimate how they would do in a similar task to the one they had done. Thus, we expected that entity theorists would tend to be more pessimistic with lower self-confidence than more optimistic incremental theorists. Equally, this self-confidence should relate to effort in subsequent tests which would be self-fulfilling. Indeed task confidence can also be understood in terms of Eccles expectancy value theory, which suggests task persistence is a function of subjective task values. That is, initial task confidence is probably related to subjective task values that are related to intrinsic motivation, interest and effort.

Gender stereotypes, threats and self-confidence are all likely to play a role in HHE or the display of male hubris and female humility in estimation of abilities. Participants were asked to undertake a gender stereotype-inducing task, i.e., numerical and reasoning aptitude problems that are likely to increase hubris and humility as well as general stereotype threats (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 (Burson, Larrick, & Klayman, 2006; Carr et al., 2008; Dunning, Griffin, Milojkovic, & Ross, 1990; Pallier, 2003). After each block, participants were asked to estimate their task success confidence.

Various hypotheses were tested some essentially replicating previous studies. It was predicted that HHE will be confirmed on DMIQ at the pre-task (T1) and post-task (T2) estimating conditions (H1). However, a more important study-specific hypothesis was that there will be a significant decrease in DMIQ estimates from T1 to T2 following the gender stereotype-inducing task (H2). This could be seen to be a manifestation of stereotype threat.

Existing literature suggests that males have higher self-confidence in general but particularly with respect to intelligence, despite being inaccurate about their (math) skills or underperforming, whereas females often lack confidence, while being accurate or outperforming males (Carr et al., 2008; Eccles-Parsons, Adler, & Meece, 1984; Pallier, 2003). Consequently replicating other research, males are expected to provide significantly higher task success probability estimations (TSP) (i.e., self-confidence) than females (H3).

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

The more important experimental hypotheses are these: gender is expected to be the best predictor of before and after self-estimates, namely, DMIQ T1 (H5) and DMIQ T2 (H6) over and above TSP and TCAP. Finally, gender is presumed to influence the relationship between TSP and DMIQ T1 (H7) and DMIQ T2 (H8). Gender is also expected to affect the relationship between TCAP and both DMIQ T1 (H9) and DMIQ T2 (H10).

2. Methods

2.1. Participants

A total of 488 participants from general public 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.

2.2. Measures

2.2.1. Repeated measure of domain-masculine intelligence type (DMIQ)

Based on the self-estimated intelligence measure (Furnham, 2001), this shortened version had the same properties and layout, but only included mathematical/logical and spatial intelligences that together form the domain-masculine intelligence type. Participants were shown a bell curve with IQ scores and asked to estimate their mathematical/logical and spatial intelligences, which were provided with detailed descriptions. Participants were asked to estimate their mathematical/logical and spatial intelligences on two occasions, prior (T1) and post (T2) to completing a psychometric task (TCAP) and assessing their task success confidence (TSP). Individual scores for DMIQ were computed. Alphas for DMIQ T1 and DMIQ T2 were 0.82 and 0.88, respectively.

2.2.2. Psychometric aptitude task—total correct aptitude problems (TCAP)

2.2.2.1. 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). Participants were informed that items in each block varied in difficulty level, ranging from elementary to difficult. A time limit of 90 s was given for each block of problems. Participants were advised to leave unanswered problems blank, in order not to exceed the time limit, or face disqualification. The time limit was set to reflect a real-life intelligence-testing situation, with the entire task taking 7.5 min to complete. Correct answers were available at the end of the survey. Alpha for the fifteen items was 0.93.

2.2.2.2. Task success probability estimation measure (TSP) (Storek & Furnham, 2012). 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