STEREOTYPICAL SCIENCE: EXPLORING HIGH SCHOOL OCCUPATIONAL
PREFERENCES FOR SCIENCE BY SEX, PERSONALITY,
AND COGNITIVE ABILITY

Sarah Ferguson, M.S.

Dissertation Prepared for the Degree of
DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2016

APPROVED:

Darrell M. Hull, Major Professor
Robin K. Henson, Committee Member
Todd Kettler, Committee Member
Endia Lindo, Committee Member
Abbas Tashakkori, Chair of the Department
of Educational Psychology
Jerry R. Thomas, Dean of the College of
Education

Circumscription and Compromise theory suggests self-concept and sex stereotype explain occupational preferences, including preferences for science, technology, engineering and mathematics (STEM). Support exists for sex differences between males and females in both science degrees and science careers. The main thrust of observed sex differences in science lies in the development of occupational interest, as it has been suggested females are encouraged away from science due to stereotypes and social pressure. The present study evaluates high school juniors and seniors \((n = 295)\) to explore their preference for science as indicated by science motivation, attitude, academic experience, and interest. Latent Profile Analysis was used to model profiles of preferences for science with a person-centered approach. Then, the impact of self-concept variables was explored and four profiles of science interest were identified. Sex differences were identified based on science interest, but were not always in favor of males. Covariate analysis indicates vocabulary ability and personality as significantly different for students in the high science interest profile. Implications of these results and future research directions are discussed.
ACKNOWLEDGEMENTS

Throughout my graduate school career I have benefited tremendously from the support and encouragement of a number of people. Dr. Darrell Hull has been an amazing advisor, dissertation chair, and mentor to me throughout my doctoral degree. I know that I have benefitted greatly from his guidance and I endeavor to follow his example as I move forward in my career. I would also like to acknowledge my dissertation committee members Dr. Robin Henson, Dr. Endia Lindo, and Dr. Todd Kettler. Their support throughout this process was much appreciated, and I know my work is better due to the guidance they provided.

There are many other faculty, staff, and colleagues in the Department of Educational Psychology and across the University of North Texas I could acknowledge given the space. I would like to thank Dr. Abbas Tashakkori for his continued support and guidance as Department Chair, and Dr. Anne Rinn-McCann for her mentorship as Assistant Chair for Doctoral Mentorships. Additionally, my colleagues Marvin Powell and Katrina Hovey were a great help to me throughout data collection and the editing process, spending hours of their time to help me with this study.

Finally, I have to thank my family and friends for supporting me and putting up with me throughout this crazy process. Thank you Abe, Kathy, Jasmine, and Hannah for your patience and emotional support. Huong, Rachel, and TJ for keeping me sane and letting me talk about my research for hours. My extended family and friends for checking in on me and giving me a place to get away when I needed a break. This dissertation is a culmination of years of my life, but it is equally a testament to the amazing support and love of the people in my life. Thank you!
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>Stereotypical Science: Exploring High School Occupational Preferences for Science by Sex, Personality, and Cognitive Ability</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Method</td>
<td>14</td>
</tr>
<tr>
<td>Results</td>
<td>20</td>
</tr>
<tr>
<td>Discussion</td>
<td>32</td>
</tr>
<tr>
<td>Conclusion and Future Directions</td>
<td>37</td>
</tr>
<tr>
<td>References</td>
<td>38</td>
</tr>
<tr>
<td>Appendix A: Extended Literature Review</td>
<td>48</td>
</tr>
<tr>
<td>Appendix B: Extended Methods</td>
<td>64</td>
</tr>
<tr>
<td>Appendix C: Extended Results and Discussion</td>
<td>80</td>
</tr>
<tr>
<td>Appendix D: IPIP Personality Items</td>
<td>102</td>
</tr>
<tr>
<td>Appendix E: Demographic Survey</td>
<td>105</td>
</tr>
<tr>
<td>Combined Reference List</td>
<td>108</td>
</tr>
</tbody>
</table>
STEREOTYPICAL SCIENCE: EXPLORING HIGH SCHOOL OCCUPATIONAL PREFERENCES FOR SCIENCE BY SEX, PERSONALITY, AND COGNITIVE ABILITY

Introduction

Education in America over the last 30 years has seen a steady increase in focus on science, technology, engineering, and math (STEM) education. This increased focus is in response to the importance of STEM content mastery for career and college preparation (NAS, 2011; NRC, 2014) and the fact that science and engineering are key factors for economic growth in the U.S. (National Academy of Sciences, 2011). However, there remains a consistent sex difference in STEM with males much more likely to choose STEM careers and degrees than females (Hill, Corbett, & Rose, 2010; NSF, 2015). Women made up almost 50% of the college-educated workforce in 2010, but only 27.5% of the workforce in science and engineering fields (NSF, 2015). In biology-focused fields, women were better represented at 48.2%, but in engineering women made up only 12.7%, and only 25.1% in computer and mathematical sciences (NSF, 2015).

The problem of sex differences in regard to STEM career choice requires additional investigation with more detail and a focus on better understanding the source of sex differences, in addition to identifying ways in which the differences can be minimized (Valian, 2014). However, some have questioned the validity of the sex differences in STEM concern. Some critics of this topic have cited increases in women receiving STEM degrees and working in STEM fields, particularly in biological science and social science areas, as evidence the sex disparity in STEM careers is abating (need a citation for these “critics” here). Additionally, some have suggested the goal should not be equal representation, as there are many fields where men are not equally represented (Valian, 2014). This argument appears regularly in popular media
and online discussions of sex differences in science careers. The sex disparity in any area, STEM or otherwise, has been connected to a lack of interest by a specific sex (Hill, et al, 2010; NSF, 2015) and therefore may not be a true disparity based on sex. However, the issue is not truly the lack of women in STEM, although logically diversity in a field may help strengthen the output of that field. Rather, one concern is that women who might be interested in STEM are being encouraged away from the field due to either explicit or implicit sex stereotypes or discrimination (Hill, et al., 2010; Valian, 2014). If women are choosing other careers on their own there may not be a stereotype effect, but the research suggests women feel unwelcome in many STEM fields, particularly physics, engineering, and computer science areas (Deemer, Thoman, Chase, & Smith, 2014; Gardner, 1974; Hill, et al., 2010; Jones, Howe, & Rua, 2000; Kessels, 2015). The suggestion is that women interested in science are “forced” away from the field due to outside influence.

Theory of Circumscription and Compromise

Gottfredson’s (1981, 1996) theory of circumscription and compromise (C&C) was conceptualized to explain developmentally how individuals’ identify occupational preferences. Previous theories of career preference examined either personal psychological variables or environmental and social variables. Gottfredson (1981) argued “a better integration of the psychological and non-psychological approaches would provide a more comprehensive explanation of the development of vocational aspirations” (p. 546). Therefore, C&C theory integrates both psychological self-concept and social occupational images into the process of occupational preference identification.

The self-concept aspect of C&C theory includes environmental components such as sex and social class as well as psychological attributes such as personality, values, and intelligence
Self-concept has been repeatedly studied as a key component in education research (Marsh, 1990; Shavelson, Hubner, & Stanton, 1976). While the exact definition of self-concept varies in the literature, a general definition is the way in which an individual understands himself or herself (Shavelson, et al., 1976). In C&C theory, self-concept develops naturally throughout an individual’s lifetime. As individuals become aware of their unique attributes such as sex or cognitive ability, they form a sense of self that drives their interactions and experiences with the world around them. These experiences then feed back into their self-concept and continue the development process (Gottfredson, 1981; 1996; Marsh, 1990; Shavelson, et al., 1976).

C&C theory understands occupational images to be “a generalization a person makes about a particular occupation,” sometimes referred to in other literature as occupation stereotypes (Gottfredson, 1981, p. 547). C&C theory avoids the stereotype label due to the negative connotations attached to the term. Occupational images can include generalizations regarding the personality of those involved in a specific occupation, the type of work they do and the life these individuals have, what defines an individual considered to be a typical case of the occupation, etc. (Gottfredson, 1981; 1996). Finally, an occupational preference in C&C theory is understood as the occupations an individual prefers or wishes for in the future (Gottfredson, 1981; 1996). This does not reflect what occupations the individual might be best suited for in “reality” but rather the occupations the individual currently prefers in his or her own conception (Gottfredson, 1981).

Occupational preferences align with an individual’s self-concept such that the ideal occupation of an individual at any given developmental stage aligns with whom the individual understands himself or herself to be. This is the circumscription component of C&C theory.
However, occupational images may not align with the individual’s occupational preference as decided by self-concept alone. In this situation, the individual’s occupation preferences adjust to better align with both his or her self-concept and understanding of occupational images. This is the compromise component of C&C theory. These two components are described in more detail in the next section.

Circumscription and Compromise

C&C theory presents circumscription as the process of narrowing the field of occupational preferences based on developmentally appropriate decisions. It is defined as “the developmental process by which children and adolescents progressively eliminate from further consideration whole sections of the occupational world as incompatible with their developing self-concepts” (Gottfredson & Lapan, 1997). For instance, a key developmental stage in C&C theory as it applies to the present study is stage four where children orient their occupational preferences to their interests (Gottfredson, 1981; 1996). They therefore remove occupations that do not align with their interests or abilities from their mental occupational preferences list.

There are four stages in C&C theory and the individual circumscribes occupational preferences based on a different level of awareness (Gottfredson, 1981; 1996; Gottfredson & Lapan, 1997). In Stage 1, children become aware of the adult world and the concept of jobs. In Stage 2, children are aware of social factors such as sex roles, and they begin to orient their occupational preferences towards jobs seen as appropriate for their sex. Stage 3 expands social understanding into concepts of prestige and the value different professions hold in society. Individuals then orient their occupational preferences about jobs they perceive as being more prestigious within their perceived range of ability. Finally, as mentioned previously, it is in Stage
4 that adolescents narrow their occupational preferences based on interests and alignment with internal self-concept.

C&C theory understands compromise as the decisions made when occupational preferences are challenged due to social barriers or problems such as lack of job opportunities, difficulty in receiving training, etc. (Gottfredson, 1981; 1996; Gottfredson & Lapan, 1997). Circumscription focuses on the developmental process of narrowing occupational preferences towards a final occupational aspiration. However, the process of career selection is not limited to only perceived self-concept in the circumscription stages. C&C theory also accounts for environmental and external impacts on an individual’s occupational preferences, labeled compromise in this theory. If a person showed interest in a job that was not hiring in the area or where training was difficult to obtain, this person might compromise interest in this position and remove it from his or her occupation preferences list (Gottfredson, 1981). Typically, individuals make compromises in a pattern such that “vocational interests are sacrificed first, job level second, and sextype last” (Gottfredson, 1981, p. 549). Gottfredson (1996) revised compromise in C&C theory to include allowances for different levels of compromise (major, moderate, and minor). Gottfredson (1996) argues people compromise their self-concept at increasing levels as necessary in reaction to their circumstances. For instance, sextype is more integral to the individual’s self-concept and therefore more difficult to compromise, while a compromise based on interest might be less threatening (Gottfredson, 1996; Gottfredson & Lapan, 1997). Circumscription and compromise therefore have a coactive and reciprocal relationship in C&C theory to consolidate viable career choices as “circumscription involves identifying one’s most desirable options” while “compromise is the process by which individuals relinquish their most preferred futures” in response to external barriers (Gottfredson & Lapan, 1997. p. 426).
Self-Concept

As self-concept is a complex construct, multiple individual variables are needed to represent the larger concept (Marsh, 1990; Shavelson, et al., 1976). As the present study is particularly interested in occupational preferences, attention is given specifically to dimensions of self-concept most related to development of occupational interest. A wealth of literature exists on the role sex plays in individual self-concept and occupational preference, supporting the use of sex in C&C theory (Cross & Madson, 1997; Twenge, 1997; Wilgenbusch & Merrell, 1999). Sex is a known factor in the identification of occupational preferences according to C&C theory, and can influence both the circumscription and compromise processes (Gottfredson, 1981; 1996). In relation to science careers, males consistently demonstrate preference for science (Hill, Corbett, & Rose, 2010; NSF, 2015).

Socio-economic status is another well-established predictor of educational outcomes (see Sirin, 2005; White, 1982). Generally, low socio-economic status correlates with lower levels of academic achievement and reduced likelihood to continue on to college (Sirin, 2005). However, the relationship between socio-economic status and occupational preferences is mixed. Some studies have shown higher levels of science interest for students from lower socio-economic status backgrounds (Breakwell, 1992), particularly if they have increased levels of ambition and educational attainment (Croll, 2008). Alternatively, a study by Gorard and See (2009) found individuals were less likely to pursue science in college if they were in a lower socio-economic status group.

Cognitive ability has also been widely studied in application to self-concept and occupation preferences. Studies have shown differences in cognitive ability affect occupational preferences (see Ackerman & Beier, 2003). Specifically as it relates to science, increased spatial
ability has been shown to result in increased preference for science careers (Wai, Lubinski & Benbow, 2009). Research on spatial ability often addresses sex differences, as males tend to exhibit higher spatial ability (see Geary, Sault, Liu & Hoard, 2000; Jensen, 2006; Maeda & Yoon, 2013; Reilly & Neumann, 2013).

The relationship of personality constructs to occupational preference has been investigated to determine whether or not personality traits differentiate those that choose different occupations or vocations; or whether occupations suit those that possess similar personality trait constellations (see for example Hansen, 1994; Holland, 1973; and Strong, 1943). While Gottfredson (1981; 1996) specifically mentions personality in combination with self-concept, research utilizing C&C theory tends to ignore personality factors (see Blanchard & Lichtenberg, 2003; Cochran, et al., 2011; Gottfredson & Lapan, 1997). Perhaps the most widely accepted conceptualization of personality traits is the Big 5 Factor model of personality (Costa & McRae, 1992; McRae & Costa, 1987; Digman, 1990; Goldberg, 1990). In studies exploring the correlation of the Big 5 personality factors to college and career preferences for science, Openness, Conscientiousness, Introversion (negative Extraversion), and Agreeableness predicted science interest, though the effects were generally small (Ackerman & Heggestad, 1997; Feist, 2012; Hong & Lin, 2011). Meta-analytic results also support the prediction of occupational preference by personality, favoring Agreeableness as a predictor of science and math career paths (Ackerman & Heggestad, 1997).

**Occupational Image**

Science and math careers are generally considered masculine fields of work (Gottfredson, 1981, 1996, 2005; Gottfredson & Lapan, 1997; Kessels, 2015). Additionally, many consider the skills necessary in science; such as spatial reasoning ability, to be innate, and therefore the only
way to be successful in science is through natural talent or genetic proclivity as opposed to hard
work or effort (Jones, et al., 2000; Kessels, 2015). By contrast, female stereotypes portray
females as feminine with abilities such as increased verbal ability and psychological inclinations
such as nurturing, which do not align with science careers (Kessels, 2015). The perception
therefore exists that males are more suited to careers in science than are females (Jones, et al.,
2000). The differences between science stereotypes and female stereotypes create a perception of
misfit between the self-concept of female students and occupational interests in science.

For more than 40 years, research in science career interest, attitudes, and motivation has
resulted in the conclusion that course enrollment, spatial ability, attitudes and interest in science
are attributable to sex differences (Gardner, 1974). Prior to 1974, researchers were primarily
focused on the existence of a sex difference in science interest. Studies had concluded attitudes
and interest were more important than cognitive factors in predicting the sex difference in
science. Additionally, mention was made of innate factors such as personality and
environmental factors such as social pressure. However, the claim was presented that researchers
could not truly test the effects of these personal and environmental elements until females had
experienced equal education and employment opportunities in science for many generations
(Gardner, 1974). Further, a review of research up to 1974 suggested many studies of females
already enrolled in science courses in higher education utilized biased sampling, as they were not
evaluating females who did not choose to take science courses in college.

Attention has shifted away from the existence of sex differences in science interest, and
appears more or less presumed. Instead, interest lies in the sources of the sex differences and
possible solutions. Although some argue the sex gap has narrowed (Lane, Goh & Driver-Linn,
2011; Leaper, Farkas, & Brown, 2012), differences persist on a variety of self-concept variables
and infrequently debated in the literature. Sex differences in science achievement have been shown to increase throughout high school (Bacharach, Baumeister, & Furr, 2003; Legewie & DiPrete, 2011; Reid & Skryabina, 2003). This aligns with C&C theory as compromise increases in adolescence and circumscription to perceived sextype interests and occupations occurs during adolescence (Gottfredson, 1981; 1996). Also in alignment with self-concept components of C&C theory, some studies have combined sex differences research with ethnic group considerations. Bacharach, Baumesiter, and Furr (2003) found white students of both sexes score higher on average than black students. Additionally, male students within each racial group scored higher than female students in the same racial group. Riegle-Crumb, Moore, and Ramos-Wada (2010) evaluated TIMSS data on 8th grade science and math achievement and attitude. They found male attitudes towards science and math were largely consistent across racial groups, while achievement was lower for Hispanic and Black males. However, female attitudes towards science varied across ethnicities (Riegle-Crumb, et al., 2010). These studies speak to the complexity of sex differences in science and highlight the impact of social factors.

In accordance with C&C theory, social factors have consistently been identified as a source of sex differences in science (Leaper, et al., 2012; Legewie & DiPrete, 2011; Robnett & Leaper, 2012; Valian, 2014). Valian (2014) suggests a noticeable increase in women showing interest in choosing STEM careers will only happen when women feel comfortable and are assured of the likelihood of their success in the field of science. This social effect is often referred to as gender stereotype or gender-science stereotype, or sex-science stereotype. Stereotype effect has been found in international examinations and was present even in countries considered more egalitarian in terms of female roles (Miller, Eagly, & Linn, 2015; Nosek, et al., 2009). The sex-science stereotype appears most closely connected with male domination relative
to the proportion of males participating in science within a country, and in college (Miller, et al., 2015). Additionally, when personal stereotypes were evaluated they were not found to provide predictive validity beyond the average national stereotypes (Nosek, et al., 2009).

The relationship between sex-science stereotypes and sex differences in science attitude, motivation, interest, and achievement appears to be reciprocal. As sex-science stereotypes in a society suggest females will not be successful in science, females feel they will not be successful in science and therefore do not perform well in science (Nosek, et al., 2009). Other support exists for this reciprocal relationship (Cundiff, Vescio, Loken, & Lo, 2013; Deemer, et al., 2014; Hayes & Bigler, 2012; Jones, et al., 2000; Lane, et al., 2011). Findings suggest the perceived value of science is predictive of female satisfaction in science education, but not predictive of male satisfaction (Hayes & Bigler, 2012). Cundiff, et al. (2013) found female undergraduate participants had stronger stereotypes towards science than males, and these stereotypes related to weaker science identification and lower science career intentions. The opposite was found for male participants such that stronger sex-stereotypes were related to stronger identification and increased career intentions in science. In Lane, et al. (2011), sex differences in science intentions were completely explained by sex stereotypes. Additionally, sex identity moderated the relationship between female’s stereotypes and their career intentions, and was strongest for females identifying strongly as feminine. Studies have shown sex stereotypes are more prevalent in physical and engineering sciences, traditionally more male-dominated fields, and less so in biological and social sciences, which are seen as more accessible to females (Deemer, et al., 2014; Gardner, 1974; Jones, et al., 2000). The reciprocal relationship between social expectations and science occupation outcomes aligns directly with C&C theory in discussing how compromise influences occupational preference.
It is worth noting that one recent study did fail to identify a sex difference in science interest (Buday, Stake, & Peterson, 2012). However, the participants in this study were high ability students, and Buday, et al. (2012) argue the lack of sex difference could be related to the unique sample of students. Gottfredson (1981, 1996) predicted high ability and/or high social class students would be less sensitive to sex stereotype. As noted in the changes to C&C theory made by Gottfredson (1996), circumscription does not apply equally to all students. Those with high ability or increased access to science and/or career resources may not be sensitive to occupational images and stereotypes.

**Occupational Preference**

The outcome in C&C theory is the occupational preference and ultimately the occupational choice of an individual. Interest is a complex variable and research on the topic often includes mention of motivation and attitude in addition to or in combination with interest. Motivation is a widely studied concept in educational psychology and is typically understood as an internal force describing why an individual gives one response or chooses one behavior over another or more frequently than another (Pintrich, 2003; Braver, et al., 2014). Bandura’s definition of motivation is commonly referenced in motivation literature, defined as an internal state that arouses, directs, and sustains goal-oriented behavior (Glynn, Brickman, Armstrong, & Taasoobshirazi, 2011; see Bandura, 1986). Attitude is generally defined as containing three components (cognitive, affective, and behavioral) and addressing positive or negative orientations towards an object (Potvin & Hasni, 2014). Finally, definitions of interest in the literature are often vague and occasionally include attitude and motivation as sub-constructs (Potvin & Hasni, 2014). Interest is understood as a key factor in career decisions and focuses on
the strength of a typically positive relationship between the individual and an object, or career in this application (Potvin & Hasni, 2014).

Two systematic reviews of motivation, interest, and attitude towards science were located (Osborne, Simon, & Collins, 2003; Potvin & Hasni, 2014). Potvin and Hasni (2014) reviewed 12 years of research on interest, motivation and attitude towards science. A total of 233 articles were reviewed discussing some combination of motivation, interest, and attitudes related to science. They specifically identified 50 articles studying sex differences. Most of the studies reviewed found slight or non-significant sex differences in motivation, attitude, and interest in science in the general population. The sex differences found typically favored males. When specific science discipline was accounted for, the results were stronger, with physics, technology, and often chemistry preferred by males while females preferred biology (Potvin & Hasni, 2014). Sex differences were stronger as students grew older, and articles identified by Potvin and Hasni (2014) only sometimes presented social stereotypes as an explanation for sex differences.

Osborne, et al. (2003) reviewed over 20 years of literature on attitudes about science. Sex and quality of teaching were the most important factors in predicting student attitudes about science. Osborne, et al. (2003) concluded research over the last 20 years still supports females as having less positive attitudes on science, suggesting social norms and access to science opportunities might contribute to sex differences in attitude, and lack of understanding for cultural values of science might be a contributing factor to lack of female interest.

Overall, the literature of motivation, attitude, and interest in science has shown males in the general population tend to score higher on all three constructs (Osborne, et al., 2003; Potvin & Hasni, 2014; Simpkins, Davis-Kean, & Eccles, 2006; Simpson & Oliver, 1990). However, some studies have found motivation and achievement in science to be higher for females (Baker,
1985; Simpson & Oliver, 1990). Attitude, motivation, and interest in science decline as students get older, particularly for students with average ability (Baker, 1985; Potvin & Hasni, 2014; Simpson & Oliver, 1990). Social and personal variables influence the relationship between science achievement and attitude, motivation, and interest constructs in a complex relationship (Osborne, et al., 2003; Potvin & Hasni, 2014; Simpkins, et al., 2006; Simpson & Oliver, 1990; Singh, Granville, & Dika, 2002). School factors of instruction style and curriculum focus have been identified as important in supporting student motivation, attitude, and interest in science (Bathgate, Schunn, & Correnti, 2013; Potvin & Hasni, 2014).

**Purpose and Research Questions**

The present study is designed to deepen understanding of sex differences in science occupational preference as a result of individual’s self-concept and the occupational images they possess as a test of C&C theory (Gottfredson, 1981; 1996). For the present study, self-concept variables: sex, socio-economic status, cognitive ability, and personality are examined collectively to explore the function of these variables in the circumscription and compromise of science career preferences. Because C&C theory expresses a relationship between self-concept and occupational preference, the present study includes a combination of attitudes towards science, science motivation, interest in science careers, and science academic experiences to represent the occupational images and career preference of participants, (i.e. their proclivity for science careers).

The present study therefore explores occupational preferences as indicated by a latent combination of science attitude, motivation, interest, and achievement. The effects of the self-concept variables on occupational preferences are directly evaluated, and the theoretical impact of occupational images is discussed. There are two research questions:
1. Based on measures of science motivation, science attitude, science interest, and science achievement, do latent profiles exist that support underlying sex differences in occupational preference for science career choice?

Hypothesis 1: Distinct profiles will be seen such that males show higher representation in a profile with high science interest, high science motivation, high science attitude, and high science achievement. Females will likewise show higher representation in a profile with low science interest, low science motivation, low science attitude, and low science achievement. Hypothesis 2: To support compromise due to sex stereotypes in science, a profile will also be seen with high representation of females showing low science interest but high science motivation, high science attitude, and high science achievement.

2. What is the magnitude of group differences in occupational preference for science by self-concept facets: sex, socio-economic status, personality latent profile, and cognitive ability?

Hypothesis 3: Sex, socio-economic status, personality, and cognitive ability will all be statistically significant predictors of occupational preference for science.

Hypothesis 4: Sex will be the strongest predictor of occupational preference, followed by socio-economic status, cognitive ability and then personality profile.

Method

The present study consisted of 11th and 12th grade students at three public high schools in a district in the southwestern United States. The top two grades in high school were selected for three reasons. One, students in these grades are better able to self-report their science attitudes, motivations, and interests in a way consistent with adults. This supports the instrumentation in this study, which has largely been validated with adult and/or college samples. Two, the students
in the upper grades of high school are closer to making decisions regarding career and/or college preferences. Three, this age is supported as being developmentally appropriate for an exploration of occupational preference according to Gottfredson’s (1981, 1996) C&C theory.

**Participant Characteristics**

Upon approval by both the university Institutional Review Board and the independent school district, participants were solicited at three participating high schools. Responses from students were kept completely confidential and names were not attached to responses in the study. No limits were set for sampling based on sex or ethnicity. The study used a cluster-convenience sampling strategy, as participants were volunteers conveniently sampled within school and grade clusters (Johnson & Christensen, 2014).

The three participating high schools were all large public high schools in a Southwestern state in the United States. School student populations ranged from 2,157 to 2,404 for the 2015-2016 academic year. Ethnicity percentages, students classified as economically disadvantaged, and students noted as English language learners are somewhat varied between the different schools, as shown in Table 1. All three schools where relatively comparable on state standardized exams in science, with 91-93% of students on the 2014-2015 test scoring at least satisfactory, 57-73% scoring college ready, and 14-27% scoring at an advanced level of achievement.
Table 1

Demographic Information on Participating High Schools

<table>
<thead>
<tr>
<th></th>
<th>High School 1</th>
<th>High School 2</th>
<th>High School 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Population</td>
<td>2,157</td>
<td>2,395</td>
<td>2,404</td>
</tr>
<tr>
<td>Student Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>16%</td>
<td>9%</td>
<td>15%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>31%</td>
<td>22%</td>
<td>35%</td>
</tr>
<tr>
<td>White</td>
<td>47%</td>
<td>63%</td>
<td>47%</td>
</tr>
<tr>
<td>Asian</td>
<td>3%</td>
<td>3%</td>
<td>0%</td>
</tr>
<tr>
<td>Economically Disadvantaged</td>
<td>40%</td>
<td>21%</td>
<td>48%</td>
</tr>
<tr>
<td>English Language Learner</td>
<td>10%</td>
<td>5%</td>
<td>11%</td>
</tr>
</tbody>
</table>

Note. Data from state education reporting agencies and publicly available district information. Sources not cited to protect confidentiality of schools and students.

Sampling Procedures

Data was collected at the participating high schools either during school or during free periods as agreed upon with school administration. The testing process was identical in all schools and at all testing times. A script was used and read by the researcher to the participants with a brief introduction and instructions on completing the measures. The testing process started with the cognitive ability measure as this was the only timed test in the study. Then the other measures were administered and completed at the participant’s own pace. Finally, a demographic questionnaire to gather basic information on participants was included as the final page of the assessment package. There were four individual measures plus the demographic questionnaire in
In this study, resulting in approximately 180 items. Testing time did not exceed 50 minutes and typically took 30-45 minutes for students to complete all materials.

In order to encourage participation and repay the participants for their time, two incentive strategies were used. First, all participants received a $5 gift card to a local restaurant. Second, a drawing was held to randomly award one student from each participating school with a $50 gift card.

Sample Size and Power

Power considerations in LPA are complex and no method of calculating power is present in the literature. Required sample size is dependent on number of profiles identified and distance between the profiles, which is unknown in advance and can only be estimated based on the literature (Tein, et al., 2013). A simulation study on power in LPA found the median sample size in the literature to be $n = 377$, and in their simulation study they found sample size did not really influence power except in a few situations (Tein, et al., 2013). In the present study, the total sample was $n = 375$ with between 100 and 130 students at each school. However, 80 of these students were in 9th or 10th grade and were removed from the dataset for the latent profile analyses for a functional $n = 295$.

Measures and Covariates

Assessments in the present study measured self-concept (cognitive ability, personality), occupational preference (motivation, attitude), as well as science academic experiences (science courses taken, overall GPA). Demographic information such as age, grade, sex, ethnicity, socio-economic status, and college and/or career plans were collected in a demographic questionnaire.

Science Academic Experiences
To assess science academic experiences in high school, participants were asked to self-report information about courses they have taken. Items focused on the number of science courses taken in high school and overall high school GPA. This combination represents science academics without necessitating a science content exam. The differences in high school science requirements and course choices make a science content exam too complicated for this study.

Cognitive Ability

This study used the Shipley-2 Vocabulary and Block Patterns scales (Shipley, Gruber, Martin, & Klein, 2009). The Vocabulary and Block Patterns scales combine to represent a composite measure of general cognitive ability. Vocabulary contains 40 items and takes 10-15 minutes while Block Patterns contains 26 items and is administered in 12 minutes. Standardization studies for this measure resulted in median reliability estimates of $\alpha = .84$ for the Vocabulary scores and $\alpha = .85$ for the Block Patterns scores (Shipley, et al., 2009).

Personality

To assess personality for this study, items from the International Personality Item Pool (IPIP) were selected. This is a freely available pool of items based on well-supported personality models. For this study, a pre-developed 50-item measure of the Big 5 constructs was selected: Neuroticism, Extraversion, Openness to New Experiences, Agreeableness, and Conscientiousness. These items follow the format of the NEO-PI measure (McCrae & Costa, 1987; Costa & McCrae, 1992). For scores on each subscale of the IPIP measure, reliability estimates are typically seen in the range of $\alpha = .78-.86$. Personality was evaluated using a higher-order composite created using LPA as described in Merz and Roesch (2011). This results in three personality profiles: Reserved, or individual’s with relatively low scores on all five factors, Well-
Adjusted, individual’s with low Neuroticism but medium to high levels of the other four factors, and Excitable, those individual’s with high levels of all five factors.

Science Motivation

To assess science motivation of participants, the choice was made to use the Science Motivation Questionnaire II (SMQ-II, Glynn, Brickman, Armstrong, & Taasoobshirazi, 2011). This scale assesses five aspects of motivation: Intrinsic Motivation, Career Motivation, Self-Determination, Self-Efficacy, and Grade Motivation. The multiple facets of this scale are particularly beneficial for this study as they touch on multiple aspects of motivation in science. The SMQ-II is freely available for research purposes and contains 25 items. The development study for the second edition of the SMQ reported reliability estimates of $\alpha = .81-.92$ for scores on the measure’s subscales (Glynn, et al., 2011).

Attitudes Toward Science

In addition to motivation, attitudes toward science were assessed using the Attitude Toward Science Scale (Vitale & Johnson, 1988). This scale assesses four factors: Instrumental Value of Science, Active Participation in Science, Difficulty and Complexities of Science, and General Attitude Toward School. The instrument has 31 total items and is freely available for use in research studies. Reliability estimates from the development study for this assessment ranged from $\alpha = .55-.99$, with the fourth factor representing the General Attitude Toward School showing low reliability at $\alpha = .55$ (Vitale & Johnson, 1988).

Demographic Questionnaire

Student age, grade, socio-economic status as represented by mother’s and father’s level of education and occupation, number of science courses completed, and GPA were all collected using a single-page questionnaire. Additionally, a question on science interest was included
which asked students “Thinking about your plans for your future career, at this point how interested are you in a career in a science field?”. Responses were coded 1 for “Very Interested”, 2 for “Somewhat Interested”, 3 for “Somewhat Not Interested”, and 4 for “Very Not Interested”.

Results

*Analytic Approach*

To address the research questions in the present study, two separate analyses were conducted. Initially, descriptive statistics were used to evaluate responses and check for normality. Then, Latent Profile Analysis (LPA) was used as the inferential analysis method (Bauer & Curran, 2004; Bergman & Magnusson, 1997; Bergman, et al., 2003; Marsh, Lüdtke, Trautwein, & Morin, 2009; Sterba, 2013). LPA is used in two applications: first in exploring latent profiles of science occupational preference, and second as covariates are added to examine group differences in the latent profiles. As LPA is a model testing process, multiple models are fit with varying levels of classes or profiles, typically 1-5 dependent on topic (see Tein, Coxe, & Cham, 2013). Each model is then compared against the previous model or models to make a decision regarding the number of latent profiles in the data (Marsh, et al., 2009).

Model retention decisions in the present study are based on Bayesian Information Criterion (BIC), Sample-Adjusted BIC (SABIC), and Akaike’s Information Criterion (AIC) (Celeux & Soromenho, 1996; Marsh, et al., 2009; Tein, et al., 2013). Additionally, the Lo, Mendell, and Rubin (LMR) test is used to compare models, in a similar fashion to the $\chi^2$ difference test in other model testing analyses (Lo, Mendell, & Rubin, 2001; Marsh, et al., 2009; Tein, et al., 2013). Finally, as in any model testing analysis, theoretical support must exist for the final model retained, and the patterns and profiles uncovered must be interpretable (Marsh, Hau, & Wen, 2004; Marsh, et al., 2009).
For the covariate analysis step, there are two possible methods. If computationally possible, the best method of covariate inclusion in LPA is to include the covariates in the LPA model itself (Clark & Muthen, 2009; Marsh, et al., 2009). This method can be computationally complex and time consuming, but is the most robust method in simulation (Clark & Muthen, 2009). Alternatively, the participants can be assigned to the profile for which they have the highest probability and subsequent group comparison analyses can be conducted (Marsh, et al., 2009). In the present study, the full covariate model converged, allowing for a robust covariate analysis.

**Descriptive Statistics**

Table 2 provides means, standard deviations, value ranges, skewness and kurtosis information, and internal consistency reliability estimates for all of the continuous scales and variables. Reliability was evaluated for the scores on each scale using Cronbach’s alpha, looking for a minimum value of .70 for measures in development and used for research purposes, and .80 for previously validated clinical measures (Henson, 2001). For categorical or nominal variables, modes and frequencies of responses are provided in Table 3.
Table 2

Descriptive Statistics for Continuous Study Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATS_AP</td>
<td>22.34</td>
<td>5.63</td>
<td>7-35</td>
<td>-.23</td>
<td>.14</td>
<td>.81</td>
</tr>
<tr>
<td>ATS_DC</td>
<td>21.38</td>
<td>4.62</td>
<td>3-38</td>
<td>-.09</td>
<td>1.29</td>
<td>.62</td>
</tr>
<tr>
<td>ATS_GS</td>
<td>9.94</td>
<td>2.78</td>
<td>3-15</td>
<td>-.08</td>
<td>-.40</td>
<td>.74</td>
</tr>
<tr>
<td>ATS_IV</td>
<td>51.18</td>
<td>8.90</td>
<td>13-70</td>
<td>-.61</td>
<td>1.46</td>
<td>.88</td>
</tr>
<tr>
<td>Shipley-BP</td>
<td>16.70</td>
<td>5.20</td>
<td>0-26</td>
<td>-.51</td>
<td>.51</td>
<td>.88</td>
</tr>
<tr>
<td>Shipley-Vocab</td>
<td>16.83</td>
<td>4.81</td>
<td>7-39</td>
<td>-.39</td>
<td>.37</td>
<td>.78</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>29.25</td>
<td>9.11</td>
<td>7-50</td>
<td>.05</td>
<td>-.77</td>
<td>.90</td>
</tr>
<tr>
<td>Extraversion</td>
<td>33.79</td>
<td>8.27</td>
<td>10-50</td>
<td>-.28</td>
<td>-.35</td>
<td>.89</td>
</tr>
<tr>
<td>Openness</td>
<td>36.92</td>
<td>6.32</td>
<td>17-50</td>
<td>-.44</td>
<td>.12</td>
<td>.77</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>36.34</td>
<td>6.42</td>
<td>10-50</td>
<td>-.68</td>
<td>.84</td>
<td>.79</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>35.45</td>
<td>6.39</td>
<td>18-50</td>
<td>-.30</td>
<td>-.26</td>
<td>.82</td>
</tr>
<tr>
<td>SMQ_CM</td>
<td>15.32</td>
<td>6.11</td>
<td>5-25</td>
<td>.18</td>
<td>-1.13</td>
<td>.93</td>
</tr>
<tr>
<td>SMQ_GM</td>
<td>18.20</td>
<td>4.69</td>
<td>5-25</td>
<td>-.47</td>
<td>-.25</td>
<td>.86</td>
</tr>
<tr>
<td>SMQ_IM</td>
<td>16.58</td>
<td>4.60</td>
<td>5-25</td>
<td>-.23</td>
<td>-.54</td>
<td>.87</td>
</tr>
<tr>
<td>SMQ_SD</td>
<td>14.95</td>
<td>4.42</td>
<td>5-25</td>
<td>.02</td>
<td>-.27</td>
<td>.85</td>
</tr>
<tr>
<td>SMQ_SE</td>
<td>18.02</td>
<td>4.35</td>
<td>5-25</td>
<td>-.44</td>
<td>-.25</td>
<td>.86</td>
</tr>
<tr>
<td>Mother’s SES</td>
<td>29.88</td>
<td>15.20</td>
<td>1-52</td>
<td>-.52</td>
<td>-.93</td>
<td></td>
</tr>
<tr>
<td>Father’s SES</td>
<td>31.94</td>
<td>14.72</td>
<td>1-55</td>
<td>-.79</td>
<td>-.51</td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>3.41</td>
<td>.70</td>
<td>1.6-5.0</td>
<td>.24</td>
<td>-.09</td>
<td></td>
</tr>
</tbody>
</table>

Note. ATS = Attitude Towards Science, SMQ = Science Motivation Questionnaire
Table 3

Descriptive Statistics for Nominal and Categorical Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Response</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science Career Interest</td>
<td>Very Interested</td>
<td>21.8%</td>
</tr>
<tr>
<td></td>
<td>Somewhat Interested</td>
<td>31.9%</td>
</tr>
<tr>
<td></td>
<td>Somewhat Not Interested</td>
<td>20.0%</td>
</tr>
<tr>
<td></td>
<td>Very Not Interested</td>
<td>26.3%</td>
</tr>
<tr>
<td>Sex</td>
<td>Male</td>
<td>43.5%</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>56.5%</td>
</tr>
<tr>
<td>Grade</td>
<td>11th</td>
<td>26.9%</td>
</tr>
<tr>
<td></td>
<td>12th</td>
<td>73.1%</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>African-American</td>
<td>11.3%</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>3.2%</td>
</tr>
<tr>
<td></td>
<td>Caucasian</td>
<td>51.4%</td>
</tr>
<tr>
<td></td>
<td>Hispanic</td>
<td>29.2%</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>4.9%</td>
</tr>
</tbody>
</table>

Data Cleaning

Data cleaning and assumptions checks were preformed prior to beginning the LPA. First, data was evaluated for values out of range. For instance, if a GPA was recorded above 5.0, which is the highest possible value, the false value was removed from the dataset. These errors are generally keystroke or response errors and can falsely skew data in analyses (Osborne, 2012). Five values out of range were removed from the dataset.

Next, assumptions were checked to evaluate the univariate and multivariate normality of the variables to be included in the analysis. While normality is not necessarily an assumption of LPA, it is a common assumption in inferential statistics and was checked to better support statistical conclusion validity of the present study. Supporting the multivariate normality of the variables in the LPA also supports external validity, suggesting these results are replicable and represent a complex construct. Univariate normality was examined using the skewness and
kurtosis values as shown in Table 2, with a conservative benchmark of values between -1.00 and 1.00 representing normally distributed data (Osborne, 2012). Kurtosis on two subscales of the Attitude Towards Science measure and one subscale of the Science Motivation measure were slightly outside of this range. However, due to the robust nature of the analysis being conducted, these values were deemed acceptable without transformation. Multivariate normality was evaluated using Mahalanobis distance as described in Henson (1999). A Q-Q plot of Mahalanobis distance with an associated $\chi^2$ value was evaluated, and the values follow a linear trend as expected to support multivariate normality (see Figure 1).

Figure 1. Q-Q plot for variables in LPA to support multivariate normality

Finally, missing data was handled in the present study using maximum likelihood estimation concurrently with the analysis process in Mplus (Osborne, 2012). Cases were removed from analysis if all variables involved in the analysis were missing. Missing data in the
present study was minimal, with most variables having no more than 5% missing values. Parent education and occupation questions did have missingness as high as 28%.

**Latent Profile Analysis**

To address Research Question 1, a LPA was conducted to explore latent profiles of individuals on the measures of science motivation, science attitude, science interest, and science achievement. Figure 2 shows the path diagram of this model. All latent profile analyses were conducted using Mplus 6 with a maximum likelihood estimation method. Composite variables were used as opposed to item-level data to simplify the model and support convergence.

![Figure 2. Latent profile analysis Model 1 for Research Question 1](image)

Model fit statistics are provided in Table 4. Model 1 was estimated with only one profile, Model 2 with two profiles, and so on to Model 5 with five profiles. Model 4 was retained as the
best model to fit the data based on the lowest loglikelihood value, lower AIC, BIC, and SABIC values, a high entropy value, results of the Lo-Mendel Ruben test, and profiles that are supported by theory. The hypothesized three-profile model was not directly supported in this analysis, though the retained four-profile model may be an expansion of the hypothesized structure.

Table 4

LPA Model Fit Summary for Research Question 1

<table>
<thead>
<tr>
<th>Model</th>
<th>Loglikelihood</th>
<th>AIC</th>
<th>BIC</th>
<th>SABIC</th>
<th>Entropy</th>
<th>Smallest Class %</th>
<th>LMR p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-8909.94</td>
<td>17867.88</td>
<td>17955.96</td>
<td>17879.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-8507.38</td>
<td>17088.75</td>
<td>17224.54</td>
<td>17107.21</td>
<td>0.89</td>
<td>45%</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>3</td>
<td>-8370.30</td>
<td>16840.60</td>
<td>17024.10</td>
<td>16865.53</td>
<td>0.90</td>
<td>30%</td>
<td>0.002</td>
</tr>
<tr>
<td>4</td>
<td>-8290.15</td>
<td>16706.31</td>
<td>16937.51</td>
<td>16737.73</td>
<td>0.89</td>
<td>15%</td>
<td>0.010</td>
</tr>
<tr>
<td>5</td>
<td>-8253.86</td>
<td>16659.72</td>
<td>16938.63</td>
<td>16697.62</td>
<td>0.90</td>
<td>2%</td>
<td>0.614</td>
</tr>
</tbody>
</table>

Note. The Lo-Mendell Ruben (LMR) test compares the current model to a model with \( k-1 \) profiles.

The four-profile model that was retained is detailed in Table 5. The means and standard deviations of the variables used to create the profiles are presented for each latent profile. Note that the standard deviations are the same as they are constrained in the analysis by default. The differences between the four latent groups are largely due to differences in interest, motivation, and attitude towards science, which supports the theoretical approach used in the present study. Profile 1 contains students with the lowest level of interest in science, a more negative attitude and low motivation towards science, and a low GPA, on average. Profile 2 contains those students with the highest interest in science, high attitude and motivation scores, and the highest
average GPA. Profile 3 students are low in terms of science interest, but more towards the 
middle in science attitude and motivation, with a GPA almost as high on average as seen in 
Profile 2. Finally, Profile 4 contains students who are somewhat interested in science, have mid-
level attitudes and motivations towards science, and a GPA lower than Profile 3 but higher than 
Profile 1. Profile 1 could be referred to as “Lowest Science Interest and GPA”, Profile 2 as 
“Highest Science Interest and GPA”, Profile 3 as “Low Science Interest, High GPA”, and Profile 
4 as “Medium Science Interest and GPA”.
Table 5

*Retained 4-Profile LPA Model Results*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Profile 1</th>
<th>Profile 2</th>
<th>Profile 3</th>
<th>Profile 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lowest</td>
<td>Highest</td>
<td>Low/High</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>$M$ (SD)</td>
<td>$M$ (SD)</td>
<td>$M$ (SD)</td>
<td>$M$ (SD)</td>
</tr>
<tr>
<td>Number of Classes</td>
<td>3.74 (0.88)</td>
<td>4.31 (0.88)</td>
<td>3.78 (0.88)</td>
<td>3.71 (0.88)</td>
</tr>
<tr>
<td>GPA</td>
<td>2.95 (0.63)</td>
<td>3.77 (0.63)</td>
<td>3.59 (0.63)</td>
<td>3.11 (0.63)</td>
</tr>
<tr>
<td>Science Interest</td>
<td>3.65 (0.52)</td>
<td>1.32 (0.52)</td>
<td>3.49 (0.52)</td>
<td>2.00 (0.52)</td>
</tr>
<tr>
<td>SMQ – Intrinsic Motivation</td>
<td>10.31 (3.05)</td>
<td>21.03 (3.05)</td>
<td>15.58 (3.05)</td>
<td>16.91 (3.05)</td>
</tr>
<tr>
<td>SMQ – Career Motivation</td>
<td>8.18 (2.84)</td>
<td>23.06 (2.84)</td>
<td>11.27 (2.84)</td>
<td>16.27 (2.84)</td>
</tr>
<tr>
<td>SMQ – Self-Determination</td>
<td>9.19 (3.00)</td>
<td>19.37 (3.00)</td>
<td>14.50 (3.00)</td>
<td>14.43 (3.00)</td>
</tr>
<tr>
<td>SMQ – Self-Efficacy</td>
<td>12.63 (3.22)</td>
<td>21.75 (3.22)</td>
<td>18.23 (3.22)</td>
<td>17.26 (3.22)</td>
</tr>
<tr>
<td>SMQ – Grade Motivation</td>
<td>12.74 (3.70)</td>
<td>21.80 (3.70)</td>
<td>18.37 (3.70)</td>
<td>17.62 (3.70)</td>
</tr>
<tr>
<td>ATS – Instrumental Value</td>
<td>41.36 (6.75)</td>
<td>59.05 (6.75)</td>
<td>48.80 (6.75)</td>
<td>51.67 (6.75)</td>
</tr>
<tr>
<td>ATS – Academic</td>
<td>19.20 (5.39)</td>
<td>23.89 (5.39)</td>
<td>21.73 (5.39)</td>
<td>23.24 (5.39)</td>
</tr>
<tr>
<td>ATS – Difficulties &amp; Complexities</td>
<td>20.15 (4.52)</td>
<td>21.25 (4.52)</td>
<td>20.88 (4.52)</td>
<td>22.75 (4.52)</td>
</tr>
<tr>
<td>ATS – General School</td>
<td>11.23 (2.71)</td>
<td>9.433 (2.71)</td>
<td>9.54 (2.71)</td>
<td>10.16 (2.71)</td>
</tr>
</tbody>
</table>

*Note.* SMQ = Science Motivation Questionnaire, ATS = Attitude Towards Science

For the second research question, the profiles identified in research question one were further analyzed to evaluate the effects of covariates. This analysis was conducted in combination with the original LPA by adding the covariates to be modeled concurrently as is shown in Figure 3 and as recommended in the literature (Clark & Muthen, 2009; Marsh, et al., 2009). The covariates of interest were: personality; cognitive ability, both verbal and spatial; sex;
and socio-economic status, as defined by parent education level and occupation. The results of this analysis are presented in Table 6.

Figure 3. Latent profile analysis Model 2 with covariates for Research Question 2
Table 6

Covariate Analysis Results for the 4-Profile Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Profile 1 Lowest Interest &amp; GPA</th>
<th>Profile 3 Low Interest, High GPA</th>
<th>Profile 4 Medium Interest &amp; GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>-0.96</td>
<td>-0.32</td>
<td>-1.373*</td>
</tr>
<tr>
<td>Shipley-2 Vocabulary</td>
<td>-0.19*</td>
<td>-0.10</td>
<td>-0.23*</td>
</tr>
<tr>
<td>Shipley-2 Block Patterns</td>
<td>-0.07</td>
<td>-0.002</td>
<td>-0.05</td>
</tr>
<tr>
<td>Mother’s SES</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.003</td>
</tr>
<tr>
<td>Father’s SES</td>
<td>-0.02</td>
<td>0.01</td>
<td>-0.002</td>
</tr>
<tr>
<td>Personality</td>
<td>-1.08*</td>
<td>-0.12</td>
<td>-0.556</td>
</tr>
</tbody>
</table>

*Note. *=p<.05, Profile 2 (Highest Science Interest & GPA) served as reference group.

Based on this analysis, some of the covariates are statistically significantly different between the profiles. Results suggest sex was statistically significantly different between Profile 2 (Highest Science Interest & GPA) and Profile 4 (Medium Science Interest & GPA), with more female’s falling into the high science interest profile. Vocabulary scores were also statistically significantly different between Profile 2 (Highest Science Interest & GPA) and both Profile 1 (Lowest Science Interest & GPA) and Profile 4 (Medium Science Interest & GPA). The negative coefficients show the high science interest group tended to have higher vocabulary scores than these other two profiles. Finally, personality was shown to be statistically significantly different for Profile 1 (Lowest Science Interest & GPA) as compared to Profile 2 (Highest Science Interest & GPA). A cross-tabulation shows the difference is between students considered “Well-Adjusted,” as significantly more of them fall into the high science interest profile (see Table 7).
Table 7

Cross-Tabulation of Statistically Significant Covariate Differences

<table>
<thead>
<tr>
<th>Profile</th>
<th>Reserved</th>
<th>Excitable</th>
<th>Well-Adjusted</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Lowest Science Interest &amp; GPA</td>
<td>3</td>
<td>26</td>
<td>16</td>
<td>21(17%)</td>
<td>21(13%)</td>
</tr>
<tr>
<td>2 Highest Science Interest &amp; GPA</td>
<td>2</td>
<td>27</td>
<td>49</td>
<td>24(20%)</td>
<td>51(32%)</td>
</tr>
<tr>
<td>3 Low Science Interest, High GPA</td>
<td>1</td>
<td>44</td>
<td>44</td>
<td>35(28%)</td>
<td>54(34%)</td>
</tr>
<tr>
<td>4 Medium Science Interest &amp; GPA</td>
<td>1</td>
<td>47</td>
<td>30</td>
<td>43(35%)</td>
<td>34(21%)</td>
</tr>
</tbody>
</table>
Discussion

The present study provides support for elements of C&C theory to explore the development of science interest. For research question one, latent profiles were identified and show a pattern of students with varying levels of interest in science, collectively understood to include attitude and motivation, as well as differences in GPA. Hypothesis one regarding the predominance of males in the high science interest profile and females in the low science interest profile was not supported. However, hypothesis two regarding the existence of a profile made up largely of females with a low science interest and a high motivation and attitude towards science was supported and is represented by Profile 3.

Profile 1, named “Lowest Science Interest & GPA,” is made up of students with the lowest level of science interest on average across almost all measures. Additionally, this profile had the lowest average GPA of all four profiles. This was the smallest profile in the present study, and is made up of males and females equally. Of the 42 students in this group, 26 are considered Excitable, making up about 62% of the profile. Students in this profile were found to be statistically significantly different from those in the “Highest Science Interest & GPA” profile based on personality profile as well as vocabulary ability. This group of students is not interested in science, holds less than a 3.00 average in GPA, has a low vocabulary ability, and is generally Excitable in personality.

Profile 2 is the “Highest Science Interest & GPA” profile, and is characterized by students with the highest interest, attitude, and motivation towards science. These are students that believe in the value of science, are confident in their abilities in science, and see science as a strong possibility for their future career. Additionally, these students have the highest GPA on average of all of the profiles. The profile contains primarily female students, making up 68% of
the profile and containing 32% of the females in the present study. Students in this profile are also primarily considered Well-Adjusted in terms of personality, and have significantly higher vocabulary scores than those in the “Lowest Science Interest & GPA” profile.

Profile 3 was named “Low Science Interest & High GPA”, as the science interest scores for this profile are almost as low as in Profile 1 but the GPA average is the second highest. These are students that do well in school, as seen in the GPA average and relatively high scores on the Grade Motivation and Academic Attitude scales. However, subscales and questions specifically related to science interest are lower than Profile 2 and often lower than Profile 4, while still being higher than Profile 1. This suggests these students are not very interested in science, other than as necessary for their academic goals. Profile 3 is the largest profile in the present study with 89 students, and it is primarily made up of female students ($n = 54$). This is also the largest percentage of female students for any of the four profiles (34%), and most closely matches the hypothesized profile of students highly motivated and positive about science but not interested in science as a career. In terms of personality, Well-Adjusted and Excitable students equally characterize this profile, and personality was not a statistically significant factor. No aspects of self-concept evaluated in the present study are statistically significantly different for this profile as compared to the “Highest Science Interest & GPA” profile.

Finally, Profile 4 is the “Medium Science Interest & GPA” profile. This profile is made up of students with scores around the middle of the responses in the present study. The GPA average is 3.11, and answers on the science interest questions and scales were typically higher than Profile 1 but lower than Profiles 2 and 3. These students are somewhat ambivalent towards school and their career decisions, and science in particular. This is the second largest profile with 77 total students, and is the only profile with more males than females (males $n = 43$). There are
also more males in this profile than in any other, with 35% of the males in the present study falling into this group. Profile 4 was statistically significantly different from the “Highest Science Interest & GPA” profile (Profile 2) in terms of sex and vocabulary scores. Specifically, Profile 4 contains more males and averages lower vocabulary scores when compared to Profile 2. This profile also contains the largest group of Excitable students in terms of personality profiles.

For research question two, the effects of the self-concept covariates on the identified latent profiles were mixed. Theoretically, sex was expected to affect the latent profiles such that males were more prevalent in the high science interest profiles with females more prevalent in low science interest profiles (Hill et al., 2010; NSF, 2015). This disparity has been connected to sex stereotype affect in previous work (Leaper, et al., 2012; Legewie & DiPrete, 2011; Robnett & Leaper, 2012; Valian, 2014). This was somewhat supported in the present study, as the “Low Science Interest and High GPA” profile (Profile 3) contains the highest percentage of females (34%). However, the second highest female representation is in the “Highest Science Interest and GPA” profile (Profile 2), suggesting science stereotype may not be a strong influence on high school females’ science interest development. The statistically significant difference between the “Highest Science Interest & GPA” profile and the “Medium Science Interest & GPA” profile, with more females present in the high science interest group, further weakens support for sex stereotype in the present study. Additionally, the largest male presence was in the “Medium Science & GPA” profile, as opposed to the “Highest Science Interest & GPA” profile as might be expected. These results are contrary to expectations and previous literature, though consideration could be given to the possibility of a publication bias.

One possible explanation for the lack of strong sex stereotype affect in the present study may be related to the observed personality profile differences. Results from previous studies on
the impact of socio-economic status of career decisions found that students with a lower socio-
-economic status can overcome the perceived barrier to science careers their status implies,
particularly if they have high levels of ambition and do well in school (Breakwell, 1992; Croll,
2008). This same logic may apply here, where female students with Well-Adjusted personality
types are able to overcome perceived sex stereotype more effectively. An alternative explanation
could be related to cognitive ability, as previous studies have also shown students with higher
cognitive ability and educational opportunities are less susceptible to sex stereotype (Buday, et
al., 2012). Both of these explanations align with C&C theory as the sextype compromise is the
hardest to overcome, and would require a stronger self-concept (Gottfredson, 1981; 1996).

The Shipley-2 Vocabulary scores were also found to be statistically significantly different
in the covariate analysis. Students in the “Highest Science Interest & GPA” profile had
significantly higher vocabulary scores as compared to those in the “Lowest Science Interest &
GPA” and “Medium Science Interest & GPA” profiles. Vocabulary ability is not as frequently
discussed as spatial ability in terms of science interest prediction (Gardner, 1974; Wai, et al.,
2009). The spatial ability component of cognitive ability was expected to show differences
between profiles of science interest based on the literature (Geary, et al., 2000; Jenson, 2006;
Maeda & Yoon, 2013; Reilly & Neumann, 2013). However, this was not supported in the present
study. It is possible the difference is related to higher cognitive ability generally as opposed to
specifically vocabulary ability, which has previously resulted in more equal sex distributions
related to science interest (Buday, et al., 2012). However, as the spatial ability measure was not
statistically significantly different in the analysis it is unclear what the source of this difference
might be.
Personality profile was found to be statistically significantly different for students in the “Lowest Science Interest & GPA” profile as compared to students in the “Highest Science Interest & GPA” profile. Specifically, those students who are considered Well-Adjusted are more likely to be interested in science and have a higher GPA. Previous studies of Big 5 personality factors and science interest typically examined the subscales only, while the present study took a higher-order personality profile approach. However, previous studies have found Openness to New Experiences, Introversion, Agreeableness, and Conscientiousness all predict science interest (Ackerman & Heggestad, 1997; Feist, 2012; Hong & Lin, 2011), which aligns well with the Well-Adjusted profile characteristics (Merz & Roesch, 2011). Integrating the five factors into a complex personality profile adds understanding to relationships between personality and other variables such as science interest in the present study.

Limitations

The present study benefits from a large sample of 11th and 12th grade high school students. However, this sample is representative of a large public school district in the Southwestern United States and may not be applicable to students in other areas of the country or in other school types. Future studies could account for more detail related to the types of science courses taken and the grades students earned in each course. An argument can be made for the difference between an “A” grade in a required on-level course and an “A” grade in an advanced science elective taken at the AP level, but this is not accounted for in this study. Finally, the literature suggests sex differences are more pronounced in some areas of science such as physics, engineering, and computer science as opposed to biology and social sciences. These different science preferences are not captured in the present study, and including this may present a clearer understanding of sex stereotype.
Conclusion and Future Directions

Overall, the present study provides an empirical evaluation of the complex construct of interest in science at the high school level. The use of LPA as an advanced modeling strategy strengthens the present study and provides a more detailed explanation of interest and the variables affecting interest. Gottfredson’s (1981, 1996) C&C theory is partially supported. The LPA person-oriented analysis method models the development of interest with four distinct profiles, but the hypothesized impact of self-concept covariates is mixed. Sex, cognitive ability, and personality were found to have an impact on science interest, though not always as hypothesized.

The results of the present study can be used in practice to suggest high school interventions and encourage interest in science. For the present study, students who were identified as highly interested in science were generally academically advanced, Well-Adjusted, and showed high vocabulary ability. For students that do not fit these characteristics, additional interventions could be considered. Students who do well academically but are not as stable in personality may need to be introduced to science careers in a different way, one that appeals to their Excitable personality profile. It is possible these students do not see their personality as a good fit for science careers. While work in a laboratory or in front of a computer may not be appealing to this type of student, they may be interested in other careers in science such as those involving working with people or in the field. Additionally, increasing science interest for students who are not as advanced in academics or who show lower vocabulary ability would need to be addressed in other ways. For instance, introducing these students to aspects of science that fit their aptitudes better may help them see how science careers are an option for them as well.
Future work could replicate this study with other high school populations to verify the current findings. Additionally, due to the lack of clear evidence to support the sex stereotype effect and the impact of self-concept on interest development, this study could be conducted with college students and science professionals. Replication of this study with older individuals can be used to identify where sex differences are occurring in science careers. Additionally, the lack of clarity regarding the possible effect of sex stereotype justifies the use of a mixed-methods approach. Mixed-methodology would allow researchers to replicate the quantitative study conducted here and add a qualitative component to evaluate student perspectives on sex stereotype and the impact it may or may not have had on their interest in science.

References


Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: evidence for overlapping traits. *Psychological Bulletin, 121*, 219-245. doi: 10.1037/0033-2909.121.2.219


Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing


APPENDIX A

EXTENDED LITERATURE REVIEW
Introduction

Education in America over the last 30 years has seen a steady increase in focus on science, technology, engineering, and math (STEM) education. Research supports the importance of STEM content mastery for career and college preparation (NAS, 2011; NRC, 2014). A National Academy of Sciences (NAS) report in 2011 concluded science and engineering will be key factors in economic growth in the US. However, there remains a consistent sex difference in STEM with males much more likely to choose STEM careers and degrees than females (Hill, Corbett, & Rose, 2010; NSF, 2015). Women made up almost 50% of the college-educated workforce in 2010, but only 27.5% of the workforce in science and engineering fields (NSF, 2015). In biology-focused fields, women were better represented at 48.2%, but in engineering women made up only 12.7%, and only 25.1% in computer and mathematical sciences (NSF, 2015).

There is a call for the research community to study the problem of sex differences in more detail with a focus on better understanding the source of sex differences in science and identifying ways in which the differences can be minimized (Valian, 2014). However, some have questioned the validity of the sex differences in STEM concern. Some critics of this topic have cited increases in women receiving STEM degrees and working in STEM fields, particularly in biological science and social science areas, as evidence the sex disparity in STEM careers is going away. Additionally, some have suggested the goal should not be equal representation, as there are many fields where men are not equally represented (Valian, 2014). This argument appears regularly in popular media and online discussions of sex differences in science careers. The sex disparity in any area, STEM or otherwise, has been connected to a lack of interest by a specific sex and therefore may not be a true disparity based on sex. However, the issue is not
truly the lack of women in STEM, although diversity in a field can only help strengthen the output of that field. Rather, the true concern is that women who might be interested in STEM are being encouraged away from the field due to either explicit or implicit sex stereotypes or discrimination (Hill, et al., 2010; Valian, 2014). If women are truly choosing other careers on their own there is not an issue, but the research suggests women feel unwelcome in many STEM fields, particularly physics, engineering, and computer science areas (Deemer, Thoman, Chase, & Smith, 2014; Gardner, 1974; Hill, et al., 2010; Jones, Howe, & Rua, 2000; Kessels, 2015). This therefore creates a sex disparity that is a concern, where women who are interested in science are “forced” away from the field due to outside influence.

The present study therefore focuses on providing more depth of understanding regarding factors at work within the observed sex disparity in science. Instead of focusing on whether or not there is a sex difference, as the literature already supports this, the present study will focus on the development of interest in science and possible movement away from this interest due to stereotype or discrimination. Deeper understanding of the complexity of the sex differences issue is sought by applying an established theoretical lens through Gottfredson’s (1981; 1996) circumscription and compromise theory. This theory specifically addresses the process of occupational interest and accounts for the outside influences of social pressure and stereotype.

Theory of Circumscription and Compromise

To provide a foundation for understanding in the current study, Gottfredson’s (1981, 1996) theory of circumscription and compromise (C&C) has been selected. C&C theory was conceptualized to explain developmentally how individuals’ identify their occupational preferences. Previous theories of career preference had examined either personal psychological variables or environmental and social variables. Gottfredson (1981) argued “a better integration
of the psychological and non-psychological approaches would provide a more comprehensive explanation of the development of vocational aspirations” (p. 546). Therefore, C&C theory integrates both psychological self-concept and social occupational images into the process of occupational preference identification.

The self-concept aspect of C&C theory includes environmental components such as sex and social class as well as psychological attributes such as personality, interests, values, and intelligence (Gottfredson, 1981, 1996, 2005; Gottfredson & Lapan, 1997). C&C theory advocates for a combined understanding of self-concept with both individual characteristics and environmental facets represented. C&C theory understands occupational images as “a generalization a person makes about a particular occupation,” sometimes referred to in other literature as occupation stereotypes (Gottfredson, 1981, p. 547). C&C theory avoids the stereotype label due to the negative connotations attached to the term. Occupational images can include generalizations regarding the personality of those involved in a specific occupation, the type of work they do and the life these individuals have, what defines an individual considered to be a typical case of the occupation, etc. (Gottfredson, 1981; 1996). Finally, an occupational preference in C&C theory is understood as the occupations an individual prefers or wishes for in the future (Gottfredson, 1981; 1996). This does not reflect what occupations the individual might be best suited for in “reality” but rather the occupations the individual currently prefers in his or her own conception (Gottfredson, 1981).

In C&C theory, self-concept develops naturally throughout an individual’s lifetime. Occupational preferences align with an individual’s self-concept such that the ideal occupation of an individual at any given developmental stage aligns with whom the individual understands himself or herself to be. This is the circumscription piece of C&C theory. However, occupational
images may not align with the individual’s occupational preference as decided by self-concept alone. In this situation, the individual’s occupation preferences adjust to better align with both his or her self-concept and understanding of occupational images. This is the compromise component of C&C theory. These two components are described in more detail in the next section.

Circumscription and Compromise

C&C theory presents circumscription as the process of narrowing the field of occupational preferences based on developmentally appropriate decisions. It is defined as “the developmental process by which youngsters progressively eliminate from further consideration whole sections of the occupational world as incompatible with their developing self-concepts” (Gottfredson & Lapan, 1997). For instance, a key developmental stage in C&C theory as it applies to the present study is stage four where children orient their occupational preferences to their interests (Gottfredson, 1981; 1996). They therefore remove occupations that do not align with their interests or abilities from their mental occupational preferences list.

There are four stages in C&C theory and the individual circumscribes occupational preferences based on a different level of awareness (Gottfredson, 1981; 1996; Gottfredson & Lapan, 1997). In stage one, children become aware of the adult world and the concept of jobs. In stage two, children are aware of social factors such as sex roles, and they begin to orient their occupational preferences towards jobs seen as appropriate for their sex. Stage three expands social understanding into concepts of prestige and the value different professions hold in society. Individuals then orient their occupational preferences about jobs they perceive as being more prestigious within their perceived range of ability. Finally, as mentioned previously, it is in stage
four that adolescents narrow their occupational preferences based on interests and alignment with internal self-concept.

C&C theory understands compromise as the decisions made when occupational preferences are challenged due to social barriers or problems such as lack of job opportunities, difficulty in receiving training, etc. (Gottfredson, 1981; 1996; Gottfredson & Lapan, 1997). Circumscription focuses on the developmental process of narrowing occupational preferences towards a final occupational aspiration. However, the process of career selection is not limited to only perceived self-concept in the circumscription stages. C&C theory also accounts for environmental and external impacts on an individual’s occupational preferences, labeled compromise in this theory. If a person showed interest in a job that was not hiring in the area or where training was difficult to obtain, this person might compromise interest in this position and remove it from his or her occupation preferences list (Gottfredson, 1981). Typically, individuals make compromises in a pattern such that “vocational interests are sacrificed first, job level second, and sextype last” (Gottfredson, 1981, p. 549). Gottfredson (1996) revised compromise in C&C theory to include allowances for different levels of compromise (major, moderate, and minor). Gottfredson (1996) argues people compromise their self-concept at increasing levels as necessary in reaction to their circumstances. The more personal elements of self-concept such as sextype are more difficult for an individual to give up, while a compromise based on amount of personal interest in a field might be less threatening (Gottfredson, 1996; Gottfredson & Lapan, 1997). Circumscription and compromise therefore work together in C&C theory as “circumscription involves identifying one’s most desirable options” while “compromise is the process by which individuals relinquish their most preferred futures” in response to external barriers (Gottfredson & Lapan, 1997, p. 426).
Previous Applications of C&C Theory

Two studies have applied C&C theory to high school occupational preferences and the impact of sex stereotypes on career decision-making. Cochran, Wang, Stevenson, Johnson, and Crews (2011) examined career aspirations in high school compared to career success in midlife. They found C&C theory fit the data well. Specifically, parents’ SES, ability, and sex predicted career aspirations directly, and indirectly influenced subsequent career success. Cochran, et al. (2011) also argued girls had less career success than boys.

In a more targeted application of C&C theory, Blanchard and Lichtenberg (2003) investigated compromise and the order of importance individuals placed on prestige, interests, and sex type. Results suggested that individuals who exhibit higher levels of compromise (i.e. those that must relinquish their occupational preferences due to external pressure) placed equal importance on prestige and sex type, while individuals who exhibit lower levels of compromise (i.e. those that do not need to relinquish much of their occupational preferences) focused on interests first and then prestige and then sex type. This suggests that for individuals whose occupational preferences do not require high levels of compromise, interests play a more important role while sex type is a minor consideration. By contrast, occupational preferences that do require high levels of compromise appear to weight prestige and sex type equally. In this higher compromise situation, personal interests are less important while social expectations and sex type are important considerations.

Self-Concept – Sex, Cognitive Ability, and Personality

In the present study, sex, socio-economic status, cognitive ability, and personality variables represent the self-concept component of C&C theory. Sex is identified as male or female to avoid the complexity of gender conceptions. A wealth of literature exists on the role
sex plays in individual self-concept and occupational preference, supporting the use of sex in C&C theory (Cross & Madson, 1997; Twenge, 1997; Wilgenbusch & Merrell, 1999). Sex is a key factor in identifying occupational preferences and influences both circumscription and compromise for individuals (Gottfredson, 1981; 1996). In relation to science careers, sex differences have been seen with males consistently showing a higher preference for science careers than females (Hill, Corbett, & Rose, 2010; NSF, 2015).

Socio-economic status is a well-established predictor of educational outcomes (see Sirin, 2005; White, 1982). Generally, students from lower socio-economic status families average lower levels of academic achievement and are less likely to continue on to college (Sirin, 2005). The relationship between socio-economic status and occupational preferences is mixed. Some studies have shown higher levels of science interest for students from lower socio-economic status backgrounds (Breakwell, 1992), particularly if they have increased levels of ambition and educational attainment (Croll, 2008). Alternatively, a study by Gorard and See (2009) found individuals were less likely to pursue science at the college level if they were in a lower socio-economic status group. More information on this relationship is needed.

Cognitive ability has also been widely studied in application to self-concept and occupation preferences. In the current study, the Cattell-Horn-Carroll (CHC) theory conceptions of cognitive ability are used (Carroll, 1993). Studies have shown differences in cognitive ability affect occupational preferences (see Ackerman & Beier, 2003). Specifically as it relates to science, an increased spatial ability has been shown to result in increased preferences for science careers (Wai, Lubinski & Benbow, 2009). Research on spatial ability often addresses sex differences, as males tend to exhibit higher spatial ability (see Geary, Sault, Liu & Hoard, 2000; Jensen, 2006; Maeda & Yoon, 2013; Reilly & Neumann, 2013).
Personality

More attention is given in the present study to the inclusion of personality as a factor in C&C theory. There is a long history of the application of personality constructs to occupational preferences including research by Holland (1973) and Strong (1943; see also Hansen, 1994). While Gottfredson (1981; 1996) specifically mentions personality in combination with self-concept, research utilizing C&C theory tends to ignore personality factors (see Blanchard & Lichtenberg, 2003; Cochran, et al., 2011; Gottfredson & Lapan, 1997).

One common theory of personality is the Big 5 factors of personality. The Big 5 was developed through a series of factor structure investigations starting with Raymond B. Cattell in the 1940s (Digman, 1990; Goldberg, 1990). Multiple investigations and revisions were conducted to develop a set of items to measure personality effectively, resulting most notably in the NEO-Five Factor Inventory of McRae and Costa in 1987. Although other theories of personality exist in literature, the Big 5 is an accepted theory of personality in the field and is supported by extensive construct validity evidence (Costa & McRae, 1992; McRae & Costa, 1987; Digman, 1990; Goldberg, 1990). Studies have previously applied the Big 5 personality factors to college and career preferences for science. Personality factors Openness, Conscientiousness, Introversion (negative Extraversion), and Agreeableness predicted science interest, though the effects were generally small (Ackerman & Heggestad, 1997; Feist, 2012; Hong & Lin, 2011). Meta-analytic results also support the prediction of occupational preference by personality, favoring Agreeableness as a predictor of science and math career paths (Ackerman & Heggestad, 1997). It is of note that studies of the impact of personality on science interest, motivation, and/or attitude do not often evaluate sex differences (Feist, 2012) or address sex as a covariate instead of a component of the primary analysis (Hong & Lin, 2011).
Outside of the specific science career focus, personality is a known predictor of academic achievement and motivation in school. Meta-analyses of these studies suggest consistent prediction of academic achievement and motivation by Big 5 factors of Neuroticism negatively and Conscientiousness positively (Judge & Ilies, 2002; Richardson, Abraham, & Bond, 2012). Effect sizes in these meta-analytic studies were moderate to large with absolute values between $r = .19$ and $r = .31$.

*Occupational Image - Sex Stereotypes in Science*

The primary focus in the present study is on the problem of sex differences in science careers. Science and math careers are generally considered masculine fields of work (Gottfredson, 1981, 1996, 2005; Gottfredson & Lapan, 1997; Kessels, 2015). Additionally, many consider the skills necessary in science such as spatial reasoning ability to be innate, and therefore the only way to be successful in science is through natural talent as opposed to hard work or effort (Jones, et al., 2000; Kessels, 2015). By contrast, female stereotypes portray females as feminine with abilities such as increased verbal ability and psychological inclinations such as nurturing, which do not align with science careers (Kessels, 2015). The perception therefore exists that males are more suited to careers in science than are females (Jones, et al., 2000). The differences between science stereotypes and female stereotypes create a perception of misfit between the self-concept of female students and occupational interests in science fields.

Sex differences in science career interest, attitudes, and motivation have been studied for more than 40 years. Gardner (1974) conducted a systematic review of the literature on sex differences in science at the time. In his report, Gardner concluded sex differences did exist as evidenced by differences in science course enrollment in universities. His review associated these differences with cognitive differences between males and females, as females showed
lower spatial ability in research. Additionally, attitudes and interest towards science were higher for males in Gardner’s (1974) review. However, the difference was smaller in high school studies. Additionally, it was noted that studies of females already enrolled in science courses in higher education utilize biased sampling. Gardner concluded attitudes and interest were more important than cognitive factors in understanding the sex difference in science. Additionally, mention is made of innate factors such as personality and environmental factors such as social pressure. However, the claim was presented that researchers cannot truly test the effects of these personal and environmental elements until females have experienced equal education and employment opportunities in science for many generations (Gardner, 1974).

Recently, the focus has shifted away from a discussion of the existence of sex differences in science. Instead, researchers are interested in the sources of the sex difference and possible solutions. Although some authors argue the sex gap has narrowed (Lane, Goh & Driver-Linn, 2011; Leaper, Farkas, & Brown, 2012), the difference is still present and not debated in the literature. Sex differences in science achievement have been shown to increase throughout high school (Bacharach, Baumeister, & Furr, 2003; Legewie & DiPrete, 2011; Reid & Skryabina, 2003). This aligns with C&C theory as compromise increases in adolescence and circumscription to sextype happens at this stage (Gottfredson, 1981; 1996). Also in alignment with self-concept components of C&C theory, some studies have combined sex differences research with ethnic group considerations. Bacharach, Baumesiter, and Furr (2003) found white students of both sexes score higher on average than black students. Additionally, the male students within each racial group scored higher than the female students in the same racial group did. Riegle-Crumb, Moore, and Ramos-Wada (2010) evaluated TIMSS data on 8th grade science and math achievement and attitude. They found male attitudes towards science and math were largely
consistent across racial groups, while achievement was lower for Hispanic and Black males. However, female attitudes towards science varied across ethnicities (Riegle-Crumb, et al., 2010). These studies speak to the complexity of sex differences in science and highlight the impact of social factors.

In accordance with C&C theory, social factors have consistently been identified as a source of the sex difference in science (Leaper, et al., 2012; Legewie & DiPrete, 2011; Robnett & Leaper, 2012; Valian, 2014). Valian (2014) suggests a noticeable increase in women showing interest in choosing STEM careers will only happen when women feel comfortable and are assured of the likelihood of their success in the field of science. This social effect is often referred to as gender stereotype or gender-science stereotype, or sex-science stereotype as in the current study. This stereotype effect has been found in international examinations and was present even in countries considered more egalitarian in terms of female roles (Miller, Eagly, & Linn, 2015; Nosek, et al., 2009). The sex-science stereotype in one study was found to be most closely connected to male domination of the science field within a country, and stronger among participants with a college education (Miller, et al., 2015). Additionally, another study suggested personal stereotypes do not provide predictive validity beyond national stereotypes (Nosek, et al., 2009).

The relationship between sex-science stereotypes and sex differences in science attitudes, motivations, interests, and achievement appears to be reciprocal. As sex-science stereotypes in a society suggest females will not be successful in science, females feel they will not be successful in science and therefore do not perform well in science (Nosek, et al., 2009). Other support exists for this reciprocal relationship (Cundiff, Vescio, Loken, & Lo, 2013; Deemer, et al., 2014; Hayes & Bigler, 2012; Jones, et al., 2000; Lane, et al., 2011). Findings suggest the perceived value of
science is predictive of female satisfaction in science education, but not predictive of male satisfaction (Hayes & Bigler, 2012). Cundiff, et al. (2013) found female undergraduate participants had stronger stereotypes towards science than males, and these stereotypes related to weaker science identification and lower science career intentions. The opposite was found for male participants such that stronger sex-stereotypes were related to stronger identification and increased career intentions in science. In Lane, et al. (2011), sex differences in science intentions were completely explained by sex stereotypes. Additionally, sex identity moderated the relationship between female’s stereotypes and their career intentions, and was strongest for females identifying strongly as feminine. Studies have shown sex stereotypes are more prevalent in the physical and engineering sciences, traditionally more male-dominated fields, and less so in biological and social sciences, which are seen as more accessible to females (Deemer, et al., 2014; Gardner, 1974; Jones, et al., 2000). The reciprocal relationship between social expectations and science occupation outcomes aligns directly with C&C theory in discussing how compromise influences occupational preference.

It is worth noting that one recent study did fail to identify a sex difference in science interest (Buday, Stake, & Peterson, 2012). However, the participants in this study were high ability students, and Buday, et al. (2012) argue the lack of sex difference could be related to the unique sample of students. Gottfredson (1981, 1996) predicted high ability and/or high social class students would be less sensitive to sex stereotype. As noted in the changes to C&C theory made by Gottfredson (1996), circumscription does not apply equally to all students. Those with high ability or increased access to science and/or career resources may not be sensitive to occupational images and stereotypes.
Future research directions on the topic of sex-science stereotype have been suggested. The presence of sex differences in science is no longer a topic worthy of study on its own (Potvin & Hasni, 2014; Sonnert & Fox, 2012; Valian, 2014). Instead, calls have been made for a more detailed study of sex difference in science at the personal level, institutional level, and in relation to how science is taught (Potvin & Hasni, 2014; Sonnert & Fox, 2012).

*Occupational Preference – Motivation, Attitude, and Interest*

The current study uses the constructs of motivation, attitude, and interest to explore occupational preference. Motivation is a widely studied concept in educational psychology and is typically understood as an internal force describing why an individual gives one response or chooses one behavior over another or more frequently than another (Pintrich, 2003; Braver, et al., 2014). Bandura’s definition of motivation is commonly referenced in motivation literature, and he defined the construct as an internal state that arouses, directs, and sustains goal-oriented behavior (Glynn, Brickman, Armstrong, & Taasoobshirazi, 2011; see Bandura, 1986). Attitude is generally defined as containing three components (cognitive, affective, and behavioral) and addressing positive or negative orientations towards an object (Potvin & Hasni, 2014). Finally, definitions of interest in the literature are often vague and occasionally include attitude and motivation as sub-constructs (Potvin & Hasni, 2014). Interest is understood as a key factor in career decisions and focuses on the strength of a typically positive relationship between the individual and an object (Potvin & Hasni, 2014).

Two systematic reviews of motivation, interest, and attitude towards science were located (Osborne, Simon, & Collins, 2003; Potvin & Hasni, 2014). Potvin and Hasni (2014) reviewed 12 years of research on interest, motivation and attitude towards science. A total of 233 articles were reviewed discussing some combination of motivation, interest, and attitudes related to
science. They specifically identified 50 articles studying sex differences. Most of the studies they reviewed found slight or non-significant sex differences in motivation, attitude, and interest in science in the general population. The sex differences found typically favored males. When specific science discipline was accounted for, the results were stronger, with physics, technology, and often chemistry preferred by males while females preferred biology (Potvin & Hasni, 2014). Sex differences were stronger as students grew older, and articles identified by Potvin and Hasni (2014) only sometimes presented social stereotypes as an explanation for sex differences.

Osborne, et al. (2003) reviewed over 20 years of literature on attitudes about science. Their review found sex and quality of teaching to be the most important factors in predicting student attitudes about science. Osborne, et al. (2003) concluded research over the last 20 years still supports females as having less positive attitudes on science. They suggested social norms and access to science opportunities might contribute to sex differences in attitude, and lack of understanding for cultural values of science might be a contributing factor to lack of female interest.

Overall, the literature of motivation, attitude, and interest in science has shown males in the general population tend to have higher levels of all three constructs (Osborne, et al., 2003; Potvin & Hasni, 2014; Simpkins, Davis-Kean, & Eccles, 2006; Simpson & Oliver, 1990). However, some studies have found motivation and achievement in science to be higher for females (Baker, 1985; Simpson & Oliver, 1990). Attitude, motivation, and interest in science decline as students get older, particularly for students at an average ability level (Baker, 1985; Potvin & Hasni, 2014; Simpson & Oliver, 1990). Social and personal variables influence the relationship between science achievement and attitude, motivation, and interest constructs in a complex relationship (Osborne, et al., 2003; Potvin & Hasni, 2014; Simpkins, et al., 2006;
Simpson & Oliver, 1990; Singh, Granville, & Dika, 2002). School factors of instruction style and curriculum focus have been identified as important in supporting student motivation, attitude, and interest in science (Bathgate, Schunn, & Correnti, 2013; Potvin & Hasni, 2014).
APPENDIX B

EXTENDED METHODS
Methodological Overview

The current study utilized a non-experimental quantitative research design. Non-experimental designs lack internal validity for strong inferences and causal conclusions (Johnson & Christensen, 2014; Shadish, Cook, & Campbell, 2002). However, established theory and/or longitudinal data collection can improve these studies (Johnson & Christensen, 2014; Shadish, et al., 2002). This study utilized the theory testing approach to strengthen the explanatory validity. Although still lacking the validity to support causal claims, support from the existing C&C theory strengthens the conclusions in the study and provides a stronger contribution to the literature. Additionally, advanced statistical analysis strategies allow for modeling of theoretical effects. In the present study, latent profile analysis (LPA) is used to explore the latent combination of occupational preferences variables and personality factors to more completely represent the effects of interest in C&C theory.

Analysis Method Overview

Latent profile analysis (LPA) was used as the inferential analysis method for this study. LPA is used in two applications: first in exploring latent profiles of science occupational preference, and second in combining personality factors into a set of latent personality profiles for further analysis. Covariates are then added to examine group differences in the latent profiles uncovered through the LPA process.

Latent Profile Analysis

LPA is a person-oriented approach to factor analysis in the same family of methods as latent class analysis, cluster analysis, mixture modeling, and others (Bergman & Magnusson, 1997; Bergman, Magnusson, & El-Khouri, 2003; Collins & Lanza, 2010; Gibson, 1959; Sterba, 2013). Like other factor analysis methods such as common factor analysis (CFA), LPA uses
covariance matrices to explore relationships between observed data and uncover latent groupings (Bauer & Curran, 2004; Bergman & Magnusson, 1997; Bergman, et al., 2003; Marsh, Lüdtke, Trautwein, & Morin, 2009). However, where CFA uses a covariance matrix of items to uncover latent factors, LPA uses a matrix of individuals to uncover latent groups of people. The main difference is “the common factor model decomposes the covariances to highlight relationships among variables, whereas the latent profile model decomposes the covariances to highlight relationships among individuals” (Bauer & Curran, 2004, p. 6). The person-oriented approach is founded on assumptions that individual differences are present and important within an effect or phenomenon, these differences occur in a logical way which can be examined through patterns, and a small number of patterns (profiles in LPA) are meaningful and occur across individuals (Bergman & Magnusson, 1997; Bergman, et al., 2003; Sterba, 2013).

The overall goal of a LPA is to uncover latent profiles or groups \((k)\) of individuals \((i)\) who share a meaningful and interpretable pattern of responses on the measures of interest \((j)\) (Bergman, et al., 2003; Marsh, et al., 2009; Sterba, 2013). This is done using joint and marginal probabilities in within-class and between-class models. Two equations define the within-class model:

\[
y_{ij} = \mu_{j}^{(k)} + \varepsilon_{ij} \tag{1}
\]

\[
\varepsilon_{ij} \sim N(0, \sigma_{j}^{2(k)}) \tag{2}
\]

where \(\mu_{j}^{(k)}\) is the model implied mean and \(\sigma_{j}^{2(k)}\) is the model implied variance, which will vary across \(j = 1\ldots J\) outcomes and \(k = 1\ldots K\) classes or profiles. The assumptions of LPA include the assumption the outcome variables are normally distributed within each class, and the within-class outcomes are locally independent (Sterba, 2013). The between-class model represents the probability of membership in a given class \(k\):
where $\omega^{(k)}$ is a multinomial intercept and $c_i$ is the latent classification variable for the individual. The within-class and between-class models can therefore be combined into a single model using the law of total probability resulting in:

$$f(y_i) = \sum_{k=1}^{K} p(c_i = k) f(y_i | c_i = k)$$

which is the marginal probability density function for an individual ($i$) after summing across the joint within-class density probabilities for the $J$ outcome variables, weighted by the probability for class or profile membership from equation 3. Finally, the LPA analysis results in a posterior probability for each individual defined as:

$$p(c_i = k | y_i) = \frac{p(c_i = k) f(y_i | c_i = k)}{f_{y_i}},$$

representing the probability of an individual ($i$) being assigned membership ($c_i$) in a specific class or profile ($k$) given their scores on the outcome variables in the $y_i$ vector. A posterior probability is calculated for each individual in each profile, with values closer to 1.0 indicating higher probability of membership in a specific profile. The more distinction between the posterior probabilities for an individual, the more certainty there is around their membership assignment (Sterba, 2013).

As LPA is a model testing process, multiple models are fit with varying levels of classes or profiles, typically 1-5 dependent on topic (see Tein, Coxe, & Cham, 2013). Each model is then compared against the previous model or models to make a decision regarding the number of latent profiles in the data (Marsh, et al., 2009). Commonly, decisions regarding model retention in LPA use Bayesian Information Criterion (BIC), Sample-Adjusted BIC (SABIC), and Akaike’s Information Criterion (AIC) (Celeux & Soromenho, 1996; Marsh, et al., 2009; Tein, et al., 2013).
BIC is used for model selection decisions with a lower BIC value representing the preferred model:

$$BIC(K) = -2L(K) + v(K) \ln n$$  \hspace{1cm} (6)$$

with $v(K)$ representing the number of parameters to be estimated in the model. The BIC prefers parsimony in a model but can be too conservative. An alternative is the SABIC, which adjusts the formula to account for $n$ and is less constraining on the number of parameters in the model (Tein, et al., 2013):

$$SABIC(K) = -2L(K) + v(K) \ln n * ((n + 2)/24)$$  \hspace{1cm} (7)$$

Finally, AIC is the least consistent of the model fit measures but is freer on the parsimony constraints. AIC is calculated as:

$$AIC(K) = -2L(K) + 2v(K)$$  \hspace{1cm} (8)$$

Additionally, the Lo, Mendell, and Rubin (LMR) test is typically used to compare models, in a similar fashion to the $\chi^2$ difference test in other model testing analyses (Lo, Mendell, & Rubin, 2001; Marsh, et al., 2009; Tein, et al., 2013). LMR uses the likelihood ratio of one model as compared to another with an adjusted asymptotic distribution instead of a $\chi^2$ distribution (Lo, et al., 2001). Finally, as in any model testing analysis, theoretical support should exist for the final model retained and the patterns and profiles uncovered should be interpretable (Marsh, et al., 2009).

A common follow up to LPA is the examination of covariates to discover relationships and differences between the latent groups (Clark & Muthen, 2009; Marsh, et al., 2009). Exploring relationships with covariates provides additional information on the latent profiles and how the effects of variables on these profiles might differ. One key point here, as highlighted in Marsh, et al. (2009), is the “covariates are assumed to be strictly antecedent variables” and
should have no effect on the formation of the profiles themselves (p. 195). This allows for some freedom in the method of covariate inclusion, as these covariates should not influence the LPA.

There are two recommended methods of covariate inclusion in the literature. First, if computationally possible, the best method of covariate inclusion in LPA is to include the covariates in the LPA model itself (Clark & Muthen, 2009; Marsh, et al., 2009). This method can be computationally complex and time consuming, but is the most robust method in simulation (Clark & Muthen, 2009). If this is not possible, the next option is to examine the profiles based on each individual’s most likely membership and explore the relationship of the profiles with the covariates. Using this method, you would assign each individual to the profile for which they have the highest posterior probability, and then analyze the relationship between the profile groups and covariates. However, this option is only recommended if entropy, or the level of classification uncertainty in the model, is above 0.80 (Clark & Muthen, 2009). Entropy is less common as a model retention index due to lack of support in simulation studies (see Tein, et al., 2013). However, entropy as a measure of classification uncertainty can still be useful in supporting the need for LPA in the data as opposed to a linear analysis. Entropy is calculated as:

\[ E(K) = \sum_{k=1}^{K} \sum_{l=1}^{n} t_{lk} \ln t_{lk} \]  

where \( t_{lk} \) represents the posterior probabilities as shown in equation 5. Equation 9 is therefore a measure of how well the LPA model in question partitions the data into profiles (Celeux & Soromenho, 1996).

Method

Upon approval by both the UNT Institutional Review Board and Denton independent school district, participants were solicited at participating schools through email and mailed notices, as well as through personal appearances at the schools. Parent consent and student assent
was collected for all students under 18 years-of-age, and all students 18 years and up completed an adult consent form. Responses from students were kept completely confidential and their name was not attached to their responses in the study.

Participants

This study assessed 11th and 12th grade students at three public high schools in a district in the Southwestern United States. The top two grades in high school were selected for three reasons. One, students in these grades are better able to self-report their science attitudes, motivations, and interests in a way consistent with adults. This supports the instrumentation in this study, which has largely been validated with adult and/or college samples. Two, the students in the upper grades of high school are closer to making decisions regarding career and/or college preferences. Three, this age is supported as being developmentally appropriate for an exploration of occupational preference according to Gottfredson’s (1981, 1996) C&C theory.

No limits were set for sampling based on sex or ethnicity. The study used a cluster-convenience sampling strategy, as participants were volunteers conveniently sampled within school and grade clusters (Johnson & Christensen, 2014). Typically, when setting sample size the researcher is concerned with adequately powering the statistical analyses to discover statistically significant results, should they be present. However, power considerations in LPA are complex and no method of calculating power is present in the literature. Required sample size is dependent on number of profiles identified and distance between the profiles, which is unknown in advance and can only be estimated based on the literature (Tein, et al., 2013). A simulation study on power in LPA found the median sample size in the literature to be $n = 377$, and in their simulation study they found sample size did not really influence power except in a few situations (Tein, et al., 2013). In the present study, the total sample was $n=375$ with between
100 and 130 students at each school. However, 80 of these students were in 9th or 10th grade and were removed from the dataset for the latent profile analyses for a functional \( n = 295 \).

The three participating high schools were all large public high schools in a Southwestern state in the United States. School student populations ranged from 2,157 to 2,404 for the 2015-2016 academic year. Ethnicity percentages, students classified as economically disadvantaged, and students noted as English language learners are somewhat varied between the different schools, as shown in Table B.1. Additionally, all three schools where relatively comparable on state standardized exams in science, with 91-93% of students on the 2014-2015 test scoring at least satisfactory, 57-73% scoring college ready, and 14-27% scoring at an advanced level of achievement.
Table B.1

Demographic Information on Participating High Schools

<table>
<thead>
<tr>
<th></th>
<th>High School 1</th>
<th>High School 2</th>
<th>High School 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student Population</td>
<td>2,157</td>
<td>2,395</td>
<td>2,404</td>
</tr>
<tr>
<td>African-American</td>
<td>16%</td>
<td>9%</td>
<td>15%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>31%</td>
<td>22%</td>
<td>35%</td>
</tr>
<tr>
<td>White</td>
<td>47%</td>
<td>63%</td>
<td>47%</td>
</tr>
<tr>
<td>Asian</td>
<td>3%</td>
<td>3%</td>
<td>0%</td>
</tr>
<tr>
<td>Economically Disadvantaged</td>
<td>40%</td>
<td>21%</td>
<td>48%</td>
</tr>
<tr>
<td>English Language Learner</td>
<td>10%</td>
<td>5%</td>
<td>11%</td>
</tr>
</tbody>
</table>

Note. Data from state education reporting agencies and publicly available district information. Sources not cited to protect confidentiality of schools and students.

Procedure

Data was collected at the participating high schools either during school or during free periods as agreed upon with school administration. Rooms met specific requirements such as appropriate lighting, seating and writing space for all testers, temperature controlled, low noise level, limited distractions, etc. This supported testing conditions for the participants and allowed for control of outside variables in the testing process at each school. Measures were paper-and-pencil to allow more participants to complete the testing process and limit the need for computer access.
The testing process was identical in all schools and at all testing times. A script was used and read by the researcher to the participants with a brief introduction and instructions on completing the measures. The testing process started with the cognitive ability measure as this was the only timed test in the study. Then the other measures were administered and completed at the participant’s own pace. Finally, a demographic questionnaire to gather basic information on participants was included as the final page of the assessment package. There were four individual measures plus the demographic questionnaire in this study, resulting in approximately 180 items. Testing time did not exceed 50 minutes and typically took 30-45 minutes for students to complete all materials.

In order to encourage participation and repay the participants for their time, two incentive strategies were used. First, all participants received a $5 gift card to a local restaurant. Second, a drawing was held to randomly award one student from each participating school with a $50 gift card.

**Measures**

Assessments in this study measured self-concept (cognitive ability, personality), occupational preference (motivation, attitude), as well as science academic experiences (science courses taken, overall GPA). Demographic information such as age, grade, sex, ethnicity, socio-economic status, and college and/or career plans were collected in the demographic questionnaire (see Appendix D).

Science Academic Experiences

To assess science academic experiences in high school, participants were asked to self-report information about courses they have taken. Items focused on the number of science courses taken in high school and overall high school GPA. This combination represents science
academics without necessitating a science content exam. The differences in high school science requirements and course choices make a science content exam too complicated for this study.

Cognitive Ability Measure

This study used the Shipley-2 Vocabulary and Block Patterns scales (Shipley, Gruber, Martin, & Klein, 2009). The Vocabulary and Block Patterns scales combine to represent a composite measure of general cognitive ability. Vocabulary contains 40 items and takes 10-15 minutes while Block Patterns contains 26 items and is administered in 12 minutes. Typically, the Shipley-2 costs $42 per 25 tests for each scale. However, as this is being utilized for research purposes this cost was negotiated with the publisher to about $35 per 25 tests. Standardization studies for this measure resulted in median reliability estimates of $\alpha = .84$ for the Vocabulary scale and $\alpha = .85$ for the Block Patterns scale (Shipley, et al., 2009).

Personality

To assess personality for this study, items from the International Personality Item Pool (IPIP) were selected (see Appendix C). This is a freely available pool of items based on well-supported personality models. For this study, a pre-developed 50-item measure of the Big 5 constructs was selected: Neuroticism, Extraversion, Openness to New Experiences, Agreeableness, and Conscientiousness. These items follow the format of the NEO-PI measure (McCrae & Costa, 1987; Costa & McCrae, 1992). For each subscale of the IPIP measure, reliability estimates are typically seen in the range of $\alpha = .78-.86$.

Science Motivation

To assess science motivation of participants, the choice was made to use the Science Motivation Questionnaire II (SMQ-II, Glynn, Brickman, Armstrong, & Taasoobshirazi, 2011). This scale assesses five aspects of motivation: Intrinsic Motivation, Career Motivation, Self-
Determination, Self-Efficacy, and Grade Motivation. The multiple facets of this scale are particularly beneficial for this study as they touch on multiple aspects of motivation in science. The SMQ-II is freely available for research purposes and contains 25 items. The development study for the second edition of the SMQ reported reliability estimates of $\alpha = .81-.92$ for the measure’s subscales (Glynn, et al., 2011).

Attitudes Toward Science

In addition to motivation, attitudes toward science were assessed using the Attitude Toward Science Scale (Vitale & Johnson, 1988). This scale assesses four factors: Instrumental Value of Science, Active Participation in Science, Difficulty and Complexities of Science, and General Attitude Toward School. The instrument has 31 total items and is freely available for use in research studies. Reliability estimates from the development study for this assessment ranged from .55-.99, with the fourth factor representing the General Attitude Toward School showing low reliability at .55 (Vitale & Johnson, 1988). This fourth factor is evaluated carefully in the present study.

Data Analysis Strategy

To address the research questions in the present study, three separate analyses were conducted. Initially, descriptive statistics were used to evaluate responses and check for normality. Missing data was handled in the present study using maximum likelihood estimation concurrently with the analysis process in Mplus (Osborne, 2012). Reliability was evaluated for each scale using Cronbach’s alpha, looking for a minimum value of .70 for measures in development and used for research purposes, and .80 for previously validated clinical measures (Henson, 2001). Then, confirmatory factor analysis was conducted on each measure used in the study to support the use of composite variables in the LPA. Next, LPA was used to explore
unobserved patterns in responses for measures of occupational preference and personality. Finally, covariate analysis and descriptive statistics were used to explore the effect of covariates in the latent profiles identified using LPA.

RQ 1 Analysis – Latent Profile Analysis

To address Research Question 1, a LPA was conducted to explore latent profiles of individuals on the measures of science motivation, science attitude, science interest, and science achievement. Figure B.1 shows the path diagram of this model, which was analyzed using Mplus 6 with the mixture modeling add-on. Additionally, possible sex differences in the profiles were examined in the final model (see Figure B.2). The intention is to identify any differences between the groups based on frequency of male and female students identified as most probable for each profile. Previous LPAs have been conducted in similar studies such as science self-efficacy (Chen & Usher, 2012), perfectionism and performance in STEM (Rice, Lopez, & Richardson, 2012), academic self-concept (Marsh, et al., 2009), student motivation (Hietajärvia, Tuominen-Soinia, Hakkarainena, Salmela-Arob, & Lonkaa, 2015), and sex differences in research motivation (Smith, Deemer, Thoman, & Zazworsky, 2014).

Hypothesis 1: Distinct profiles will be seen such that males show higher representation in a profile with high science interest, high science motivation, high science attitude, and high science achievement. Females will likewise show higher representation in a profile with low science interest, low science motivation, low science attitude, and low science achievement.

Hypothesis 2: To support compromise due to sex stereotypes in science, a profile will also be seen with high representation of females showing low science interest but high science motivation, high science attitude, and high science achievement.
For Research Question 2, a separate LPA was conducted on the Big 5 personality factors Neuroticism, Extraversion, Openness to New Experiences, Agreeableness, and Conscientiousness (see Figure B.3). This LPA was then analyzed in combination with the profiles identified in research question 1 as a covariate. Personality factors are particularly well suited to LPA analysis as the factors are considered sub-facets of a higher-order personality construct (Finch & West, 1997; Merz & Roesch, 2011). LPA allows these sub-facets to be modeled together in a search for combined profiles of personality patterns (Finch & West, 1997; Merz & Roesch, 2011).
Hypothesis 3: A profile most closely related to high science occupational preference will consist of positive Conscientiousness, Openness, and Agreeableness and negative Extroversion and Neuroticism.

**Figure B.3.** Latent profile analysis Model 3 for Research Question 2

RQ 3 Analysis

Finally, for research question 3, all self-concept covariates (sex, socio-economic status, personality profile, and cognitive ability) were examined in relation to the latent profiles of science occupational interest. This analysis was conducted in combination with the original LPA by adding the covariates to be modeled concurrently as is shown in Figure B.2 and as recommended in the literature (Clark & Muthen, 2009; Marsh, et al., 2009)

Hypothesis 4: Sex, personality, and cognitive ability will all be statistically significant predictors of occupational preference for science.
Hypothesis 5: Sex will be the strongest predictor of occupational preference, followed by cognitive ability and then personality profile.

Figure B.2. Latent profile analysis Model 2 with covariates for Research Question 3
APPENDIX C

EXTENDED RESULTS AND DISCUSSION
Results

Descriptive Statistics

Descriptive statistics are used to evaluate the central tendency and spreadoutness of the collected data. Table C.2 provides means, standard deviations, value ranges, and skewness and kurtosis information, and internal consistency reliability estimates for all of the continuous scales and variables. For categorical or nominal variables, modes and frequencies of responses are provided.

Data Cleaning

Data cleaning and assumptions checks were preformed prior to beginning the LPA. First, data was evaluated for values out of range. For instance, if a GPA was recorded above 5.0, which is the highest possible value, the false value was removed from the dataset. These errors are generally keystroke or response errors and can falsely skew data in analyses (Osborne, 2012). Five values out of range were removed from the dataset.

Next, assumptions were checked to evaluate the univariate and multivariate normality of the variables to be included in the analysis. While normality is not necessarily an assumption of LPA, it is a common assumption in inferential statistics and was checked to better support the statistical conclusions validity of the present study. Univariate normality was examined using the skewness and kurtosis values as shown in Table C.2, with values between -1.00 and 1.00 typically representing normally distributed data (Osborne, 2012). Kurtosis on two subscales of the Attitude Towards Science measure and one subscale of the Science Motivation measure were outside of this range. However, due to the robust nature of the analysis being conducted, these values were deemed acceptable without transformation.
### Table C.2

**Descriptive Statistics for Study Variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATS_AP</td>
<td>22.34</td>
<td>5.63</td>
<td>7-35</td>
<td>-.228</td>
<td>.140</td>
<td>.807</td>
</tr>
<tr>
<td>ATS_DC</td>
<td>21.38</td>
<td>4.62</td>
<td>3-38</td>
<td>-.090</td>
<td>1.289</td>
<td>.620</td>
</tr>
<tr>
<td>ATS_GS</td>
<td>9.94</td>
<td>2.78</td>
<td>3-15</td>
<td>-.083</td>
<td>-.403</td>
<td>.735</td>
</tr>
<tr>
<td>ATS_IV</td>
<td>51.18</td>
<td>8.90</td>
<td>13-70</td>
<td>-.610</td>
<td>1.463</td>
<td>.875</td>
</tr>
<tr>
<td>Shipley-BP</td>
<td>16.70</td>
<td>5.20</td>
<td>0-26</td>
<td>-.514</td>
<td>.505</td>
<td>.877</td>
</tr>
<tr>
<td>Shipley-Vocab</td>
<td>16.83</td>
<td>4.81</td>
<td>7-39</td>
<td>-.391</td>
<td>.370</td>
<td>.776</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>29.25</td>
<td>9.11</td>
<td>7-50</td>
<td>.050</td>
<td>-.769</td>
<td>.893</td>
</tr>
<tr>
<td>Extraversion</td>
<td>33.79</td>
<td>8.27</td>
<td>10-50</td>
<td>-.281</td>
<td>-.349</td>
<td>.887</td>
</tr>
<tr>
<td>Openness</td>
<td>36.92</td>
<td>6.32</td>
<td>17-50</td>
<td>-.441</td>
<td>.121</td>
<td>.765</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>36.34</td>
<td>6.42</td>
<td>10-50</td>
<td>-.679</td>
<td>.840</td>
<td>.785</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>35.45</td>
<td>6.39</td>
<td>18-50</td>
<td>-.300</td>
<td>-.262</td>
<td>.818</td>
</tr>
<tr>
<td>SM_CM</td>
<td>15.32</td>
<td>6.11</td>
<td>5-25</td>
<td>.177</td>
<td>-1.132</td>
<td>.934</td>
</tr>
<tr>
<td>SM_GM</td>
<td>18.20</td>
<td>4.69</td>
<td>5-25</td>
<td>-.472</td>
<td>-.252</td>
<td>.855</td>
</tr>
<tr>
<td>SM_IM</td>
<td>16.58</td>
<td>4.60</td>
<td>5-25</td>
<td>-.229</td>
<td>-.539</td>
<td>.872</td>
</tr>
<tr>
<td>SM_SD</td>
<td>14.95</td>
<td>4.42</td>
<td>5-25</td>
<td>.018</td>
<td>-.271</td>
<td>.846</td>
</tr>
<tr>
<td>SM_SE</td>
<td>18.02</td>
<td>4.35</td>
<td>5-25</td>
<td>-.435</td>
<td>-.246</td>
<td>.858</td>
</tr>
<tr>
<td>Mother’s SES</td>
<td>29.88</td>
<td>15.20</td>
<td>1-52</td>
<td>-.523</td>
<td>-.927</td>
<td></td>
</tr>
<tr>
<td>Father’s SES</td>
<td>31.94</td>
<td>14.72</td>
<td>1-55</td>
<td>-.785</td>
<td>-.506</td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>3.41</td>
<td>.70</td>
<td>1.6-5.0</td>
<td>.240</td>
<td>-.086</td>
<td></td>
</tr>
</tbody>
</table>

**Frequency**

<table>
<thead>
<tr>
<th>Science Career Interest</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Interested</td>
<td>21.8%</td>
</tr>
<tr>
<td>Somewhat Interested</td>
<td>31.9%</td>
</tr>
<tr>
<td>Somewhat Not Interested</td>
<td>20.0%</td>
</tr>
<tr>
<td>Very Not Interested</td>
<td>26.3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>43.5%</td>
</tr>
<tr>
<td>Female</td>
<td>56.5%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>11th</td>
<td>26.9%</td>
</tr>
<tr>
<td>12th</td>
<td>73.1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>African-American</td>
<td>11.3%</td>
</tr>
<tr>
<td>Asian</td>
<td>3.2%</td>
</tr>
<tr>
<td>Caucasian</td>
<td>51.4%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>29.2%</td>
</tr>
<tr>
<td>Other</td>
<td>4.9%</td>
</tr>
</tbody>
</table>

82
Multivariate normality was evaluated using Mahalanobis distance as described in Henson (1999). A Q-Q plot of Mahalanobis distance with an associated $\chi^2$ value was evaluated, and the values follow a linear trend as expected to support multivariate normality (see Figure C.4).

Figure C.4. Q-Q plot for variables in LPA to support multivariate normality

Missing data in the present study was handled using the Mplus default, which is to estimate missing values based on all available information. Cases were removed from analysis if all variables involved in the analysis were missing. Missing data in the present study was minimal, with most variables having no more than 5% missing. Parent education and occupation questions did have missingness as high as 28%.
Validity and Reliability

Confirmatory factor analysis (CFA) was conducted on the measures used in the present study to evaluate the factor structure and support the use of the composite values. Model fit was evaluated based on structural equation modeling fit indices CFI, TLI, RMSEA, and WRMR. Cut off values for these indices exist, though they have been debated repeatedly in the literature (Hu & Bentler, 1999; Marsh, Hau, & Wen, 2007). Values for CFI and TLI closer to .90, RMSEA values below .06, and WRMR values closer to 1.0 are considered indicative of good model fit. Additionally, internal consistency reliability was assessed using Cronbach’s alpha, with values above .80 showing good reliability for previously validated measures, and values above .70 supporting reliability for measures still in development.

Shipley-2. The Shipley-2 Vocabulary and Block Patterns scales were evaluated using CFA. The Weighted Least Squares Means and Variances (WLSMV) estimator was used as items on these measures are dichotomously scored as either correct or incorrect. Results of all CFA attempts are reported in Table C.3. The vocabulary measure failed to exhibit acceptable model fit with a one-factor structure as designed. Review of the technical manual for the Shipley-2 revealed the test publisher’s argument that this measure cannot be evaluated using a CFA due to the dichotomously scored nature of the tests and the type of measure (Shipley, et al., 2009). While the WLSMV estimator can account for the dichotomous items, the type of measure does not appear to conform to the CFA expectations. This difficulty has been discussed in the literature (Floyd & Widaman, 1995) and one possible solution is to parcel items and create more variance in each observed indicator in the model (see Little, Cunningham, Shahar, & Widaman, 2002).

This approach was attempted with the Vocabulary test by combining the items into eight parcels of five items, which resulted in an acceptable measurement model. Item parceling was
also attempted with the Block Patterns scale, with responses on a shared stimulus parceled into a composite. This resulted in 12 item parcels for the Block Patterns measure. An alternate two-factor model was also fit with the Block Patterns scale based on prior psychometric studies (Beaujean, Hull, Sheng, Worrell, Bolen, & Verdisco, 2016; Ferguson, Powell, & Hull, in progress). Reliability estimates for these two measures were $\alpha=.78$ for Vocabulary and $\alpha=.88$ for Block Patterns, supporting the internal consistency of the items. Based on prior validation results combined with the CFA results and reliability estimates reported here, the measures are supported for the present study.

Table C.3

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
<th>WRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary, One Factor</td>
<td>1053.49</td>
<td>740</td>
<td>0.78</td>
<td>0.77</td>
<td>0.03</td>
<td>1.31</td>
</tr>
<tr>
<td>Vocabulary, Parceled</td>
<td>52.95</td>
<td>20</td>
<td>0.93</td>
<td>0.90</td>
<td>0.08</td>
<td>0.06*</td>
</tr>
<tr>
<td>Block Patterns, One Factor</td>
<td>769.61</td>
<td>252</td>
<td>0.90</td>
<td>0.89</td>
<td>0.07</td>
<td>1.79</td>
</tr>
<tr>
<td>Block Patterns, Parceled</td>
<td>221.05</td>
<td>27</td>
<td>0.74</td>
<td>0.65</td>
<td>0.16</td>
<td>0.09*</td>
</tr>
<tr>
<td>Block Patterns, Two Factor</td>
<td>617.65</td>
<td>251</td>
<td>0.93</td>
<td>0.92</td>
<td>0.06</td>
<td>1.56</td>
</tr>
</tbody>
</table>

Note. Parceled models estimated using Maximum Likelihood, SRMR values reported in place of WRMR. Good fit indicated by values lower than 0.08 per Hu and Bentler (1999).

IPIP. The IPIP Personality measure was evaluated using a CFA with a WLSMV estimator due to the categorical nature of Likert scaled items. Model fit for this CFA is reported in Table C.4 and did not meet the fit index conventions suggested in Hu and Bentler (1999). However, many scholars in personality assessment have discussed this issue and support relaxed fit standards in CFA due to the complex nature of personality constructs (Hopwood &
Donnellan, 2010; Marsh, et al., 2004). Additionally, the five-factor model exhibited significantly better fit than a single factor model, and the pattern coefficients for the items on their designated factors were above .40 except for three items, two on the Openness factor (0.32 and 0.36) and one on the Agreeableness factor (0.30). Reliability for the full IPIP scales was estimated at $\alpha=.86$, with subscale estimates ranging from $\alpha=.77-.89$ (see Table C.2). Based on the combined results, the measure is supported for use in the present study.

Table C.4

**CFA Model Fit Summary for IPIP**

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
<th>WRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Five Factor</td>
<td>3050.48</td>
<td>1165</td>
<td>0.83</td>
<td>0.83</td>
<td>0.06</td>
<td>1.86</td>
</tr>
<tr>
<td>One Factor</td>
<td>7145.57</td>
<td>1175</td>
<td>0.48</td>
<td>0.45</td>
<td>0.12</td>
<td>3.30</td>
</tr>
</tbody>
</table>

Science Motivation Questionnaire-II. The SMQ-II was published with a five-factor structure, so this structure was evaluated in the present study. Model fit for the measure was good overall as shown in Table C.5, although the RMSEA value was higher than preferred. An alternative one-factor model was also estimated, and the fit was noticeably worse than the five-factor model. Reliability for the full SMQ-II was found to be $\alpha=.95$ in the present study. Reliability was also evaluated for all five subscales in the SMQ-II, and results ranged from $\alpha=.85-.93$. Therefore, the five-factor model is supported for the present study.
Table C.5

*CFA Model Fit Summary for SMQ-II*

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
<th>WRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Five Factor</td>
<td>1102.28</td>
<td>265</td>
<td>0.95</td>
<td>0.95</td>
<td>0.09</td>
<td>1.46</td>
</tr>
<tr>
<td>One Factor</td>
<td>2949.68</td>
<td>275</td>
<td>0.85</td>
<td>0.83</td>
<td>0.16</td>
<td>3.02</td>
</tr>
</tbody>
</table>

Attitude Towards Science. For the ATS measure, the published four-factor model had noted concerns in the development process (Vitale & Johnson, 1988). In this study, the published four-factor model was estimated as well as an alternative one-factor model. In addition, due to the concerns with the fourth factor (General Attitude Towards School), an alternative three-factor model was also estimated removing the three items on the fourth factor. Model results for these analyses are report in Table C.6. Good model fit was not achieved for this scale with any of the three models estimated. Further work needs to be done to effectively evaluate this scale. Reliability estimates for all four scales ranged from $\alpha=0.62$ to $0.88$ with the full-scale reliability estimated at $\alpha=0.83$, indicating good internal consistency. The Difficulties and Complexities subscale was the lowest reliability with $\alpha=0.62$, though this is not the same factor to show low reliability in the development study. Reliability estimates are considered adequate for a scale still in development (Henson, 1999), but due to the lack of construct validity evidence, values on this measure should be interpreted carefully.
Table C.6

*CFA Model Fit Summary for ATS*

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
<th>WRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Four Factor</td>
<td>1741.93</td>
<td>458</td>
<td>0.74</td>
<td>0.72</td>
<td>0.12</td>
<td>2.04</td>
</tr>
<tr>
<td>Three Factor</td>
<td>1243.24</td>
<td>347</td>
<td>0.80</td>
<td>0.78</td>
<td>0.12</td>
<td>1.83</td>
</tr>
<tr>
<td>One Factor</td>
<td>1212.29</td>
<td>464</td>
<td>0.72</td>
<td>0.70</td>
<td>0.09</td>
<td>1.76</td>
</tr>
</tbody>
</table>

*Note.* Four-factor model resulted in non-positive definite matrix due to General School factor

Inferential Statistics

Latent profile analysis was used as the primary analysis to address the research questions. All latent profile analyses were conducted using Mplus 6 with a maximum likelihood estimation method. Composite variables were used as opposed to item-level data to simplify the model and support convergence. Missing values were estimated during the analysis process, except for cases where all variables in the analysis were missing. The cases were removed from the analysis if values on all variables in the analysis were missing ($n=4$).

Research Question 1

A latent profile analysis was used for Research Question 1 to identify the presence of latent profiles on measures of science interest, motivation, attitude, and academic experiences. The model is shown in Figure B.1 and models an increasing number of latent profiles in an iterative model building process. Model 1 was estimated with only one profile, Model 2 with two profiles, and so on to Model 5 with five profiles. Mode fit statistics are provided in Table C.7. Model 4 was retained as the best model to fit the data based on the lowest loglikelihood value, lower AIC, BIC, and SABIC values, a high entropy value, results of the Lo-Mendel Ruben test, and profiles that are supported by theory. The hypothesized 3-profile model was not directly
supported in this analysis, though the retained 4-profile model may be an expansion of the hypothesized structure.

Table C.7

*LPA Model Fit Summary for Research Question 1*

<table>
<thead>
<tr>
<th>Model</th>
<th>Loglikelihood</th>
<th>AIC</th>
<th>BIC</th>
<th>SABIC</th>
<th>Entropy</th>
<th>Smallest Class %</th>
<th>LMR p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-8909.94</td>
<td>17867.88</td>
<td>17955.96</td>
<td>17879.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-8507.38</td>
<td>17088.75</td>
<td>17224.54</td>
<td>17107.21</td>
<td>0.89</td>
<td>45%</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>3</td>
<td>-8370.30</td>
<td>16840.603</td>
<td>17024.10</td>
<td>16865.538</td>
<td>0.90</td>
<td>30%</td>
<td>0.002</td>
</tr>
<tr>
<td>4</td>
<td>-8290.15</td>
<td>16706.31</td>
<td>16937.51</td>
<td>16737.73</td>
<td>0.89</td>
<td>15%</td>
<td>0.010</td>
</tr>
<tr>
<td>5</td>
<td>-8253.86</td>
<td>16659.72</td>
<td>16938.63</td>
<td>16697.62</td>
<td>0.90</td>
<td>2%</td>
<td>0.614</td>
</tr>
</tbody>
</table>

*Note.* The Lo-Mendell Ruben (LMR) test compares the current model to a model with *k*-1 profiles.

The 4-profile model that was retained is detailed in Table C.8. The means and standard deviations of the variables used to create the profiles are presented for each latent profile. Note that the standard deviations are the same as they are constrained in the analysis by default. The differences between the four latent groups are largely due to differences in interest, motivation, and attitude towards science, which supports the theoretical approach used in the present study.

Profile 1 contains students with the lowest level of interest in science, a more negative attitude and low motivation towards science, and a low GPA, on average. Profile 2 contains those students with the highest interest in science, high attitude and motivation scores, and the highest average GPA. Profile 3 students are low in terms of science interest, but more towards the middle in science attitude and motivation, with a GPA almost as high on average as seen in
Profile 2. Finally, Profile 4 contains students who are somewhat interested in science, have mid-
level attitudes and motivations towards science, and a GPA lower than Profile 3 but higher than
Profile 1. Profile 1 could be referred to as “Lowest Science Interest and GPA”, Profile 2 as
“Highest Science Interest and GPA”, Profile 3 as “Low Science Interest, High GPA”, and Profile
4 as “Medium Science Interest and GPA”.

Table C.8

4-Profile Model Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Profile 1</th>
<th>Profile 2</th>
<th>Profile 3</th>
<th>Profile 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Classes</td>
<td>3.74 (0.88)</td>
<td>4.31 (0.88)</td>
<td>3.78 (0.88)</td>
<td>3.71 (0.88)</td>
</tr>
<tr>
<td>GPA</td>
<td>2.95 (0.63)</td>
<td>3.77 (0.63)</td>
<td>3.59 (0.63)</td>
<td>3.11 (0.63)</td>
</tr>
<tr>
<td>Science Interest</td>
<td>3.65 (0.52)</td>
<td>1.32 (0.52)</td>
<td>3.49 (0.52)</td>
<td>2.00 (0.52)</td>
</tr>
<tr>
<td>SM – Intrinsic Motivation</td>
<td>10.31 (3.05)</td>
<td>21.03 (3.05)</td>
<td>15.58 (3.05)</td>
<td>16.91 (3.05)</td>
</tr>
<tr>
<td>SM – Career Motivation</td>
<td>8.18 (2.84)</td>
<td>23.06 (2.84)</td>
<td>11.27 (2.84)</td>
<td>16.27 (2.84)</td>
</tr>
<tr>
<td>SM – Self-Determination</td>
<td>9.19 (3.00)</td>
<td>19.37 (3.00)</td>
<td>14.50 (3.00)</td>
<td>14.43 (3.00)</td>
</tr>
<tr>
<td>SM – Self-Efficacy</td>
<td>12.63 (3.22)</td>
<td>21.75 (3.22)</td>
<td>18.23 (3.22)</td>
<td>17.26 (3.22)</td>
</tr>
<tr>
<td>SM – Grade Motivation</td>
<td>12.74 (3.70)</td>
<td>21.80 (3.70)</td>
<td>18.37 (3.70)</td>
<td>17.62 (3.70)</td>
</tr>
<tr>
<td>ATS – Instrumental Value</td>
<td>41.36 (6.75)</td>
<td>59.05 (6.75)</td>
<td>48.80 (6.75)</td>
<td>51.67 (6.75)</td>
</tr>
<tr>
<td>ATS – Academic</td>
<td>19.20 (5.39)</td>
<td>23.89 (5.39)</td>
<td>21.73 (5.39)</td>
<td>23.24 (5.39)</td>
</tr>
<tr>
<td>ATS – Difficulties &amp; Complexities</td>
<td>20.15 (4.52)</td>
<td>21.25 (4.52)</td>
<td>20.88 (4.52)</td>
<td>22.75 (4.52)</td>
</tr>
<tr>
<td>ATS – General School</td>
<td>11.23 (2.71)</td>
<td>9.433 (2.71)</td>
<td>9.54 (2.71)</td>
<td>10.16 (2.71)</td>
</tr>
</tbody>
</table>

*Note.* SM = Science Motivation, ATS = Attitude Towards Science
Research Question 2

For the second research question on the existence of latent profiles of personality types, a second latent profile analysis was conducted on the five personality scales from the IPIP. The model for this analysis can be seen in Figure B.3. The model fitting process followed the same procedure as in Research Question 1, and the retained model contained three latent profiles (see Table C.9). An argument could be made for Model 1 due to the non-statistically significant LMR test result for Model 2. However, Model 3 showed good model fit based on the indices and loglikelihood values, did show a statistically significant improvement over Model 2, and theoretically aligns with previous work (see Merz & Roesch, 2011). Therefore, the hypothesized latent personality profiles were supported in this analysis.

Table C.9

<table>
<thead>
<tr>
<th>Model</th>
<th>Loglikelihood</th>
<th>AIC</th>
<th>BIC</th>
<th>SABIC</th>
<th>Entropy</th>
<th>Smallest Class %</th>
<th>LMR p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-6414.44</td>
<td>12848.89</td>
<td>12888.13</td>
<td>12856.40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-6284.62</td>
<td>12601.25</td>
<td>12664.03</td>
<td>12613.27</td>
<td>0.87</td>
<td>12%</td>
<td>0.728</td>
</tr>
<tr>
<td>3</td>
<td>-6185.28</td>
<td>12414.55</td>
<td>12500.89</td>
<td>12431.09</td>
<td>0.79</td>
<td>4%</td>
<td>0.005</td>
</tr>
<tr>
<td>4</td>
<td>-6171.02</td>
<td>12398.04</td>
<td>12507.92</td>
<td>12419.08</td>
<td>0.79</td>
<td>4%</td>
<td>0.355</td>
</tr>
</tbody>
</table>

Note. The Lo-Mendell Ruben (LMR) test compares the current model to a model with \(k-1\) profiles.

The means and standard deviations of the three latent profiles are presented in Table C.10. Students in Profile 1 align with what Merz & Roesch (2011) termed “Reserved” students, those with relatively low scores on the personality traits. The Profile 2 participants tend to be
higher across the board and have the highest Neuroticism scores, matching the “Excitable” profile from Merz and Roesch (2011). Finally, the Profile 3 student averages match the profile termed “Well-Adjusted” in the 2011 study.

Table C.10

3-Profile Model Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Profile 1 Reserved</th>
<th>Profile 2 Excitable</th>
<th>Profile 3 Well-Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroticism</td>
<td>20.03 (8.79)</td>
<td>31.22 (8.79)</td>
<td>27.87 (8.79)</td>
</tr>
<tr>
<td>Extraversion</td>
<td>19.93 (6.32)</td>
<td>29.45 (6.32)</td>
<td>38.99 (6.32)</td>
</tr>
<tr>
<td>Openness</td>
<td>21.35 (4.76)</td>
<td>34.18 (4.76)</td>
<td>40.42 (4.76)</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>19.92 (4.80)</td>
<td>33.74 (4.80)</td>
<td>39.97 (4.80)</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>22.72 (5.23)</td>
<td>32.95 (5.23)</td>
<td>38.75 (5.23)</td>
</tr>
</tbody>
</table>

Research Question 3

For the final research question, the profiles identified in Research Question 1 were further analyzed to evaluate the effects of covariates. The covariates of interest were: personality, as defined by the personality profiles in research question two; cognitive ability, both verbal and spatial; sex; and socio-economic status, as defined by parent education level and occupation. The covariate affects were analyzed in the latent profile analysis process as shown in Figure B.2. The results of this analysis are presented in Table C.11.
Table C.11

*Covariate Analysis Results for the 4-Profile Model*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Profile 1 Lowest Interest &amp; GPA</th>
<th>Profile 3 Low Interest, High GPA</th>
<th>Profile 4 Medium Interest &amp; GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>-0.96</td>
<td>-0.32</td>
<td>-1.373*</td>
</tr>
<tr>
<td>Shipley-2 Vocabulary</td>
<td>-0.19*</td>
<td>-0.10</td>
<td>-0.23*</td>
</tr>
<tr>
<td>Shiple-2 Block Patterns</td>
<td>-0.07</td>
<td>-0.002</td>
<td>-0.05</td>
</tr>
<tr>
<td>Mother’s SES</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.003</td>
</tr>
<tr>
<td>Father’s SES</td>
<td>-0.02</td>
<td>0.01</td>
<td>-0.002</td>
</tr>
<tr>
<td>Personality</td>
<td>-1.08*</td>
<td>-0.12</td>
<td>-0.556</td>
</tr>
</tbody>
</table>

*Note.* *=p<.05, Profile 2 (Highest Science Interest & GPA) served as reference group.

Based on this analysis, some of the covariates do have a statistically significant impact on the profiles. Results suggest sex was statistically significantly different between Profile 2 (Highest Science Interest & GPA) and Profile 4 (Medium Science Interest & GPA), with more female’s falling into the high science interest profile. Vocabulary scores were also statistically significantly different between Profile 2 (Highest Science Interest & GPA) and both Profile 1 (Lowest Science Interest & GPA) and Profile 4 (Medium Science Interest & GPA). The negative coefficients show the high science interest group tended to have higher vocabulary scores than these other two profiles. Finally, personality was shown to be statistically significantly different for Profile 1 (Lowest Science Interest & GPA) as compared to Profile 2 (Highest Science Interest & GPA). A cross-tabulation shows the difference is between students considered “Well-Adjusted”, as significantly more of them fall into the high science interest profile (see Table C.12).
Table C.12

*Cross-Tabulation of Statistically Significant Covariate Differences*

<table>
<thead>
<tr>
<th>Profile</th>
<th>Reserved</th>
<th>Excitable</th>
<th>Well-Adjusted</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Lowest Science Interest &amp; GPA</td>
<td>3</td>
<td>26</td>
<td>16</td>
<td>21(17%)</td>
<td>21(13%)</td>
</tr>
<tr>
<td>2 Highest Science Interest &amp; GPA</td>
<td>2</td>
<td>27</td>
<td>49</td>
<td>24(20%)</td>
<td>51(32%)</td>
</tr>
<tr>
<td>3 Low Science Interest, High GPA</td>
<td>1</td>
<td>44</td>
<td>44</td>
<td>35(28%)</td>
<td>54(34%)</td>
</tr>
<tr>
<td>4 Medium Science Interest &amp; GPA</td>
<td>1</td>
<td>47</td>
<td>30</td>
<td>43(35%)</td>
<td>34(21%)</td>
</tr>
</tbody>
</table>
Discussion

The present study provides support for elements of C&C theory to explore the development of science interest. For research question one, latent profiles were identified and show a pattern of students with varying levels of interest in science, collectively understood to include attitude and motivation, as well as differences in GPA. Hypothesis one regarding the predominance of males in the high science interest profile and females in the low science interest profile was not supported. However, hypothesis two regarding the existence of a profile made up largely of females with a low science interest and a high motivation and attitude towards science was supported and is represented by Profile 3.

Profile 1, named “Lowest Science Interest & GPA”, is made up of students with the lowest level of science interest on average across almost all measures. The only exception is in General Attitude Towards School, which is a divergent concept from specific science interest. Additionally, this profile had the lowest average GPA of all four profiles. This was the smallest profile in the present study, and is made up of males and females equally. Of the 42 students in this group, 26 are considered Excitable, making up about 62% of the profile. Students in this profile were found to be statistically significantly different from those in the Highest Science Interest & GPA profile based on personality profile as well as vocabulary ability. This group of students is not interested in science, holds less than a 3.00 average in GPA, has a low vocabulary ability, and is generally Excitable in personality.

Profile 2 is the “Highest Science Interest & GPA” profile, and is characterized by students with the highest interest, attitude, and motivation towards science. These are students that believe in the value of science, are confident in their abilities in science, and see science as a strong possibility for their future career. Additionally, these students have the highest GPA on
average of all of the profiles. The profile contains primarily female students, making up 68% of the profile and containing 32% of the females in the present study. Students in this profile are also primarily considered Well-Adjusted in terms of personality, and have significantly higher vocabulary scores than those in the Lowest Science Interest & GPA profile.

Profile 3 was named “Low Science Interest & High GPA”, as the science interest scores for this profile are almost as low as in Profile 1 but the GPA average is the second highest. These are students that do well in school, as seen in the GPA average and relatively high scores on the Grade Motivation and Academic Attitude scales. However, subscales and questions specifically related to science interest are lower than Profile 2 and often lower than Profile 4, while still being higher than Profile 1. This suggests these students are not very interested in science, other than as necessary for their academic goals. Profile 3 is the largest profile in the present study with 89 students, and it is primarily made up of female students (n=54). This is also the largest percentage of female students for any of the four profiles (34%), and most closely matches the hypothesized profile of students highly motivated and positive about science but not interested in science. In terms of personality, Well-Adjusted and Excitable students equally characterize this profile, and personality was not a statistically significant factor. No aspects of self-concept evaluated in the present study are statistically significantly different for this profile as compared to the Highest Science Interest & GPA profile.

Finally, Profile 4 is the “Medium Science Interest & GPA” profile. This profile is made up of students with scores around the middle of the responses in the present study. The GPA average is 3.11, and answers on the science interest questions and scales were typically higher than Profile 1 but lower than Profiles 2 and 3. These are students that are somewhat ambivalent towards school and their career decisions, and science in particular. This is the second largest
profile with 77 total students, and is the only profile with more males than females (males n=43). There are also more males in this profile than in any other, with 35% of the males in the present study falling into this group. Profile 4 was statistically significantly different from the Highest Science Interest & GPA Profile 2 in terms of sex and vocabulary scores. Specifically, Profile 4 contains more males and averages lower vocabulary scores when compared to Profile 2. This profile also contains the largest group of Excitable students in terms of personality profiles.

For research question two, the effects of the self-concept covariates on the identified latent profiles were mixed. Theoretically, sex was expected to affect the latent profiles such that males were more prevalent in the high science interest profiles with females more prevalent in low science interest profiles (Hill et al., 2010; NSF, 2015). This disparity has been connected to sex stereotype affect in previous work (Leaper, et al., 2012; Legewie & DiPrete, 2011; Robnett & Leaper, 2012; Valian, 2014). This was somewhat supported in the present study, as the Low Science Interest and High GPA profile (Profile 3) contains the highest percentage of females (34%). However, the second highest female representation is in the High Science Interest and GPA profile (Profile 2), suggesting science stereotype may not be a strong influence on high school females’ science interest development. The statistically significant difference between the Highest Science Interest & GPA profile and the Medium Science Interest & GPA profile, with more females present in the high science interest group, further weakens support for sex stereotype in the present study. Additionally, the largest male presence was in the Medium Science & GPA profile, as opposed to the High Science Interest & GPA profile as might be expected.

One possible explanation for the lack of strong sex stereotype affect in the present study may be related to the observed personality profile differences. Results from previous studies on
the impact of socio-economic status of career decisions found that students with a lower socio-
-economic status can overcome the perceived barrier to science careers their status implies,
particularly if they have high levels of ambition and do well in school (Breakwell, 1992; Croll,
2008). This same logic may apply here, where female students with Well-Adjusted personality
types are able to overcome perceived sex stereotype more effectively. An alternative explanation
could be related to cognitive ability, as previous studies have also shown students with higher
cognitive ability and educational opportunities are less susceptible to sex stereotype (Buday, et
al., 2012). Both of these explanations align with C&C theory as the sextype compromise is the
hardest to overcome, and would require a stronger self-concept (Gottfredson, 1981; 1996).

The Shipley-2 Vocabulary scores were also found to be statistically significantly different
in the covariate analysis. Students in the Highest Science Interest & GPA profile had
significantly higher vocabulary scores as compared to those in the Lowest Science Interest &
GPA and Medium Science Interest & GPA profiles. This does not follow the theoretical pattern
implied in the literature, where vocabulary ability is typically not predictive of science interest
(Wai, et al., 2009). The spatial ability component of cognitive ability was expected to show
differences between profiles of science interest based on the literature (Geary, et al., 2000;
Jenson, 2006; Maeda & Yoon, 2013; Reilly & Neumann, 2013). However, this was not
supported in the present study. It is possible the difference is related to higher cognitive ability
generally as opposed to specifically vocabulary ability, which has previously resulted in more
equal sex distributions related to science interest (Buday, et al., 2012). However, as the spatial
ability measure was not statistically significantly different in the analysis it is unclear what the
source of this difference might be.
For the personality profile analysis in research question two, three latent profiles were identified. This result replicates a prior study by Merz and Roesch (2011), who also found three profiles of personality types. Their profiles were referred to as Reserved, Excitable, and Well-Adjusted, and the patterns in their study closely match the patterns observed in the present study. The LPA process allows for a holistic discussion of personality from a higher-order conception as opposed to discussing only the Big 5 subscales. As personality is a complex construct with a well-supported higher-order structure (Digman, 1990; Goldberg, 1990; Hopwood & Donnellan, 2010; McRae & Coast, 1987), modeling like LPA improves the literature and theory of personality and its impact on other important variables such as interest, as evidenced in the present study.

Personality profile was found to be statistically significantly different for students in the Lowest Science Interest & GPA profile as compared to students in the Highest Science Interest & GPA profile. Specifically, those students who are considered Well-Adjusted are more likely to be interested in science and have a higher GPA. Previous studies of Big 5 personality factors and science interest typically examined the subscales only, while the present study took a higher-order personality profile approach. However, previous studies have found Openness to New Experiences, Introversion, Agreeableness, and Conscientiousness all predict science interest (Ackerman & Heggestad, 1997; Feist, 2012; Hong & Lin, 2011), which aligns well with the Well-Adjusted profile characteristics (Merz & Roesch, 2011). Integrating the five factors into a complex personality profile adds understanding to relationships between personality and other variables such as science interest in the present study.
Limitations

The present study benefits from a large sample of 11th and 12th grade high school students. However, this sample is representative of a large public school district in the Southwestern United States and may not be applicable to students in other areas of the country or in other school types. Future studies could account for more detail related to the types of science courses taken and the grades students earned in each course. An argument can be made for the difference between an A in a required normal-level course and an A in an advanced science elective taken at the AP level, but this is not accounted for in this study. Finally, the literature suggests sex differences are more pronounced in some areas of science such as physics, engineering, and computer science as opposed to biology and social sciences. These different science preferences are not captured in the present study, and including this may present a clearer understanding of sex stereotype.

Conclusion and Future Directions

Overall, the present study provides an empirical evaluation of the complex construct of interest in science at the high school level. The use of LPA as an advanced modeling strategy strengthens the present study and provides a more detailed explanation of interest and the variables affecting interest. Gottfredson’s (1981, 1996) C&C theory is partially supported. The LPA person-oriented analysis method models the development of interest with four distinct profiles, but the hypothesized impact of self-concept covariates is mixed. Sex, cognitive ability, and personality were found to have an impact on science interest, though not always as hypothesized.

The results of the present study suggest high school interventions to encourage female interest in science may be effective, and efforts to minimize the effect of stereotype on
occupational preferences should continue. For the present study, students who were identified as highly interested in science were generally academically advanced, Well-Adjusted, and showed high vocabulary ability. For students that do not fit these characteristics, additional interventions could be considered. Students who do well academically but are not as stable in personality may need to be introduced to science careers in a different way, one that appeals to their Excitable personality profile. It is possible these students do not see their personality as a good fit for science careers. While work in a laboratory or in front of a computer may not be appealing to this type of student, they may be interested in other careers in science such as those involving working with people or in the field. Additionally, increasing science interest for students who are not as advanced in academics or who show lower vocabulary ability would need to be addressed in other ways. For instance, introducing these students to aspects of science that fit their aptitudes better may help them see how science careers are on option for them as well.

Future work could replicate this study with other high school populations to verify the current findings. Additionally, due to the lack of clear evidence to support the sex stereotype effect and the impact of self-concept on interest development, this study could be conducted with college students and science professionals. Replication of this study with older individuals can be used to identify where sex differences are occurring in science careers. Additionally, the lack of clarity regarding the possible effect of sex stereotype justifies the use of a mixed-methods approach. Mixed-methodology would allow researchers to replicate the quantitative study conducted here and add a qualitative component to evaluate student perspectives on sex stereotype and the impact it may or may not have had on their interest in science.
APPENDIX D

IPIP PERSONALITY ITEMS
IPIP Personality Items

Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence.

1. I feel threatened easily.
2. I feel comfortable around people.
3. I believe in the importance of art.
4. I have a good word for everyone.
5. I am always prepared.
6. I dislike myself.
7. I make friends easily.
8. I have a vivid imagination.
9. I believe that others have good intentions.
10. I pay attention to details.
11. I change my mood a lot.
12. I am skilled in handling social situations.
13. I love to think up new ways to do things.
15. I get chores done right away.
16. I have frequent mood swings.
17. I am the life of the party.
18. I carry the conversation to a higher level.
19. I love to help others.
20. I carry out my plans.
21. I panic easily.
22. I know how to captivate people.
23. I enjoy hearing new ideas.
24. I make people feel at ease.
25. I make plans and stick to them.
26. I am filled with doubts about things.
27. I start conversations.
28. I enjoy thinking about things.
29. I am concerned about others.
30. I complete tasks successfully.
31. I feel threatened easily.
32. I warm up quickly to others.
33. I can say things beautifully.
34. I trust what people say.
35. I do things according to a plan.
36. I get stressed out easily.
37. I talk to a lot of different people at parties.
38. I enjoy wild flights of fantasy.
39. I sympathize with others' feelings.
40. I am exacting in my work.
41. I fear for the worst.
42. I don't mind being the center of attention.
43. I get excited by new ideas.
44. I am easy to satisfy.
45. I finish what I start.
46. I worry about things.
47. I cheer people up.
48. I have a rich vocabulary.
49. I treat all people equally.
50. I follow through with my plans.
APPENDIX E

DEMOGRAPHIC SURVEY
Demographic Survey

What school do you currently attend?

______________________________________________________________________

Please indicate your sex from the following:

○ Male
○ Female

What is your current grade level?

○ 11th
○ 12th
○ Other

Please enter your year of birth: ______________

Please indicate your ethnicity from the options below:

○ African American
○ Asian
○ Caucasian
○ Hispanic
○ Other

Please indicate your mother/female guardian’s highest completed level of education:

○ Less than 7th grade
○ Junior high / Middle school (9th grade)
○ Partial high school (10th or 11th grade)
○ High school graduate
○ Partial college (at least one year)
○ College education
○ Graduate degree

What is your mother/female guardian’s occupation?

______________________________________________________________________

Please indicate your father/male guardian’s highest completed level of education:

○ Less than 7th grade
○ Junior high / Middle school (9th grade)
○ Partial high school (10th or 11th grade)
○ High school graduate
○ Partial college (at least one year)
○ College education
○ Graduate degree

What is your father/male guardian’s occupation?

______________________________________________________________________
What is your current high school GPA? (If you do not know, give an estimate)
______________

Which of the following science courses have you taken?
- Biology 1
- Pre-AP Biology 1
- Chemistry 1
- Pre-AP Chemistry 1
- Physics 1
- Pre-AP Physics 1
- Integrated Physics and Chemistry
- Anatomy & Physiology
- Aquatic Science
- Environmental Systems
- AP Chemistry
- AP Biology
- AP Environmental Science
- AP Physics 1
- AP Physics 2
- AP Physics C
- Scientific Research and Design
- Laboratory Management
- Other: ____________________________

Thinking about your plans for your future career, at this point how interested are you in a career in a science field?
- Very Interested
- Somewhat Interested
- Somewhat Not Interested
- Very Not Interested
COMBINED REFERENCE LIST


Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: evidence for overlapping traits. *Psychological Bulletin, 121*, 219-245. doi: 10.1037/0033-2909.121.2.219


108
Bergman, L. R., & Magnusson, D. (1997). A person-oriented approach in research on
developmental psychopathology. *Development and Psychopathology, 9*, 291-319. doi:
10.1017/S095457949700206X


Gottfredson’s theory. *Journal of Vocational Behavior, 62*, 250-271. doi: 10.1016/S0001-
8791(02)00026-X


attitudes and activities. *Public Understanding of Science, 1*(2), 183-197. doi: 10.1088/0963-
6625/1/2/003

Buday, S. K., Stake, J. E., & Peterson, Z. D. (2012). Gender and the choice of a science career:
The impact of social support and possible selves. *Sex Roles, 66*, 197-209.

MA: Cambridge University Press.

in a mixture model. *Journal of Classification, 13*, 195-212.


