Test anxiety and performance-avoidance goals explain gender differences in SAT-V, SAT-M, and overall SAT scores

Brenda Hannon *

Department of Psychology, Texas A&M – Kingsville, Kingsville, Texas, United States

A R T I C L E   I N   P R E S S

ARTICLE INFO

Article history:
Received 2 April 2012
Received in revised form 7 June 2012
Accepted 7 June 2012
Available online xxxx

Keywords:
SAT
Gender differences
Test anxiety
Performance-avoidance goals

A B S T R A C T

This study uses analysis of co-variance in order to determine which cognitive/learning (working memory, knowledge integration, epistemic belief of learning) or social/personality factors (test anxiety, performance-avoidance goals) might account for gender differences in SAT-V, SAT-M, and overall SAT scores. The results revealed that none of the cognitive/learning factors accounted for gender differences in SAT performance. However, the social/personality factors of test anxiety and performance-avoidance goals each separately accounted for all of the significant gender differences in SAT-V, SAT-M, and overall SAT performance. Furthermore, when the influences of both of these factors were statistically removed simultaneously, all non-significant gender differences reduced further to become trivial by Cohen's (1988) standards. Taken as a whole, these results suggest that gender differences in SAT-V, SAT-M, and overall SAT performance are a consequence of social/learning factors.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

For students in the United States the SAT (i.e., the Scholastic Assessment Test, formerly called the Scholastic Aptitude Test) is one of the most, if not the most, important measure of academic achievement because of its dominant role in high stakes decisions of college admissions (Hannon & McNaughton-Cassill, 2011). However, the SAT is also plagued with controversies, including those about construct validity and gender differences. This latter controversy is particularly relevant to the present study because the primary goal here is to examine whether social/personality and cognitive/learning factors might explain gender differences in verbal SAT (i.e., SAT-V), math SAT (i.e., SAT-M), and overall SAT scores. Social/personality and cognitive/learning factors were selected because recent research suggests that factors from these two content areas (e.g., test anxiety, working memory, integrating information from prior knowledge and the text) account for as much as 44.6% of the variance in SAT scores (Hannon & McNaughton-Cassill, 2011). Consequently, it is possible that one or more of these factors might also account for gender differences in SAT performance.

1.1. Background

At the heart of the gender difference controversy is the finding that females routinely score lower than males on both the SAT-V and SAT-M (Halpern et al., 2007). Indeed, data for the 20-year period from 1987 to 2006 indicates that the mean SAT-V score was 508 for males and 501 for females while the mean SAT-M score was 528 for males and 492 for females (Kobrin, Sathy, & Shaw, 2007). Moreover, the 30–40 point advantage for males on the SAT-M, which routinely shows the largest gender difference, has remained relatively unchanged for nearly 40 years (Halpern et al., 2007). Gender differences in SAT scores persist even though females routinely achieve higher grades in every subject in school, including mathematics and science, and even though females represent the majority of college enrollments in the United States (Halpern et al., 2007; Mau & Lynn, 2001). In other words, a grade–test disparity exists in successful achievement such that females achieve higher grades in school whereas males achieve higher scores on standardized tests designed for admissions to colleges, universities, and graduate programs (Halpern et al., 2007; Mau et al., 2001). One outcome of this grade-test disparity is that standardized tests slightly under-predict college performance for females but over-predict college performance for males (Halpern et al., 2007).

Given these findings, it is clear that it would be a considerable benefit to students, teachers, researchers, and the public in general to know which factors contribute to gender differences in SAT-V, SAT-M, and overall SAT scores. Unfortunately, although a number of studies have examined gender differences in cognitive abilities most studies have used other measures rather than the SAT (Casey, Nuttall, & Pezaris, 2001; Gallagher & Delisi, 1994; see Halpern et al., 2007 for a review). Moreover, the few studies that have examined gender differences in SAT scores have primarily focused on the SAT-M (Benbow & Stanley, 1980; Casey, Nuttall, & Pezaris,
1997; Casey, Nuttall, Pezaris, & Benbow, 1995; Casey et al., 2001), students who have high-ability (Benbow & Stanley, 1983; Casey et al., 1997; Lubinski & Benbow, 1992), or pre-adolescents (Benbow & Stanley, 1980; Casey et al., 2001). For example, Casey et al. (1995) observed that the gender difference in SAT-M scores was eliminated for both high-ability, college-bound high-school students and a diverse sample of college students (i.e., 48% of their participants) when mental rotation skills, measured by the Vandenberg Test of Mental Rotation (Vandenberg & Kuse, 1978), were controlled. However, their low-ability group (28% of their participants) revealed no gender difference in SAT-M scores and their talented pre-adolescent group (24% of their participants) revealed a gender difference in SAT-M scores even after mental rotation skills were controlled. In a related study, Casey et al. (1997) showed that mental rotation skills accounted for some of the gender difference in SAT-M scores in a population of high-ability, college-bound students. However, they also showed that math self-confidence accounted for some of the gender difference in SAT-M scores, whereas math anxiety did not. Thus, although individual differences in mental rotation skills might explain some of the gender difference in SAT-M scores, certainly mental rotation skills are not the sole explanation. Nor is it clear that individual differences in mental rotation skills account for gender differences in SAT-M scores for low-ability students, a group of students who frequently fail to enter colleges because of low SAT scores.

Other studies have suggested that gender differences in SAT scores are, in part, a consequence of demographic differences, such as level of parental education, parental income, number of high school math and English classes, and race/ethnicity. For example, Young and Fisler (2000) observed that adjusting SAT scores for the above-mentioned demographic factors decreased the gender difference in SAT-M scores by approximately 11.5 points. Unexpectedly though, these authors also observed that adjusting for these same demographic factors increased the gender difference in SAT-V scores by approximately 5.0 points. Thus, although demographic factors might explain some of the gender difference in SAT-M scores, clearly they are not the only explanation. Nor do demographic factors explain the gender difference in SAT-V scores.

1.2. Summary and the present study

In summary, studies have attempted to explain gender differences in SAT scores in terms of individual differences in mental rotation skills, math self-confidence, test anxiety, and demographics. However, the quantity of studies has been quite limited and to date, no study has demonstrated that one or a combination of these factors account for all of the gender differences in SAT-V, SAT-M, and overall SAT scores in a population of college students. The present study addresses this shortcoming by determining whether gender differences in SAT scores might be explained in terms of individual differences in cognitive/learning and social/personality factors. These factors were selected because recent research suggests that factors from these two content areas account for 44.6% of the variance in SAT-V, SAT-M, and overall SAT scores (e.g., Hannon & McNaughton-Cassill, 2011). Consequently, it is possible that these factors might also account for gender differences in SAT scores.

The influences of three cognitive/learning factors were investigated: (i) working memory, a cognitive resource that is shared by many cognitive processes, (ii) knowledge integration, characterized as the ability to integrate prior knowledge from long-term memory with new text-based information, and (iii) episodic belief of learning, also known as metacognitive awareness. These three cognitive/learning factors were selected because previous research suggests that they routinely account for large amounts of variance in SAT scores (Daneman & Carpenter, 1980; Daneman & Hannon, 2001; Hannon & Daneman, 2001; Rukavina & Daneman, 1996; Turner & Engle, 1989). Presumably in a test-taking situation with greater cognitive abilities are able to integrate and process test questions quicker and more accurately than students with poorer cognitive abilities.

Two social/personality factors were investigated: (i) test anxiety, characterized as a mental state that includes behavioral, cognitive, emotional, and/or bodily reactions (McIlroy, Bunting, & Adamson, 2000; see Hembree, 1988 for a review), and (ii) performance-avoidance goals, characterized as one’s desire to not perform poorly on measures of achievement (Elliott & Church, 1997). These two social/personality factors were selected because previous research suggests that higher test anxiety/performance avoidance scores result in lower SAT-V, SAT-M, and SAT scores, $r = −.33$ to $−.41$ (Hannon & McNaughton-Cassill, 2011). Furthermore, test anxiety as an explanation for gender differences in SAT scores warrants revisiting because previous research examining the gender-test anxiety-SAT relationship used a small subset of math test anxiety questions adopted from a little-known questionnaire. Thus, it is possible that gender differences in SAT scores might be explained by individual differences in test anxiety when the test anxiety measure is more widely-accepted. A second reason for re-examining the gender-test anxiety-SAT relationship is because females have greater test anxiety than males (e.g., Chapell et al., 2005). Consequently test anxiety is a good candidate for explaining gender differences in SAT scores. Finally, it should be noted that test anxiety and performance-avoidance goals are strongly related. Presumably in a test-taking situation the act of attempting to avoid a negative outcome elicits test anxiety, especially when a student focuses on normal performance in the face of possible failure (Elliot & McGregor, 1999, p. 629).

2. Methods

2.1. Participants

The 229 participants were University of Texas students who received $40.00–$50.00 for participating in one of two large scale studies: (i) a three-year study examining the relationships between test anxiety and social/attitudinal beliefs and cognitive/learning abilities in European-American versus Hispanic students (grant #5R24MH070636) and (ii) a three-year study assessing the relationships between social/personality and cognitive/learning abilities and frontal/hippocampus functioning. All participants were freshmen who had completed the SAT, were dominant English speakers, and were free of any known learning disabilities. Their mean age was 18.42 years ($std = 0.69$) and 113 were female and 116 were male.¹

2.2. Measures

Because all of the measures are explained well in other published studies, they are only briefly described below. References are provided for each measure.

2.2.1. Cognitive/Learning measures

Knowledge integration was assessed using the component processes task (Hannon, 2012). Briefly, in this task students first learn three-sentence paragraphs that described relationships among two real and three nonsense terms; for instance: A NORT resembles a JET but is faster and weighs more. A BERL resembles a CAR but is slower and weighs more. A SAMP resembles a BERL but is slower and weighs

¹ Data for 133 of the 229 participants were used in Hannon & McNaughton-Cassill, 2011.
more. Next, they answer true–false statements that assess knowl-
edge integration (e.g., Like PLANES, NORTS fly in the air). Accuracy
was the dependent measure and higher scores indicated better
performance.

The measures of working memory were the reading and opera-
tion span tasks (Daneman & Carpenter, 1980; Turner & Engle,
1989). The total number of words recalled was the dependent mea-
sure for each task.

Finally, students completed 12 items selected from two subsets
of Schommer's (1990) epistemology questionnaire. A sample state-
ment is You will just get confused if you try to integrate new ideas in a
textbook with knowledge you already have about a topic. Agreement
level for each statement was identified using a 5-point Likert scale.
Lower scores represented mature beliefs about learning whereas
higher scores represented naïve beliefs.

2.2.2. Social/Personality measures
Each measure was presented on a computer. Students selected
answers and the research assistant typed in their responses. In
order to protect the privacy of the students, the research assistant
could not see the computer screen.

Students completed Sarason's (1978) measure of test anxiety,
which included 37 true–false statements. A sample item is I wish
examinations did not bother me so much. Higher scores indicated
greater test anxiety.

Students also completed Elliot and Church's (1997) measure of
achievement motivation goals. Although this measure includes
three scales, only the performance-avoidance goals scale was used.
A sample item from this scale is I just want to avoid doing poorly in
this class. Students selected an answer for each statement using a
7-point Likert scale. Higher scores indicated a greater propensity to-
wards performance-avoidance.

2.2.3. Measures of academic achievement
SAT-V, SAT-M, and overall SAT scores were obtained from uni-
versity records.

3. Results
3.1. Descriptive Statistics and Correlations
As Table 1 shows, all the measures had large ranges and the
skewness and kurtosis statistics suggested normal distributions
(i.e., all values <3). Further, the gender-SAT-M and gender-overall
SAT correlations were significant, r = -.17 and r = -.18 respectively
while the gender-SAT-V correlation was marginaly significant,
r = -.13. These three negative correlations suggest that female have
lower scores than males, a finding which is consistent with previ-
ous research (Halpern et al., 2007). Additionally, all of the cogni-
tive/learning and social/personality factors significantly
 correlated with all three SAT scores, a finding which replicates pre-
vious research (Hannon & McNaughton-Cassill, 2011). However of
the five cognitive/learning and social/personality factors, only the
two social/personality factors, test anxiety and performance-avoid-
ancegoals, correlated with gender. The significant gender-
test anxiety and gender-performance-avoidance goals correlations
indicate that females experience more test anxiety and have higher
performance-avoidance goals than males. In the context of the
present study, these significant correlations also suggest that test
anxiety and performance-avoidance goals might account for some
of the gender differences in SAT performance.

3.2. Explaining gender differences in SAT measures
Three sets of analyses assessed whether cognitive/learning and
social/personality factors explain gender differences in SAT scores.
Set one used Analysis of Variance (i.e., ANOVA) to identify the mag-
nitude of the gender differences in SAT-V, SAT-M, and overall SAT
scores. In each ANOVA gender was a between-subjects variable.

The second and third sets of analyses used Analysis of Covari-
ance (i.e., ANCOVA). The second set of analyses determined
whether cognitive/learning factors might account for gender dif-
f erences in SAT scores. The third set of analyses determined
whether social/personality factors might account for gender differ-
ences in SAT scores.

Partial $\eta^2$ are reported for all effects. According to Cohen (1988),
a large effect, which is rare in the behavioral sciences, has an
$\eta^2 = >.14$; a medium effect has an $\eta^2$ between .06 and .139; a
small effect has any $\eta^2$ between .011 and .059; and effects that are
considered trivial have an $\eta^2 < .01$.

3.2.1. Magnitudes of gender differences in SAT-V, MAT-M, and overall
SAT scores
Three ANOVAs—one for each SAT measure—were completed. As
Table 2 and 3 show, there was a marginally significant gender dif-
ference in SAT-V scores, $F(1,227) = 3.76$, $p = .054$, partial $\eta^2 = .016$
such that males scored higher than females, 546 versus 523 respec-
tively. There were also significant gender differences in SAT-M and
overall SAT scores, $F(1,227) = 7.62$, $p = .032$, partial $\eta^2 = .032$ and
$F(1,227) = 6.76$, $p = .01$, partial $\eta^2 = .029$ respectively, such that males
scored higher than females on the SAT-M, 551 versus 522 respec-
tively, as well as the overall SAT, 1097 versus 1046 respectively.

3.2.2. Cognitive/Learning explanations of SAT scores
Three ANCOVAs—one for each SAT measure—were completed. Working memory, knowledge integration, and epistemic belief
of learning were the covariates in each ANCOVA.

When the influences of the cognitive/learning factors were sta-
tistically removed from SAT-V, SAT-M, and overall SAT scores, the
marginal gender difference for the SAT-V remained, $F(1,223) = 3.27$, $p = .072$, partial $\eta^2 = .014$, the significant gender dif-
ference for the SAT-M scores remained, $F(1,223) = 6.67$, $p = .01$
partial $\eta^2 = .029$ and the gender significant gender difference for overall
SAT scores remained, $F(1,223) = 6.32$, $p = .013$, partial $\eta^2 = .028$.
In other words, these results suggest that none of the cognitive/
learning factors account for the gender differences in SAT scores.

3.2.3. Social/Personality explanations of SAT scores
Three ANCOVAS, one for each SAT measure—were completed. Working memory, knowledge integration, and epistemic belief
of learning were the covariates in each ANCOVA.

When the influences of the social/personality factors were sta-
tistically removed from SAT-V, SAT-M, and overall SAT scores, the
marginal significant significant gender differences for all three
scores were also eliminated: SAT-V scores, $F < 1.0$, partial $\eta^2 = .000$, SAT-M scores, $F(1,225) = 1.75$, $p = .19$, partial $\eta^2 = .008$, and
overall SAT scores, $F < 1.0$, partial $\eta^2 = .003$. In other words, the social/personality factors of test anxiety and performance-
avoidance goals accounted for all of the gender differences in SAT
scores. This is the first time that all of the gender differences in
all of the SAT measures have been accounted for.

To determine whether one or both test anxiety and perfor-
mance-avoidance goals are necessary in order to eliminate the
gender differences in SAT scores, two additional sets of ANCOVAs were
completed: one with test anxiety as a single covariate and the other
with performance-avoidance goals as a single covariate. The results
revealed that each social/personality factor accounted for all of the
significant gender differences in SAT-V, SAT-M, and overall SAT

---

2 Elliot and Murayama (2008) have updated the Elliot-Church (1997) measure of
achievement motivation goals. However, the updated version is highly similar to the
original measure used in the present study.
scores. Specifically, as Table 3 shows when the influences of test anxiety were statistically removed, the gender differences were eliminated for the SAT-V, $F < 1.0$, partial $g^2 = .001$, SAT-M, $F (1,226) = 2.71$, $p = .10$, partial $g^2 = .012$, and the overall SAT, $F (1,226) = 1.34$, $p = .25$, partial $g^2 = .006$. Similarly, when the influences of performance-avoidance goals were statistically removed, the gender differences were eliminated for the SAT-V, $F < 1.0$, partial $g^2 = .003$, SAT-M, $F (1,226) = 2.85$, $p = .093$, partial $g^2 = .012$, and overall SAT scores, $F (1,226) = 2.00$, $p = .16$, partial $g^2 = .009$.

Nevertheless, the residual effect sizes for gender differences in these two single-covariate ANCOVAs were greater than the residual effect sizes for gender differences in the ANCOVA that included both test anxiety and performance-avoidance goals as co-variates. Indeed as Table 3 shows, the effect sizes for gender differences in SAT-M scores (i.e., .012) in both single-covariate ANCOVAs were small but still acceptable effects by Cohen’s standards (1988). In contrast, the effect size for gender differences in SAT-M scores in the combined ANCOVA was trivial (i.e., .008). Thus, although removing the influences of either test anxiety or performance-avoidance goals eliminated the statistically significant gender differences in SAT scores, in order to render all effect sizes as trivial, the influences of both of these factors need to be removed.

4. Discussion

This study showed that the social/personality factors of test anxiety and performance-avoidance goals explain gender differences in SAT-V, SAT-M, and overall SAT scores. Personality and Individual Differences (2012), http://dx.doi.org/10.1016/j.paid.2012.06.003
behavioral, cognitive, emotional, and/or bodily reactions (McIlroy et al., 2000).

From a theoretical perspective, the best explanation for the gender differences in SAT scores is one that includes the negative influences of both test anxiety and performance-avoidance goals. Elliot and McGregor (1999) suggest that the act of attempting to avoid a negative outcome in a test-taking situation elicits test anxiety, especially when a student focuses on normal performance in the face of possible failure. By this account, perhaps more females than males avoid studying for the SAT which, in turn, creates more anxiety as they try to complete the SAT. Alternatively, perhaps females are more likely than males to set modest achievement goals for the SAT because they don’t wish to perform poorly. These modest goals, in turn, elicit greater test anxiety, especially when females focus on normal performance in the face of possible failure. In support of the latter explanation is research that shows a student’s desire to achieve high scores for the sake of appearances predicts higher SAT scores (Rose, Hall, Bolen, & Webster, 1996).

Regardless of the theoretical explanation, the good news is that altering the goals of test takers can alleviate/remove at least some of the deleterious influences of test anxiety and performance-avoidance goals (Elliot, 1999). For example, Elliot and Harackiewicz (1996) observed that when instructions for completing a puzzle were more performance-approach oriented, students performed better than when the instructions were more performance-avoidance related. In the context of the present study perhaps altering the attitudes that female students have about the SAT might narrow the gender gap in SAT performance.

Finally, the present findings potentially inform the stereotype threat-math performance literature. According to Brodish and Devine (2009), the influence that stereotype threat exerts on math performance is mediated by performance-avoidance goals and worry/anxiety. Consequently, this account predicts higher correlations between math performance and performance-avoidance goals and test anxiety for females than males because females were found to experience stereotype threat to a greater extent. However, the present study did not support this prediction. Rather, both factors predicted math performance slightly less for females than males (performance avoidance: -.25 and -.35 and test anxiety: -.20 and -.25 respectively), a finding that is more consistent with researchers that question whether stereotype threat can explain the gender gap in math performance (Stoet & Geary, 2012).

The present study also has limitations. For example, because all of the dependent variables (SAT-V, SAT-M, overall SAT) were assessed before any of the predictors were measured, this study does not establish causality. Additionally, the present study uses a population of students at one university rather than several populations from multiple universities. Future research might wish to examine whether the present findings generalize to other populations of students.

In summary, the present study shows that the social/personality factors of test anxiety and performance-avoidance accounted for gender differences in SAT performance, the cognitive/learning factors of working memory, knowledge integration, and epistemic belief of learning did not.

Acknowledgements

I would like to thank Stephanie Keller for her invaluable help with data collection and data scoring. This research was partially funded by grant 5R24MH070636.

References