Personality and Individual Differences xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Personality and Individual Differences

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

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

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

ABSTRACT

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.

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,

* Tel.: +1 306 280 0628.
E-mail address: brenda.a.hannon@gmail.com

0191-8869/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.paid.2012.06.003

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) epistemic 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 students 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 (Elliot & 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 knowledge 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 operation span tasks (Daneman & Carpenter, 1980; Turner & Engle, 1989). The total number of words recalled was the dependent measure for each task.

Finally, students completed 12 items selected from two subsets of Schommer's (1990) epistemology questionnaire. A sample statement 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.[2] 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-pointLikert scale. Higher scores indicated a greater propensity towards performance-avoidance.

2.2.3. Measures of academic achievement

SAT-V, SAT-M, and overall SAT scores were obtained from university 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 marginally significant, $r = -.13$.These three negative correlations suggest that female have lower scores than males, a finding which is consistent with previous research (Halpern et al., 2007). Additionally, all of the cognitive/learning and social/personality factors significantly correlated with all three SAT scores, a finding which replicates previous 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-avoidance goals, 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.

[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.

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 magnitude 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 Covariance (i.e., ANCOVA). The second set of analyses determined whether cognitive/learning factors might account for gender differences in SAT scores. The third set of analyses determined whether social/personality factors might account for gender differences in SAT scores.

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

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 Tables 2 and 3 show, there was a marginally significant gender difference 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 respectively. 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 respectively, 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 statistically 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 difference for the SAT-M scores remained, $F(1,223) = 6.67$, $p = .01$, *partial* $\eta^2 = .029$ and the 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. In each ANCOVA test anxiety and performance-avoidance goals were the covariates.

When the influences of the social/personality factors were statistically removed from SAT-V, SAT-M, and overall SAT scores, the marginally significant/significant gender differences for all three measures 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 performance-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

Table 1
Descriptive statistics, cronbach alphas and correlations among gender, measures of cognitive/learning and social/personality factors, SAT-V, SAT-M, and overall SAT Scores ($n = 229$).

	1	2	3	4	5	6	7	8	9	10
1. Gender	---	-.13**	-.18*	-.17*	-.01	-.04	-.10	.01	.26*	.27*
2. SAT-V		---	.62*	.91*	.33*	.30*	.29*	-.38*	-.39*	-.30*
3. SAT-M			---	.89*	.29*	.31*	.26*	-.23*	-.31*	-.32*
4. Overall SAT				---	.34*	.34*	.31*	-.34*	-.39*	-.34*
5. Reading span					---	.71*	.25*	-.09	-.14*	-.08
6. Operation span						---	.25*	-.18*	-.21*	.03
7. High-knowledge integration							---	-.10	-.13*	-.18*
8. Epistemic belief of learning								---	.25*	.19*
9. Test Anxiety									---	.56*
10. Performance avoidance										---
Mean		535.07	536.59	1072.00	58.94	74.38	26.38	34.42	15.37	4.61
Standard deviation		85.91	80.90	150.14	10.03	11.94	5.41	4.87	7.20	1.06
Skewness		0.43	0.12	0.21	0.17	-0.20	-0.30	-0.08	0.34	-0.48
Kurtosis		-0.31	0.34	-0.15	-0.33	-0.44	-0.81	-0.50	-0.71	0.10
Lowest score		350.00	320.00	690.00	36.00	46.00	14.00	23.00	2.00	1.83
Highest score		800.00	800.00	1540.00	89.00	98.00	36.00	45.00	33.00	7.00
Maximum score		800.00	800.00	1600.00	100.00	100.00	36.00	60.00	37.00	7.00

Note. Standard deviations are in brackets.
* $p < .05$.
** $p < .06$.

Table 2
Descriptive statistics, for measures of cognitive/learning, social/personality factors, SAT-V, SAT-M, and overall SAT Scores as a function of Gender.

	Males ($n = 116$)	Females ($n = 113$)
SAT and GPA		
SAT-V	545.86 (89.06)	523.98 (81.47)
SAT-M	550.95 (82.41)	521.86 (76.92)
Overall SAT	1096.81 (155.48)	1045.84 (140.50)
Cognitive/Learning factors		
Reading span	59.06 (10.38)	58.82 (9.71)
Operations span	74.85 (12.09)	73.89 (11.81)
Knowledge integration	26.89 (5.56)	25.86 (5.22)
Epistemic belief of learning	34.39 (4.80)	34.46 (4.96)
Social/Personality factors		
Test Anxiety	13.53 (6.67)	17.26 (7.26)
Performance-avoidance goals	4.33 (1.15)	4.90 (0.87)

Note. Standard deviations are in brackets.

Table 3
Effect sizes for gender differences in SAT-V, SAT-M, and overall SAT scores as a function of covariates.

	SAT-V	SAT-M	Overall SAT
Initial effect sizes	.016*	.032*	.029*
Effect sizes after influences of cognitive/learning factors removed	.014*	.029*	.028*
Effect sizes after influences of both social/personality factors removed	.000	.008	.003
Effect sizes after influence of test anxiety is removed	.001	.012*	.006
Effect sizes after influence of performance avoidance is removed	.003	.012*	.009

Note. Effect sizes are partial η^2.
* indicates non-trivial effect sizes.

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* $\eta^2 = .001$, the SAT-M, $F\ (1,226) = 2.71$, $p = .10$, *partial* $\eta^2 = .012$, and the overall SAT, $F\ (1,226) = 1.34$, $p = .25$, *partial* $\eta^2 = .006$. Similarly, when the influences of performance-avoidance goals were statistically removed, the gender differences were eliminated for the SAT-V, $F < 1.0$, *partial* $\eta^2 = .003$, SAT-M, $F\ (1,226) = 2.85$, $p = .093$, *partial* $\eta^2 = .012$, and overall SAT scores, $F\ (1,226) = 2.00$, $p = .16$, *partial* $\eta^2 = .009$.

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 each separately accounted for all of the gender differences in SAT performance, whereas none of the cognitive/learning factors did. Furthermore, when the influences of test anxiety and performance-avoidance goals were statistically removed simultaneously, all non-significant gender differences reduced further to the point that they were trivial by Cohen's (1988) standards.

The present findings are somewhat inconsistent with the research of Casey et al. (1997) who observed that math test anxiety failed to account for the gender difference in SAT-M scores. However, Casey et al. used a subset of questions taken from a fairly unknown test anxiety questionnaire, whereas the present study used a widely-accepted measure of test anxiety, namely the Sarason. Consequently, it is possible that Casey et al.'s measure of test anxiety failed to capture all of the components of test anxiety, such as 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 this 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 orientated, 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 experience stereotype threat to a great 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 to thank Stephanie Keller for her invaluable help with data collection and data scoring. This research was partially funded by grant 5R24MH070636.

References

Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science, 210*, 1262–1264.
Benbow, C. P., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science, 222*, 1029–1031.
Brodish, A. B., & Devine, P. G. (2009). The role of performance-avoidance goals and worry in mediating and relationship between stereotype threat and performance. *Journal of Experimental Social Psychology, 45*, 180–185.
Casey, M. B., Nuttall, R. L., & Pezaris, E. (1997). Mediators of gender differences in mathematics college entrance test scores: A comparison of spatial skills with internalized beliefs and anxieties. *Developmental Psychology, 33*, 669–680.
Casey, M. B., Nuttall, R. L., & Pezaris, E. (2001). Spatial-mechanical reasoning skills versus mathematics self-confidence as mediators of gender differences on mathematics subtests using cross-national gender-based items. *Journal for Research in Mathematics Education, 32*, 28–57.
Casey, M. B., Nuttall, R. L., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology, 31*, 697–705.
Chapell, M. S., Blanding, Z. B., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., et al. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97*, 268–274.
Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). New York: Academic Press.
Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.
Daneman, M., & Hannon, B. (2001). Using working memory theory to investigate the construct validity of multiple-choice reading comprehension tests such as the SAT. *Journal of Experimental Psychology: General, 130*, 208–223.
Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*, 169–189.
Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72*, 218–232.
Elliot, A. J., & Harackiewicz (1996). Approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72*, 218–232.
Elliot, A. J., & McGregor, H. (1999). Test anxiety and the hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 74*, 628–644.
Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology, 100*, 613–628.
Gallagher, A., & DeLisi, R. (1994). Gender differences in Scholastic Aptitude Test-Mathematics problem solving among high-ability students. *Journal of Educational Psychology, 86*, 204–211.
Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Shibley Hyde, J., & Gensbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*, 1–50.
Hannon, B. (2012). Understanding the relative contributions of lower-level word processes, higher-level processes, and working memory to reading comprehension performance in proficient adult readers. *Reading Research Quarterly, 47*, 125–152.
Hannon, B., & Daneman, M. (2001). A new tool for measuring and understanding individual differences in the component processes of reading comprehension. *Journal of Educational Psychology, 93*, 103–128.
Hannon, B., & McNaughton-Cassill, M. (2011). *SAT performance: Understanding the contributions of cognitive/learning and social/personality factors Applied Cognitive Psychology, 25*, 528–535.
Hembree, R. (1988). Correlates, causes, and treatment of test anxiety. *Review of Educational Research, 58*, 47–77.
Kobrin, J.L., Sathy, V., & Shaw, E.J. (2007). A historical view of subgroup performance difference on the SAT reasoning test. Research Report No. 2006–5. New York, NY: College Board Publications.
Lubinski, D., & Benbow, C. P. (1992). Gender differences in abilities and preferences among the gifted: Implications for the math/science pipeline. *Current Directions in Psychological Science, 1*, 61–66.
Mau, W., & Lynn, R. (2001). Gender differences on the Scholastic Aptitude Test, the American College Test, and college grades. *Educational Psychology, 21*, 133–136.
McIlroy, D., Bunting, B., & Adamson, G. (2000). An evaluation of the factor structure and predictive utility of a test anxiety scale with reference to students' past performance and personality indices. *British Journal of Educational Psychology, 70*, 17–32.
Rose, R. J., Hall, C. W., Bolen, L. M., & Webster, R. E. (1996). Locus of control and college students approaches to learning. *Psychological Reports, 79*, 163–171.
Rukavina, I., & Daneman, M. (1996). Integration and its effect on acquiring knowledge about competing scientific theories from text. *Journal of Educational Psychology, 88*, 272–287.
Sarason, I. G. (1978). Test anxiety, stress, and social support. *Journal of Personality, 49*, 101–114.
Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*, 498–504.
Stoet, G., & Geary, D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Review of General Psychology, 16*(1), 93–102.
Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127–154.
Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 48*, 599–604.
Young, J. W., & Fisler, J. L. (2000). Sex differences on the SAT: An analysis of demographic and educational variables. *Research in Higher Education, 41*, 401–416.

Please cite this article in press as: Hannon, B. 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