THE COGNITIVE EFFECTS OF GENDERED LANGUAGE ON MEMORY IN COLLEGE STUDENTS: IMPLICATIONS FOR STANDARDIZED TESTING

by

Ryan S. Watkins, B.A.

A thesis submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Arts
with a Major in Psychological Research
May 2018

Committee Members:

Crystal Oberle, Chair

Shirley Ogletree

Jon Lasser

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

ACKNOWLEDGEMENTS

To my patient mentor, Dr. Crystal Oberle: I am eternally grateful to you for allowing me the independence of function, for your enthusiasm towards researching multiple subjects, and for the saintly aid you provided during this time. To Dr. Shirley Ogletree: thank you for the wonderful class discussions you facilitated, the research postulations we explored, and the research materials needed to help create this study. To Dr. Jon Lasser: thank you for creating within me the confidence to not only attempt a higher education, but also complete it.

A very special thank you goes to my first psychology professor, Cyndi Walker-Ewert: I thank you for rekindling the fires of passion I have for exploring the human mind, for helping me mold my potential, and for teaching me I am not without my own aptitude in places I wouldn't think to venture.

To John R. Angulo: thank you for bringing me into the wonderful field of Psychology and helping me keep up when I was struggling to do so. To Brittany Culp: you were always there when I needed a friend and you helped consistently reaffirm I was always much better than I ever thought I was. To Laurel and John Watkins: you always asked questions and showed support for my education and my hobbies, even when you didn't understand them, and I appreciate that more now than I ever have.

I am very grateful to my other friends and family, especially those who are no longer with me, who have offered me the support and experience I have needed in my life to arrive at this junction and to continue moving forward.

A distinctive mention goes out to Angela E. Johnson, Joseph Perez, Fatemeh Fereidooni, Sara R. Hicks, Stephanie Purol, Marieke Visser, Matthew Pacailler, Katie Stephenson, Katherine Mooney, E. Marin Cordivin, Kirstin L. Critz, Linda-Brooke Thompson, Nicole R. Stokes, and the MAPR program in general for providing the environment I needed to learn and improve my skills as a researcher.

An exceptional gratitude goes out to Texas State University. It was amazing to have the opportunity to study and conduct my research in the facilities you provided.

And finally, I am also indebted to the following university staff: Dr. Reiko Graham, Dolores Reyes-Pergioudakis, Danielle L. McEwen, Dr. Ty Schepis, and Dr. Katherine R. Warnell for their reliable support and assistance throughout my time in this project.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

ABSTRACT

This study aims to investigate the influence of gendered language on memory retention of material in a reading passage that is comparable to those used in standardized tests, hypothesizing that gender-neutral language would create contextual ambiguity and consequently lead to lower test scores in individuals with gender traditional attitudes but higher test scores in individuals with gender transcendent attitudes. Three separate versions of the same reading task were created: one containing gendered language with stereotype-consistent terminology, one containing gendered language with stereotype-inconsistent terminology, and one containing gender-neutral language with no stereotype-consistent or inconsistent terminology. 153 college students from the Department of Psychology of Texas State University were utilized for this study. There were higher scores in the Gendered Language, Stereotype-Consistent condition (GLSC), when compared to the Gendered Language, Stereotype-Inconsistent condition (GLSI) and Gender Neutral Language Condition (GNL), and there were lower scores for the GLSI condition in those who scored lower on the gender transcendent items of the SRQ. The results show, along with previous studies, an increasing trend of transcendent attitudes in education covering gender and social roles. This seems to, at times, impact the capability of an individual to learn and affects an individual's ability to recall information in those that showcase weaker transcendent attitudes and, subsequently, those that showcase strong traditional attitudes.

I. INTRODUCTION

Gendered language encompasses words that are representative of an individual's gender and creates gender biases by bringing gender to noticeable standards and adding to generalized opinions of gender. For example, gendered language would include words like *fireman* or *governess/nanny*, but gender-neutral language would include words like *firefighter* or *childcare provider/babysitter* (Thomson, Murachver, & Green 2001). Another example of gender-neutral language is the avoidance of the pronoun *he*, including the forms *him* and *his*, to refer to people of an unknown gender (Fowler, 2015). Studying this relation connecting language and gender began with three seminal books in 1975: Key's *Male/Female Language*, Lakoff's *Language and Women's Place*, and Thorne and Henley's *Language and Sex: Difference and Dominance*. These particular books review analyses on subjects and disciplines that supported gender discrimination throughout society at the time (Freeman & McElhinny, 1996). Further in 1990, Tannen provided research that would instigate popular interest in researching the relation between language and gender through various theoretical and methodological perspectives (Coates, 1985; Gordon, 2011).

However, the history of the study of the relation between language and thought is more interminable. The relation between language and thought functions in two directions, influencing each other in significant fashions governing human behavior. More specifically, this is the decisive case with linguistic relativity, which contests that human perception and thought processes can be molded by language (Bigler & Leaper, 2015). The words in any given language represent concepts and ideas that can affect how one determines how they view the world around them (Boroditsky, 2011; Gentner &

Loewenstein, 2002; Khorsroshahi, 1989). Simply put, words used in description influences individuals' cognition (Waxman, 2013).

The theory that language influences cognition associated with gender is deep-rooted and reinforced by various research (Arthur, Bigler, Liben, Gelman, & Ruble, 2008; Bigler & Leaper, 2015; Henley, 1989; Leaper, 2014; Leaper & Bigler, 2004; Rennels & Langlois, 2014; Stahlberg, Braun, Irmen, & Sczesny, 2007). This research also overwhelmingly indicates that the usage of gender-identifying terminology further enables the existence of gender bias in any given environment. Within the English language, there exists many different ways of gender identification: pronouns (e.g., *she*, *he*, *hers*, *his*), honorific titles (e.g., *Ms*., *Mrs*., *Mr*., *Sir*), nouns (e.g., *boy*, *girl*), and occupational titles (e.g., *stewardess*, *policeman*). Unfortunately, evidence on the detrimental effects of using these gender-identifying words is gradually increasing. This presents an obligation to consider reviewing and potentially reforming language practices in regards to gender-related cognition. One such consideration could possibly be assessing the language utilized in standardized tests and how gendered language within the test might affect test scores.

This review is divided into four sections. The first section holds a concise summary of gendered language, circumstances where language, particularly English, can indicate individuals' gender. The second section provides an analysis of the relationship between gendered language and gender-related models of attitude and thought. The third section discusses current and future reforms regarding the substitution of gendered language with gender-neutral language. The fourth section explains what standardized testing entails, providing information on its strengths and weakness with respect to

language. This final section is then followed by a set of conclusions and extension to the current research.

Gendered Language

The degree and means by which language reflects gender varies by the type of language that is used (Leaper, 2014). Stahlberg et al. (2007) argue for three varieties of language that differ in regard to gender representation. At one extreme are the grammatical gendered languages, such as Spanish and French, which include gender-based pronouns and gender-based forms of all personal pronouns. In Spanish, examples of gender-based pronouns include *ella* for *her and él* for *he,* and examples of gender-based personal nouns include *perro* (masculine) and *perra* (feminine) for *dog*. At the other extreme are genderless languages, such as Korean and Japanese, which include neither gender-based pronouns nor gender-based personal nouns. Finally, between the two extremes are the natural gendered languages, such as English and Swedish, which include gender-based pronouns (e.g., *she, he*) and gender-based forms of some personal nouns (e.g., *policeman, policewoman*) but not all personal nouns (e.g., *neighbor, nurse*).

Like most other grammatical and natural gendered languages, English uses a binary categorization of gender representation with numerous words to represent male and female, but very few words to reflect other gender categories, such as intersex or hermaphrodite (de Klerk & Bosch, 1996; Friedman, Leaper, & Bigler, 2007; Gelman, Taylor, & Nguyen, 2004; Henley, 1989; Van Fleet & Atwater, 1997). The multitude of ways that these two categories are represented include gender-based adjectives (e.g., *masculine, feminine*), pronouns (e.g., *she, he*), general nouns (e.g., *man, woman*),

occupational titles (e.g., *actress*, *fireman*), honorific titles (e.g., *Mrs*., *Mr*.), and names given at birth (e.g., *John*, *Jennifer*).

Gendered Language and Gender-Related Thought

Empirical evidence suggests that gendered language, which is very prevalent in the English language, may profoundly influence a child's prejudice and stereotyping by affecting a child's ability to properly conceptualize and categorize living and non-living objects (Bigler & Liben, 2006; Mulac, Bradac, & Karol-Mann, 1985; Rice, 2000; Waxman, 2013; Welch-Ross & Schmidt, 1996), which may lead to gender stereotyping and prejudice (Arthur et al., 2008; Bigler & Liben, 2006, 2007; Friedman et al., 2007; Gelman et al., 2004; Hilliard & Liben, 2010). For instance, gender-indicating nouns can facilitate untrue categorization and conceptualization of social groups (i.e., people who interact with each other and share similar characteristics and a sense of unity). This finding is congruent with the view of psychologic essentialism, whereby people think that everything has a quintessence, something (e.g., appearances, behaviors) that is both intrinsic and central to its character (Medin & Ortony, 1989). Gelman (2003; Gelman & Heyman, 1999) presented substantial evidence that category labeling can lead to children's beliefs that the affiliates of a particular category share features, attributes, or other commonalities, in spite of no presentation of conceptual or perceptual indications regarding these parallels. Two such studies found that preschool children quickly and easily made inferences about hypothetical children they didn't know based solely on the assignment of either gender-specific first names (Bauer & Coyne, 1997) or gender-specific category labels (Gelman, Collman, & Maccoby, 1986).

Such categorizing, and subsequently labeling said categories, is deemed a fundamental part of the social learning of prejudice and stereotyping (Bigler & Leper, 2015). This, in agreement with a basis of essentialist reasoning giving characterization to a child's view of category labels, posits that the discovery of gender category labels in childhood and adolescence is apt to encourage gender prejudice and stereotyping. According to Bigler and Liben's (2006, 2007) developmental intergroup theory, when gendered language is used, children immediately focus on gender as a basis for categorizing themselves and others. In testing this theory, Hilliard and Liben (2010) indeed found that compared to gender-neutral language, gender category labels taught or instigated by adults increased the children's attentiveness towards gender that then led to greater stereotyping on the basis of gender.

As well as influencing cognition, gendered language also influences the affect, behavior, and interests of children (Cimpian, Mu, & Erickson, 2012; Vervecken, Hannover, & Wolter, 2013). For instance, the gender stereotyping that stems from use of gendered language can lead to in-group favoritism, whereby children exhibit substantial biases for same-gender peers (Arthur et al., 2008; Bigler & Leaper, 2015; Bigler & Liben, 2006; Leaper & Bigler, 2004). Such favoritism is seen when, upon experiencing gender-labeled language, children devote more time playing with same-gender peers than cross-gender peers (Fagot, Leinbach, & Hagan, 1986; Hilliard & Liben, 2010).

In conclusion, gendered language demonstrates an early developmental effect on an individual's affect, behavior, interests, and cognition, consistent with the principles of linguistic relativity and the developmental intergroup theory. Commonly, the utilization of gendered labels (e.g., *boys*, *girls*) increases attention to gender as a basis for

categorization and subsequently leads to the development of an essentialist perspective on gender that each gender shares an inherent yet invisible collection of traits. This utilization of gendered language facilitates gender prejudices and stereotypes, as well as in-group favoritism and bias.

Attempting to Substitute Gendered Language

Over the past few decades, the U.S. has experienced a decline in the generic use of masculine pronouns and nouns. Regarding pronouns, the more inclusive phrase, *he or she,* is becoming a more common substitute for the generic *he* to refer to people, although researchers still believe that this phrase keeps a focus on gender, in comparison to the preferred plural pronoun, *they* (Liben & Bigler, 2015). Regarding nouns, using occupational titles with male suffixes (e.g., *fireman*, *councilman*) to represent all workers is less frequent in the present day. These words are being substituted with either gender-neutral counterparts (e.g., *firefighter*, *councilor*) or with parallel masculine and feminine (e.g., *councilman*, *councilwoman*). As with the pronouns, the gender-neutral nouns are favored for the purpose of avoiding any parochial gender bias and stereotyping with different jobs (Liben, Bigler, & Krogh, 2002).

Stringer (2011) argues that additional reforms are needed. In particular, educators should abstain from using gender category labels when referring to their students and with respect to their classroom organization. This change would permit more inclusivity of non-binary gender identities, such as transgender, pangender, and genderqueer (Liben & Bigler, 2015; Petterson, Dixson, Little, & Vasey, 2016). Moving in the right direction, several colleges have policies that serve to reduce gender category labeling and categorization, as well as all-gender restrooms and dormitories.

Strengths and Weaknesses of Standardized Testing

A standardized test is any assessment that meets the following criteria: (a) all test-takers respond to the same exact questions or to a group of equivalent questions from a general selection, (b) all test administrators follow a standard set of rules to ensure that the test-takers experience the same conditions, and (c) all tests are scored using a standard set of procedures that allows for comparative performance evaluations of individual test-takers or entire institutions (Hidden curriculum, 2014; *Olson & Sabers, 2008;* Popham, 1999; Porter, McMaken, Hwang, & Yang, 2011). Although standardized tests typically employ multiple-choice questions, they may also incorporate other question types, such as essay questions, short-answer questions, and true-or-false questions. These questions may be administered in traditional paper-and-pencil format, on a computer, or even orally. Standardized tests are utilized for a diverse array of education-related purposes, such as identifying learning disabilities of students in need of special services, and assessing a student's achievement or mastery of a particular subject in preparation for subsequent courses or grade levels (Popham, 2016; Porter et al., 2011; Ravitch, 1985).

Despite the controversial debate over the efficacy of standardized tests in the U.S., many people in the educational system believe these tests to be reliable and valid assessments in academia, in large part due to the standardized design, administration, and scoring procedures that decrease the chance of bias and favoritism on evaluations (Ravitch, 1985), but also because these tests have been subjected to empirical evaluation to determine the tests' replicability and generalizability (Kuncel & Hezlett, 2007). Regarding the latter, through the process of developing a standardized test, initial results

may reveal such problems as test bias due to subjectivity when creating the questions, measurement error, and inflation of scores (Neill, 2009). Subsequent revisions to the test are then made to increase its reliability and validity.

Nevertheless, most assessment experts advise against utilizing standardized tests as a sole evaluation of academic achievement, performance, and learning (Porter et al., 2011; Ravitch, 1985). In fact, some experts argue that using standardized tests may result in limiting the core and national curriculum, creating an environment that weakens scholastic involvement and retainment of both students and teachers (Neill, 2009; Popham, 1999). Moreover, an individual's score is affected not only by the subject matter learned inside the institution, but also by subject matter learned external to the institution and an individual's inherent intelligence (Popham, 1999; Ravitch, 1985). However, a value-added model could statistically control for these latter two extraneous factors (Hassel & Rosch, 2008; Miner, 2000).

Language is at the core of another major problem with standardized tests in the U.S., whereby non-native English-speaking students may score lower than native English-speaking students on the tests simply due to their lack of familiarity with or mastery of the English language, as opposed to a lack of knowledge (Duran, 1989; Garcia, 1991). This claim is supported by research, revealing that this language bias reduces the tests' reliability and validity (Abedi, Leon, & Mirocha, 2001; Abedi & Lord, 2001). Thus, language background should be a crucial consideration in the development, administration, and interpretation of standardized tests (Pennock-Roman & Rivera, 2011).

Conclusions

The indication of gender through linguistics is commonplace in many structures within the English language. These structures implement a binary handling of gender, which can be unpredictable in a global environment that includes a noticeable percentage of individuals whose gender identity is outside the categorization of "male" and "female". Additionally, gendered language steers individuals to categorize themselves and others specifically by gender, which then leads to gender prejudices and stereotypes. Research supports that shifting vernacular and lexical input from gender category labeling to gender-neutral linguistics would reduce these problems (Bigler & Leaper, 2015; Bigler & Liben, 2007).

In regards to standardized tests, many students and educators argue that these tests are biased due to the language of the tests. Namely, non-native English-speaking students may score lower than native English-speaking students on the tests simply due to their lack of familiarity with or mastery of the English language, as opposed to a lack of knowledge. Furthermore, cultural biases may exist for similar reasons. What is not yet known, however, is the potential impact of gendered language on such testing.

Purpose of the Current Research

This thesis project investigates the potential influence of gendered language on the memory retention of material from a reading passage, comparable to the reading passages and accompanying questions that appear on many standardized tests, such as the Scholastic Aptitude Test (SAT). For this study, participants were randomly assigned to read one of three reading passages: one with gendered language that is consistent with traditional gender stereotypes (e.g., a *woman* is a nurse), one with gendered language that

is inconsistent with traditional gender stereotypes (e.g., a *man* is a nurse), and one with gender-neutral language (e.g., a *person* is a nurse). Hypothesis 1 is that there would be higher scores in the Gendered Language, Stereotype-Consistent Condition (GLSC) when compared to the Gender Neutral Language Condition (GNL). Hypothesis 2 is that there would be lower scores in the Gendered Language, Stereotype-Inconsistent Condition (GLSI) when compared to the GNL condition. Hypothesis 3 is that the positive impact of the GLSC condition passage on recall would be greatest for those who score higher on the gender traditional items of the Social Roles Questionnaire (SRQ). Hypothesis 4 is that the negative impact of the GLSI condition on recall will be greatest for those who score lower on the gender transcendent items of the SRQ.

II. METHOD

As described below, this research involved participants completing an online study that incorporated both an experimental manipulation (i.e., random assignment into the language conditions) and assessment of individual differences (i.e., gender and both transcendent and traditional attitudes centered around gender role characteristics).

Participant Recruitment

Participants were recruited through two means. First, participation was sought from students who were enrolled in Introduction to Psychology courses and comprised the Department of Psychology Human Subjects Pool at Texas State University. Through Sona Systems, which is a participant management software system, potential subjects were presented with a brief description of the study's purpose and procedures, along with a link to the online study. In exchange for completing this study, students earned one credit to go toward the four-credit requirement for their Introduction to Psychology course. Second, participation was sought from students who were enrolled in a variety of undergraduate courses in the Department of Psychology at Texas State. Through course websites in the Teaching, Research, and Collaboration System (TRACS), potential subjects were presented with a brief description of the study's purpose and procedures, along with a link to the online study. In exchange for completing this study, students earned a small amount of extra credit that was determined by the professors teaching those courses.

Materials and Data Collection Procedures

For this research, participants completed an online survey and assessment that was created in Qualtrics. Upon clicking on the link to this study, they were presented

with a consent form that briefly described the purpose, procedures, and approximate completion time. This form also emphasized that participation is completely voluntary and that responses are anonymous, and it provided the contact information for the researcher, the Institutional Review Board (IRB) Chair, and the IRB Regulatory Manager. Once participants finished reading the consent form, they gave their electronic signature at the end of the form, acknowledging that they had read and understood the information in the consent form. Afterward, they scrolled to the bottom of the page and clicked a button labeled "Next" to continue with the survey and assessment. The participant then completed the following parts of the study in the order that they are listed.

Demographic Questionnaire

Participants were asked to describe characteristics of themselves to better assess the data that they provided. These demographics included age, biological sex, ethnicity, English as a first language or second language, estimated grade point average (GPA), and academic rank (see Appendix A).

Reading Passage

There were three versions of this passage, based on an excerpt from *20 Years at Hull-House* (Addams, 1912), where the plot remained the same but the characters' names and traits varied (see Appendix B). The GLSC passage contained gendered language that is consistent with gender stereotypes. The GLSI passage contained gendered language that is inconsistent with gender stereotypes. The GNL passage contained gender-neutral language and contained no stereotypes. Table 1 describes the specific similarities and differences among these three passages.

Table 1

Similarities and Differences Among the Reading Passages for the Different Conditions

<table>
<tr><th>Gendered Language, Stereotype-Consistent Condition</th><th>Gendered Language, Stereotype-Inconsistent Condition</th><th>Gender Neutral Language Condition</th></tr>
<tr><td colspan="2">54 instances of gendered-language words:
Pronouns (he, she, his, him, and her)
Nouns (husband, wife, man, woman, father, mother, son, daughter, boy, girl, masculine names of George and Jacob, and feminine names of Amanda and Lily)</td><td>0 instances of these words; gender-neutral names of Skyler, Taylor, Oakley, and Justice</td></tr>
<tr><td>13 behaviors, characteristics, or interests that are consistent with traditional gender norms:
Man is engineer
Man has gambling problem
Man has drinking problem
Man performs violent act
Man is aggressive
Woman is nurse
Woman is concerned with appearance
Woman is compassionate and likes helping others
Boy likes sports
Boy steals things
Girl is fearful
Girl has a doll
Girl does not like to be dirty</td><td>13 behaviors, characteristics, or interests that are inconsistent with traditional gender norms:
Woman is engineer
Woman has gambling problem
Woman has drinking problem
Woman performs violent act
Woman is aggressive
Man is nurse
Man is concerned with appearance
Man is compassionate and likes helping others
Girl likes sports
Girl steals things
Boy is fearful
Boy has a doll
Boy does not like to be dirty</td><td>All behaviors, characteristics, and interests are assigned to people whose gender is not identified</td></tr>
</table>

Participants were randomly assigned to one of these three reading-passage conditions. For each condition, the reading passage remained on the screen for a minimum of 4 minutes to help ensure that participants carefully read the passage. At the

end of that duration, participants were allowed to click on a button in order to proceed to the memory test. After clicking on that button, they were no longer able to go back to reread the passage.

Memory Test

This test included 20 multiple-choice questions based on the previous section's passage (see Appendix C). This multiple-choice format was used to increase the comparability with most standardized tests such as the STARS, TAKS, TAAS, SAT, GRE, and ACT. The total number of correctly-answered questions served as the measure of memory retention.

Social Roles Questionnaire

The Social Roles Questionnaire-Revised (SRQ; Baber & Tucker, 2006) is a 13-item assessment that measures transcendent and traditional attitudes centered on role characteristics. The questionnaire includes five items that measure gender-transcendent attitudes and eight items that measure gender-traditional attitudes, which are attitudes based on beliefs toward gender-ambiguous social roles and established gender-based social roles respectively. Participants will rate based on a 7-point Likert scale (see Appendix D). This inventory was described on the consent form as an attitude questionnaire.

Statistical Analyses

A 3 x 2 x 2 between-subjects ANOVA was used to analyze the data. For this analysis, the independent variables were condition (GLSC, GLSI, and GNL), gender traditional (high and low, based on a median split), and gender transcendent (high and low, based on a median split). The dependent variable was the memory retention score.

III. RESULTS

Participant Demographics

There were a total of 153 participants in this study. In regards to the biological sex of the participants, there were 18 male participants (11.8%) and 135 female participants (88.2%). In regards to the ethnicity of the participants, there were 70 Caucasian students (45.8%); 48 Hispanic, Latinx, or Chicanx participants (31.4%); 17 African American students (11.1%); 6 Asian American, Middle Eastern, or Pacific Islander participants (3.9%); 1 Native American participant (0.7%); 10 Biracial or Multiracial participants (6.5%); and 1 Other participant (0.7%). When asked if English was their first/native language, 124 participants answered 'Yes' (81.0%), 14 participants answered 'No' (11.8%), and 11 participants answered 'I learned English at the same time as another language' (7.2%).

In regards to academic rank, there were 31 Freshman-level participants (20.3%), 36 Sophomore-level participants (23.5%), 25 Junior-level participants (16.3%), and 61 Senior-level participants (39.9%). In regards to age, the range for the study's pool of participants was between a minimum of 18 years of age and a maximum of 44 years of age ($M = 20.95$, $SD = 3.27$). In regards to estimated GPA, the range for the participants fell amid the minimum of 1.00 and the maximum of 4.00 ($M = 3.24$, $SD = .512$). Finally, there were 49 participants in the GLSC group, 48 participants in the GLSI group, and 56 participants in the GNL group.

Descriptive and Inferential Results

For preliminary analyses, chi-square analyses were used to assess group differences in the categorical variables of biological sex, ethnicity, whether or not English

was your first/native language, and academic rank, and one-way ANOVA analyses were used to assess group differences in the continuous variables of age, GPA, gender transcendent attitudes, and gender traditional attitudes. As indicated below, there were no significant differences found and therefore, no confounding variables existed within this study.

As per the assessment of the continuous variable of age and estimated GPA, participants in the different conditions (i.e. GLSC, GLSI, and GNL) did not differ, $F(2, 150) = 0.34$, $p = .71$, $\eta_p^2 = .005$. The same lack of difference was found in the assessment of the continuous variable of estimated GPA, $F(2, 150) = 0.06$, $p = .79$, $\eta_p^2 = .003$. In assessing the transcendent attitudes of the participants, no difference was found in the participants among the different conditions, $F(2, 150) = .91$, $p = .41$, $\eta_p^2 = .012$. In assessing the traditional attitudes of the participants, no difference was found in the participants among the different conditions, $F(2, 150) = .47$, $p = .62$, $\eta_p^2 = .006$.

In reviewing the chi-square test assessing the categorical variable of biological sex, participants once again did not differ, $\chi^2 = .44$, $p = .80$. The chi-square test assessing the categorical variable of ethnicity revealed no difference in participants among the different conditions, $\chi^2 = 14.19$, $p = .29$. In going through the results of the chi-square test assessing the categorical variable of English being the participants' first/native language, there was again no difference, $\chi^2 = 1.30$, $p = .86$. In reviewing the final chi-square test assessing the categorical variable of academic rank, there was no significant difference, $\chi^2 = 8.71$, $p = .19$.

The ANOVA analysis revealed a significant main effect of language condition on memory score (see Table 2). The Tukey HSD post-hoc test revealed that participants in

the GLSC group (M = 17.73, SD = 2.03) achieved better scores in comparison than those in the GLSI group (M = 15.75, SD = 4.20; p = .004) and GNL group (M = 16.61, SD = 3.35; p = .05). The results of a signification interaction between language condition and transcendent attitudes on memory score are shown in Figure 1. From Figure 1, participants who scored high in transcendent attitudes were shown to be moderately unaffected by language condition, but participants that scored low in transcendent attitudes appear to be greatly affected by the GLSI scale.

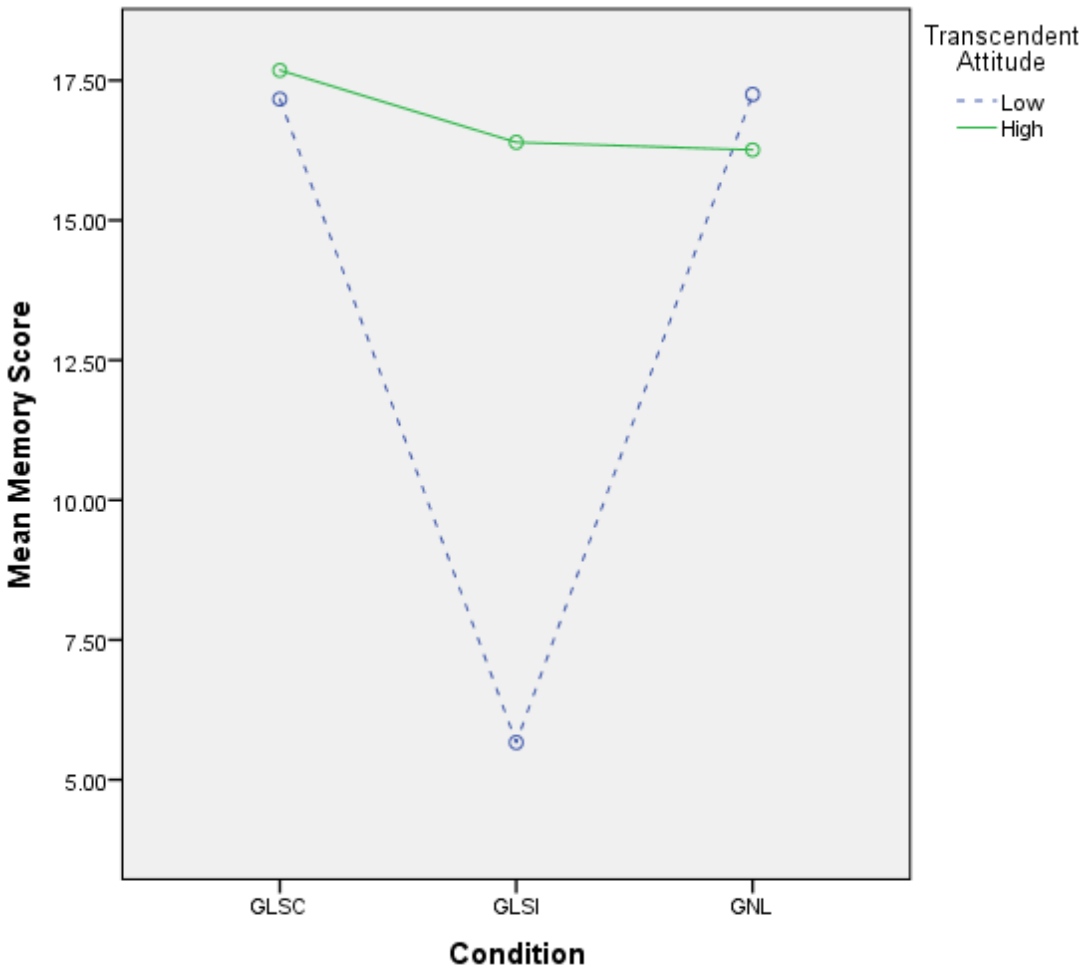
Table 2

Tests of Between-Subjects Effects

Source	**F**	**p**	**η_p^2**
Condition	17.409	.000	.197
Transcendent Attitude	6.284	.013	.042
Traditional Attitude	.058	.810	.000
Transcendent Attitude x Traditional Attitude	.017	.897	.000
Condition x Transcendent Attitude	13.575	.000	.161
Condition x Traditional Attitude	.730	.483	.010
Condition x Transcendent Attitude x Traditional Attitude	1.127	.290	.008

Figure 1

Univariate ANOVA Score of the Interaction between Transcendent Attitudes x Traditional Attitudes x Group of Participants



IV. DISCUSSION

Discussion of Results and Theoretical Implications

This thesis project explored the potential influence of gendered language on the memory retention of material from a reading passage, comparable to the reading passages and accompanying questions that appear on many standardized tests, such as the SAT. As mentioned above, there was a significant main effect of language condition on memory score. Hypothesis 1 is that there would be higher scores in the GLSC when compared to the GNL. The results of the post-hoc test in the analysis revealed that participants in the GLSC group achieved better scores in comparison than those in the GLSI group and GNL group, which supports Hypothesis 1. As mentioned in previous literature, gendered language automatically activates gender stereotype schemas. When the story details are consistent with those schemas, like they are in the GLSC condition, an individual is able to connect the new information from the story to firmly established schema that serve to strengthen their memory (Arthur et al., 2008; Welch-Ross & Schmidt, 1996).

Hypothesis 2 is that there would be lower scores in the GLSI when compared to the GNL condition. There were lower scores in the GLSI group in comparison to GNL, but the results displayed marginal difference and there was no significance that could be found to adequately support Hypothesis 2. Hypothesis 3 is that the positive impact of the GLSC condition passage on recall would be greatest for those who score higher on the gender traditional items of the SRQ. As with Hypothesis 2, there was no major interaction or significance to support the hypothesis. As theorized in previous literature, a movement of teaching transcendent attitudes towards gender and social roles throughout levels of education has begun increasing in appearance in recent years (Bigler & Leaper,

2015, Bigler & Liben, 2006; 2007). This suggests that there could fluctuating differences in the impact of gendered language and stereotypes on memory recall due to the increased awareness of differing attitudes in those who score high in gender traditional attitudes. Thus, this could explain the lack of support for both hypotheses.

Hypothesis 4 is that the negative impact of the GLSI condition on recall will be greatest for those who score lower on the gender transcendent items of the SRQ. The results of this signification interaction between language condition and transcendent attitudes on memory score are shown in participants that scored low in transcendent attitudes appear to be greatly affected by the GLSI scale (see Figure 1). In regards to gendered language activating gender stereotype schemas, when the story details are inconsistent with those schemas, an individual is either poorly able to or unable to link new information they have gathered from reading the story to a firmly established schema (Arthur et al., 2008; Baber & Tucker, 2006; Welch-Ross & Schmidt, 1996). This poor connection or lack of one can be detrimental to their ability of memory recall as it weakens the connection to previously gathered information. Additionally, low scores in the gender transcendent attitudes, while not necessarily meaning high scores in gender traditional attitudes, permit that an individual is supportive to some extent of non-traditional gender roles, but the individual is aware of traditional roles and subsequently subjected to the same gender stereotype schema activation in the mind. Therefore, this results in an increased likelihood of answering the question incorrectly due to a weaker connection to previously gathered information that does not match up with gender stereotypes, and the individual's score suffers for it.

In reviewing the data and previous literature, there is an apparent rising trend of transcendent attitudes in education towards gender and social roles and there are varying levels of how they're introduced throughout the levels of education. This could possibly be due to this trend's growth exceeding the response of the education system to react to such a trend, disrupting an individual's capability to learn, and affecting the memory in those that feel weaker transcendent attitudes and, potentially, for those who feel strong traditional attitudes. However, for norm-referenced tests, this shouldn't affect scores because the normative sample would display the same issues. This could theoretically be due to an inconsistent structure of education that varies on its use of and approach to gendered language and gender-neutral language.

Practical Applications

As discussed by Dr. Bigler and Dr. Leaper (2015), introducing gender-neutral language and elimination of gendered language in early adolescence would lead to a wider range of gender identity as well as a better comprehensiveness in the absence of gender. Teachers and parents who cancel out gender labels in the classroom and home respectively is possible form of practical application during the period of growth. Businesses and organizations can amend legal language to be less gender-exclusive by reducing gender labeling and classifications. Colleges and universities can establish free, incentivized workshops and events that help illustrate and educate on this topic, as well as providing public service announcements to the differences between gender-neutral language and gendered language in English and other languages.

From the evidence gathered in this experiment and those in the literature regarding this topic, if gender-neutral language were introduced at the collegiate stage of

education in standardized testing, then it could possibly affect the ability to recall or correctly recognize information gathered in reading passages. This would consequently affect test scores of test-takers. However, introducing gender-neutral language in elementary or primary levels of education and maintained throughout stages of education could potentially reduce or completely diminish the effects that gender-neutral language can have cognitively on test-takers.

Strengths and Limitations

A strength of the study is that there are those who have not yet heard of what gendered language and gender-neutral language is, so in exploring this subject, the knowledge of either language and their effects might spread and help those in improving their communication and understanding of gender-based concepts. In the potential of helping those improve their communication, it might help those realize that the language, or languages, they know might be different for others, especially those who are categorized as English as a Second Language (ESL) and didn't grow up learning English as a child. In that same vein, a strength could be in the awareness the experiment offers in that ESL individuals have a more difficult time with English-based assessments and often suffer a stereotype of being seen as having far less intelligence than the standardized average.

Another strength of this test is that the reading excerpt was taken from a book that is not widely known, is not required reading, no royalties were required for utilizing it, and no downloadable version could easily be found. The reading excerpt was also split into three versions and each was customized for usage in this experiment, bearing only some resemblance to its original passage. This meant that the students were limited, if not

restricted, to the information they read and could not find the passage, even if they had the book. Additionally, the questions were also customized for each passage along with the answers, so the participants had no way of knowing the right answers unless the passage was read and could be either recalled or recognized by memory. Moreover, the random assignment of the passage and following questions decreased the likelihood of participants taking the same exact test as another participant and preventing the student from taking the same test should they decide to exit and re-take for improved grade.

Within the test, a trend was noticed halfway through the experiment that participants seemed to quit the study after reading the prompt and without answering any questions, citing a technical issue where the button that allowed participants to proceed past the reading would not appear. This led to half the first 135 participants' data being unusable. Upon investigation, it became apparent that no such technical issue existed, but rather, the participants neglected to read the prompt warning of a four-minute waiting period that came on a page immediately before proceeding to page where the reading excerpt was. This led to a second prompt being placed at the bottom of the reading warning of the four-minute waiting period, but before where the button to proceed would appear.

A limitation of the study, although serving as another purpose of the study, is that it is entirely based on stereotypes and gendered language found within English, which can be different to those found in other gendered languages (i.e. French, Italian, Portuguese, Spanish). This is especially troublesome for individuals classified as ESL as they cannot always comprehend the stereotypes and language found in readings and

excerpts in a timely manner, if ever, when compared to those who are not ESL, which could potentially cause lower test scores.

Another limitation comes in the form of the population that was gathered for this assessment. The assessment was limited to the undergraduate student population of the Department of Psychology at Texas State University. The population is predominantly comprised of Caucasian and Hispanic/Latinx/Chicanx students. Most students in the college population of central Texas in general only know one or two languages, which usually ends up being Spanish (i.e. another gendered language), meaning limitations on bilingualism, a great lack of multilinguism, and limitations on having ESL participants whose first language is gender-neutral for comparison.

A gigantic limitation upon this study is that, while gendered language studies have been conducted for three decades and have been conducted in an educational perspective for at least the two decades, gender neutrality in language was only made popular in the middle of the 20th century and has still been slow to garner attention until recently. Consequently, there only exists a few assessments to properly gauge gendered language as a whole and even fewer to assess the comprehensive effects of gendered language in a testing environment. Therefore, the foundation in building from previous research is firm, but not strong or completely cohesive, which means more research needs to be done.

A further limitation of the study was the lack of similar environment to taking a standardized test. Students could take this survey on their personal computer or their phone, could potentially take the test anywhere with an internet connection, and were not technologically blocked from using other applications or services on their devices. This

means they were not restricted from sharing answers with other test-takers, nor were they restricted in taking a screenshot of reading.

Directions for Future Research

As mentioned above, the rising trend of transcendent attitudes in education towards gender affecting the capability to learn and an individual's memory is only a theoretical assessment based on this experiment and previous experiments and literature. To fully understand this and add validity to the statement, the next experiment that could be done should be one of longitudinal design. As mentioned above, there is virtually no consistent structure in education revolving around gendered language and gender-neutral language. This potential experiment would involve two groups to directly affect and one control group over four points in education: elementary/primary level, middle/intermediate level, high/secondary level, and collegiate/university level. For maximum coverage of the subject, the experiment should be done simultaneously in several different labs in several areas, with several areas grouped together by similarity of demographics and/or styles of education.

A similar experiment should be done to facilitate the testing environment the average student finds themselves in whenever they are required to take a standardized test. This could be constructed in a small computer lab or classroom with multiple participants, timing restrictions for each section, and restrictions on technology on the use of personal technology.

One option for future research is that of designing a similar standardized test that utilizes another gendered language, such as Spanish or French, in order to test if that same results can be found outside of an English-based test. Following that, another option

is creating a test of inverted design utilizing a gender-neutral language (e.g. Armenian, Chinese, Estonian, Finnish, Hungarian, Japanese, Korean, Persian, Thai) and its subsequent stereotypes to explore if there are relevant results. Consequently, a following analysis to compare and contrast should be done on all three experiments to assess limitations and improvements in order to create better versions of assessments separately or together. A separate study from this could potentially be conducted into exclusively studying languages that carry three genders (i.e. Afrikaans, Belarusian, Dutch, German, Greek, Russian) or more than three genders (i.e. Chechen, Czech, Polish, Swahili).

Another improvement in a possible future study should seek to reduce the difference in number of participants between genders. As aforementioned in the previous section, there were only 12 males and 82 females in this study, which may have added to the absence of statistical power in measuring differences between genders. Additionally, the current study's sample was not very diverse in ethnicity (e.g. 75.6% of the participants registering as Caucasian or Hispanic/Latinx/Chicanc), which might influence the results' generalizability. Moreover, 76.6 % of the participants registered that English was their first/native language, which definitively influences the findings of the experiment. Future research ought to pursue a more varied sample and take into account evaluating ethnic differences in language assessment in regards to gendered language.

Summary and Conclusion

Despite these limitations, this study delivers helpful understanding of the dynamics between language, gender, and standardized tests in a college sample. Specifically, the findings of this study demonstrate that a population that has been strictly taught gendered language has increased trouble comprehending, recalling, or correctly

recognizing information offered through gender-neutral language, but has no trouble with comprehension and memory recall/recognition in gendered language-based tests. As such, those who are not used to gender-neutral language and those who are not used to gendered language that is inconsistent with gender norms will have test scores that suffer because of this.

APPENDIX SECTION

APPENDIX A

Demographic Survey

Instructions: Please respond to the following demographic questions.

1. What is your sex?
    a. Male
    b. Female

2. What is your ethnicity?
    a. Caucasian
    b. Hispanic/Latinx/Chicanx
    c. African American
    d. Asian American, Middle Eastern, or Pacific Islander
    e. Native American
    f. Biracial or Multiracial
    g. Other

3. What is your age? _____

4. Is English your first/native language?
    a. Yes
    b. No
    c. I learned English at the same time as another language.

5. What is your academic rank?
    a. Freshman
    b. Sophomore
    c. Junior
    d. Senior

6. What is your estimated current grade point average (GPA) in college? _____

APPENDIX B

Reading Passages

Gendered Language, Stereotype-Consistent Condition

I remember one summer when I was caring for a neighboring family of three that used to be a family of four. The husband, George, was a 45-year-old Italian man who had everything going for him. Before, he was a successful engineer, and had an amazing wife, Amanda, and two wonderful children, Jacob and Lily. However, he made some mistakes that led to imprisonment. He developed a gambling problem, which also came with a drinking problem. Both of these began to increase and led to an incident that his family rarely spoke of.

I can recall a string of events on the day when I took his wife and two children to visit him in the Illinois State Penitentiary. When we finally saw him, his wife, much to my annoyance, talked about nothing but his unfashionable striped clothing and his overall unkempt appearance. His son couldn't be bothered by any of it and just asked when he was going to play soccer with him again. Meanwhile, his daughter spoke very little, seemingly frightened by the prison atmosphere and clutching her favorite doll to her chest.

On our return to Chicago, Amanda wanted to stop by a grocery store to pick up a few things for her shift at the hospital. While she frequently complained to me about many annoyances of being a nurse, she was an overall compassionate woman who loved helping her patients. Anyway, when the family was walking out of the store, the son presented his sister with two oranges. When his sister refused the oranges, not wanting to potentially dirty her outfit, he decided to eat one of them, and he asked me to play catch with the other one. Just then, a stock boy came out of the store and ran after the family, violently ripping the stolen fruit from the son's hands and threatening to arrest the child, a mindless threat which was a bit ironic given that the son just visited his father in prison.

Stranger than any of the day's events was the fact that no one for a moment considered George a criminal. He had merely gotten too excited over a game cards and had stabbed his opponent with a fork. "Why should a man who lost his temper be kept from his family forever?" was his family's reiterated inquiry.

Gendered Language, Stereotype-Inconsistent Condition

I remember one summer when I was caring for a neighboring family of three that used to be a family of four. The wife, Amanda, was a 45-year-old Italian woman who had everything going for her. Before, she was a successful engineer, and had an amazing husband, George, and two wonderful children, Jacob and Lily. However, she made some mistakes that led to imprisonment. She developed a gambling problem, which also came with a drinking problem. Both of these began to increase and led to an incident that her family rarely spoke of.

I can recall a string of events on the day when I took her husband and two children to visit her in the Illinois State Penitentiary. When we finally saw her, her husband, much to my annoyance, talked about nothing but her unfashionable striped clothing and her overall unkempt appearance. Her daughter couldn't be bothered by any of it and just asked when she was going to play soccer with her again. Meanwhile, her son spoke very little, seemingly frightened by the prison atmosphere and clutching his favorite doll to his chest.

On our return to Chicago, George wanted to stop by a grocery store to pick up a few things for his shift at the hospital. While he frequently complained to me about many annoyances of being a nurse, he was an overall compassionate man who loved helping his patients. Anyway, when the family was walking out of the store, the daughter presented her brother with two oranges. When her brother refused the oranges, not wanting to potentially dirty his outfit, she decided to eat one of them, and she asked me to play catch with the other one. Just then, a stock girl came out of the store and ran after the family, violently ripping the stolen fruit from the daughter's hands and threatening to arrest the child, a mindless threat which was a bit ironic given that the daughter just visited her mother in prison.

Stranger than any of the day's events was the fact that no one for a moment considered Amanda a criminal. She had merely gotten too excited over a game cards and had stabbed her opponent with a fork. "Why should a woman who lost her temper be kept from her family forever?" was her family's reiterated inquiry.

Gender Neutral Language Condition

I remember one summer when I was caring for a neighboring family of three that used to be a family of four. One parent, Skyler, was a 45-year-old Italian who had everything going right. Before, Skyler was a successful engineer, and had an amazing spouse, Taylor, and two wonderful children, Oakley and Justice. However, Skyler made some mistakes that led to imprisonment. Skyler developed a gambling problem, which also came with a drinking problem. Both of these began to increase and led to an incident that the family rarely spoke of.

I can recall a string of events on the day when I took the spouse and two children to visit Skyler in the Illinois State Penitentiary. When we finally saw Skyler, the spouse, much to my annoyance, talked about nothing but Skyler's unfashionable striped clothing and overall unkempt appearance. The child, Oakley, couldn't be bothered by any of it and just asked when Skyler was going to play soccer with Oakley again. Meanwhile, the other child, Justice, spoke very little, seemingly frightened by the prison atmosphere and clutching a favorite doll to the chest.

On our return to Chicago, Taylor wanted to stop by a grocery store to pick up a few things for the next shift at the hospital. While Taylor frequently complained to me about many annoyances of being a nurse, Taylor was an overall compassionate person who loved helping the patients. Anyway, when the family was walking out of the store,

Oakley presented Justice with two oranges. When Justice refused the oranges, not wanting to potentially dirty Justice's outfit, Oakley decided to eat one of the oranges, and asked me to play catch with the other one. Just then, a stock person came out of the store and ran after the family, violently ripping the stolen fruit from the child's hands and threatening to arrest the child, a mindless threat which was a bit ironic given that the kid just visited a parent in prison.

Stranger than any of the day's events was the fact that no one for a moment considered Skyler a criminal. Skyler had merely gotten too excited over a game cards and had stabbed the opponent with a fork. "Why should a person who lost their temper be kept from their family forever?" was the family's reiterated inquiry.

APPENDIX C

Memory Tests

Gendered Language, Stereotype-Consistent Condition	**Gendered Language, Stereotype-Inconsistent Condition**	**Gender Neutral Language Condition**
What does Amanda do for a living? a. Architect b. Engineer c. Hospital administrator **d. Nurse**	What does George do for a living? a. Architect b. Engineer c. Hospital administrator **d. Nurse**	What does Taylor do for a living? a. Architect b. Engineer c. Hospital administrator **d. Nurse**
What did Jacob talk about during the visit to the prison? **a. Playing soccer** b. Reasons for imprisonment c. The frightening prison atmosphere d. The striped outfits of the prisoners	What did Lily talk about during the visit to the prison? **a. Playing soccer** b. Reasons for imprisonment c. The frightening prison atmosphere d. The striped outfits of the prisoners	What did Oakley talk about during the visit to the prison? **a. Playing soccer** b. Reasons for imprisonment c. The frightening prison atmosphere d. The striped outfits of the prisoners
Who has a drinking problem? a. Amanda **b. George** c. Jacob d. Lily	Who has a drinking problem? **a. Amanda** b. George c. Jacob d. Lily	Who has a drinking problem? a. Justice b. Oakley **c. Skyler** d. Taylor
In which state does the story take place? **a. Illinois** b. New York c. Pennsylvania d. Texas	In which state does the story take place? **a. Illinois** b. New York c. Pennsylvania d. Texas	In which state does the story take place? **a. Illinois** b. New York c. Pennsylvania d. Texas

Who stole something from the grocery store? a. Amanda b. George **c. Jacob** d. Lily	Who stole something from the grocery store? a. Amanda b. George c. Jacob **d. Lily**	Who stole something from the grocery store? a. Justice **b. Oakley** c. Skyler d. Taylor
Who was frightened by the prison atmosphere? a. Amanda b. George c. Jacob **d. Lily**	Who was frightened by the prison atmosphere? a. Amanda b. George **c. Jacob** d. Lily	Who was frightened by the prison atmosphere? **a. Justice** b. Oakley c. Skyler d. Taylor
What is George's ancestry? a. French b. German **c. Italian** d. Spanish	What is Amanda's ancestry? a. French b. German **c. Italian** d. Spanish	What is Skyler's ancestry? a. French b. German **c. Italian** d. Spanish
What does George do for a living? a. Architect **b. Engineer** c. Hospital administrator d. Nurse	What does Amanda do for a living? a. Architect **b. Engineer** c. Hospital administrator d. Nurse	What does Skyler do for a living? a. Architect **b. Engineer** c. Hospital administrator d. Nurse
What was Lily carrying during the visit to the prison? a. A deck of cards **b. A doll** c. A soccer ball d. A stuffed animal	What was Jacob carrying during the visit to the prison? a. A deck of cards **b. A doll** c. A soccer ball d. A stuffed animal	What was Justice carrying during the visit to the prison? a. A deck of cards **b. A doll** c. A soccer ball d. A stuffed animal

Who has a gambling problem? a. Amanda **b. George** c. Jacob d. Lily	Who has a gambling problem? **a. Amanda** b. George c. Jacob d. Lily	Who has a gambling problem? a. Justice b. Oakley **c. Skyler** d. Taylor
In which season does the story take place? a. Spring **b. Summer** c. Fall d. Winter	In which season does the story take place? a. Spring **b. Summer** c. Fall d. Winter	In which season does the story take place? a. Spring **b. Summer** c. Fall d. Winter
Who is described as a compassionate person who loves helping others? **a. Amanda** b. George c. Jacob d. Lily	Who is described as a compassionate person who loves helping others? a. Amanda **b. George** c. Jacob d. Lily	Who is described as a compassionate person who loves helping others? a. Justice b. Oakley c. Skyler **d. Taylor**
What was stolen from the grocery store? a. A candy bar b. A soda c. Apples **d. Oranges**	What was stolen from the grocery store? a. A candy bar b. A soda c. Apples **d. Oranges**	What was stolen from the grocery store? a. A candy bar b. A soda c. Apples **d. Oranges**
In which city does the family live? **a. Chicago** b. Dallas c. Pittsburg d. New York City	In which city does the family live? **a. Chicago** b. Dallas c. Pittsburg d. New York City	In which city does the family live? **a. Chicago** b. Dallas c. Pittsburg d. New York City

What did Amanda talk about during the visit to the prison? a. Playing soccer b. Reasons for imprisonment c. The frightening prison atmosphere **d. The striped outfits of the prisoners**	What did George talk about during the visit to the prison? a. Playing soccer b. Reasons for imprisonment c. The frightening prison atmosphere **d. The striped outfits of the prisoners**	What did Taylor talk about during the visit to the prison? a. Playing soccer b. Reasons for imprisonment c. The frightening prison atmosphere **d. The striped outfits of the prisoners**
How old is George? a. 30 b. 35 c. 40 **d. 45**	How old is Amanda? a. 30 b. 35 c. 40 **d. 45**	How old is Skyler? a. 30 b. 35 c. 40 **d. 45**
Who didn't want to get their outfit dirty? a. Amanda b. George c. Jacob **d. Lily**	Who didn't want to get their outfit dirty? a. Amanda b. George **c. Jacob** d. Lily	Who didn't want to get their outfit dirty? **a. Justice** b. Oakley c. Skyler d. Taylor
How does the story's narrator know the family? a. Co-worker b. Cousin **c. Neighbor** d. Parent	How does the story's narrator know the family? a. Co-worker b. Cousin **c. Neighbor** d. Parent	How does the story's narrator know the family? a. Co-worker b. Cousin **c. Neighbor** d. Parent
Who threatened to arrest the person who stole something from the grocery store? a. A cashier b. A security guard **c. A stock person** d. The store manager	Who threatened to arrest the person who stole something from the grocery store? a. A cashier b. A security guard **c. A stock person** d. The store manager	Who threatened to arrest the person who stole something from the grocery store? a. A cashier b. A security guard **c. A stock person** d. The store manager

What is the reason for George's imprisonment?	What is the reason for Amanda's imprisonment?	What is the reason for Skyler's imprisonment?
a. Stabbing someone with a fork b. Stabbing someone with a knife c. Stealing something from work d. Tax fraud	**a. Stabbing someone with a fork** b. Stabbing someone with a knife c. Stealing something from work d. Tax fraud	**a. Stabbing someone with a fork** b. Stabbing someone with a knife c. Stealing something from work d. Tax fraud

APPENDIX D

Social Roles Questionnaire

Instructions: Below are a number of statements covering attitudes on social roles. Please indicate, using the scale provided, your level of agreement, disagreement, or neutrality.

Scale:
1 – strongly disagree
2 – disagree
3 – neutral
4 – agree
5 – strongly agree

Items:
1. People can be both aggressive and nurturing regardless of sex.
2. People should be treated the same regardless of sex.
3. The freedom that children are given should be determined by their age and maturity level and not by their sex.
4. Tasks around the house should not be assigned by sex.
5. We should stop thinking about whether people are male or female and focus on other characteristics.
6. A father's major responsibility is to provide financially for his children.
7. Men are more sexual than women.
8. Some types of work are just not appropriate for women.
9. Mothers should make most decisions about how children are brought up.
10. Mothers should work only if necessary.
11. Girls should be protected and watched over more than boys.
12. Only some types of work are appropriate for both men and women.
13. For many important jobs, it is better to choose men instead of women.

Key:
Gender transcendent items: 1, 2, 3, 4, 5
Gender traditional items: 6, 7, 8, 9, 10, 11, 12, 13

REFERENCES

Abedi, J., Leon, S., & Mirocha, J. (2001). *Examining ELL and non-ELL student performance differences and their relationship to background factors: Continued analyses of extant data.* Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*, 219–234. doi:10.1207/S15324818AME1403_2

Arthur, A. E., Bigler, R. S., Liben, L. S., Gelman, S. A., & Ruble, D. N. (2008). Gender stereotyping and prejudice in young children: A developmental intergroup perspective. In S. R. Levy & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 66–86). Oxford, UK: Oxford University Press.

Baber, K. M., & Tucker, C. J. (2006). The social roles questionnaire: A new approach to measuring attitudes toward gender. *Sex Roles*, *54*, 459–467. Doi: 10.1007/s11199-006-9018-y

Bauer, P. J., & Coyne, M. J. (1997). When the name says it all: Preschoolers' recognition and use of the gendered nature of common proper names. *Social Development*, *6*, 271–291. doi:10.1111/j.1467-9507.1997.tb00106.x

Bigler, R., & Leaper, C. (2015). Gendered language: Psychological principles, evolving practices, and inclusive policies. *Policy Insights from the Behavioral and Brain Sciences*, *2*, 187–194. doi: 10.1177/2372732215600452.

Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. *Advances in Child Development and Behavior*, *34*, 39–89. doi:10.1016/S0065-2407(06)80004-2

Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science*, *16*, 162–166. doi:10.1111/j.1467-8721.2007.00496.x

Boroditsky, L. (2011). How languages construct time. In S. Dehaene & E. Brannon (Eds.), *Space, time and number in the brain: Searching for the foundations of mathematical thought* (pp. 333–342). San Diego, CA: Elsevier.

Cimpian, A., Mu, Y., & Erickson, L. C. (2012). Who is good at this game? Linking an activity to a social category undermines children's achievement. *Psychological Science*, *23*, 533–541. doi:10.1177/0956797611429803

Coates, J. (1985). Women, men, and language: A sociolinguistic account of gender differences in language. London: Longman.

de Klerk, V., & Bosch, B. (1996). Nicknames as sex-role stereotypes. *Sex Roles*, *35*, 525–544. doi:10.1007/BF00290049

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children). *Child Development*, *82*(3), 796–811. doi: 10.1111/j.1467-8624.2011.01577.x

Duran, R. P. (1989). Assessment and instruction of at-risk Hispanic students. *Exceptional Children*, *56*, 154–158. doi:10.1177/001440298905600207

Fagot, B. I., Leinbach, M. D., & Hagan, R. (1986). Gender labeling and the adoption of sex-typed behaviors. *Developmental Psychology*, *22,* 440–443. doi:10.1037/0012-1649.22.4.440

Fowler, H.W. (2015). *Fowler's dictionary of modern English usage.* New York, NY: Oxford University Press.

Freeman, R., & McElhinny, B. (1996). Language and gender. In S. Mckay & N. Hornberger (Eds.), *Sociolinguistics and language teaching* (pp. 218-280). Cambridge: Cambridge University Press.

Friedman, C. K., Leaper, C., & Bigler, R. S. (2007). Do mothers' gender-related attitudes or comments predict young children's gender beliefs? *Parenting: Science and Practice, 7,* 357–366. doi:10.1080/15295190701665656

Garcia, G. E. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. *Reading Research Quarterly*, *26,* 371–391. doi:10.2307/747894

Gelman, S. A. (2003). *The essential child*. New York, NY: Oxford University Press.

Gelman, S. A., Collman, P., & Maccoby, E. E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development, 57,* 394–404. doi:10.2307/1130595

Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science, 10,* 489–493. doi:10.1111/1467-9280.00194

Gelman, S. A., Taylor, M. G., & Nguyen, S. P. (2004). Mother-child conversations about gender. *Monographs of the Society for Research in Children Development, 69*, 1–145. doi:10.1111/j.1540-5834.2004.06901002.x

Gentner, D., & Loewenstein, J. (2002). Relational language and relational thought. In E. Asel & J. P. Byrnes (Eds.), *Language, literacy, and cognitive development: The development and consequences of symbolic communication* (pp. 87–120). Mahwah, NJ: Lawrence Erlbaum.

Gordon, C. (2011). Gumperz and interactional sociolinguistics. In R. Wodak, B. Johnstone, & P. Kerswill (Eds.), *The Sage handbook of sociolinguistics* (pp. 67-84). Los Angeles: Sage.

Hassel, B., & Rosch, J. (2008) *Ohio value-added primer*. Washington, DC: Fordham Foundation.

Henley, N. M. (1989). Molehill or mountain? What we know and don't know about sex bias in language. In M. Crawford & M. Gentry (Eds.), *Gender and thought: Psychological perspectives* (pp. 59–78). New York, NY: Springer.

Hidden curriculum. (2014). In S. Abbott (Ed.), *The glossary of education reform*. Retrieved on December 10, 2016, from http://edglossary.org/hidden-curriculum

Hilliard, L. J., & Liben, L. S. (2010). Differing levels of gender salience in preschool classrooms: Effects on children's gender attitudes and intergroup bias. *Child Development, 81*, 1787-1798. doi:10.1111/j.1467–8624.2010.01510.x

Key, M. R. (1975). *Male/female language.* Metuchen, NJ: The Scarecrow Press.

Khorsoroshahi, F. (1989). Penguins don't care, but women do: A social identity analysis of a Whorfian problem. *Language in Society, 18,* 505–525. doi:10.1017/S0047404500013889

Kuncel, N. R., & Hezlett, S. A. (2007). Assessment: Standardized tests predict graduate students' success. *Science, 315,* 1080–1081. doi:10.1126/science.1136618

Lakoff, R. (1975). *Language and woman's place: Text and commentaries.* New York, NY: Oxford University Press.

Leaper, C. (2014). Gender similarities and differences in language. In T. Holtgraves (Ed.), *Oxford handbook of language and social psychology* (pp. 62-81). New York, NY: Oxford University Press.

Leaper, C., & Bigler, R. S. (2004). Gendered language and sexist thought. *Monographs of the Society for Research in Child Development, 69,* 128–142. doi:10.1111/j.1540-5834.2004.06901012.x

Lee, J. (1989). Teacher expectations: Self-fulfilling prophecies, perceptual bias, and accuracy. *Journal of Personality and Social Psychology, 57,* 469–480. doi:10.1037/0022-3514.57.3.469

Liben, L. S., & Bigler, R. S. (2015). Understanding and undermining the development of gender dichotomies: The legacy of Sandra Lipsitz Bem. *Sex Roles, 76,* 544–555. doi:10.1007/s11199-015-0519-4

Liben, L. S., Bigler, R. S., & Krogh, H. R. (2002). Language at work: Children's gendered interpretations of occupational titles. *Child Development, 73,* 810-828. doi:10.1111/1467-8624.00440Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179-195). New York, NY: Cambridge University. doi:10.1017/CBO9780511529863.009

Miner, B. (2000). Standardized minds: The high price of America's testing culture and what we can do to change it//Contradictions of school reform: Educational costs of standardized testing. *The Progressive, 64,* 40–43.

Mulac, A., Bradac, J. J., & Karol-Mann, S. (1985). Male/Female language differences and attributional consequences in children's television. *Human Communication Research, 11,* 481–506.

Neill, M. (2009). *Standardized tests are unfair and harmful.* Detroit, MI: Greenhaven Press.

Olson, A., & Sabers, D. (2008). Standardized tests. In T. L. Good (Ed.), *21st century education: A reference handbook* (pp. 423-431). Los Angeles: Sage.

Pennock-Roman, M., & Rivera, C. (2011). Mean effects of test accommodations for ELLs and Non-ELLs: A meta-analysis of experimental studies. *Educational Measurement: Issues and Practice, 30,* 10–28. doi: 10.1111/j.1745-3992.2011.00207.x

Petterson, L. J., Dixson, B. J., Little, A. C., & Vasey, P. L. (2016). Reconsidering male bisexuality: Sexual activity role and sexual attraction in Samoan men who engage in sexual interactions with Fa'afafine. *Psychology of Sexual Orientation and Gender Diversity, 3,* 11–16. doi:10.1037/sgd0000160

Popham, W.J. (1999). Why standardized tests don't measure educational quality. *Educational Leadership, 56*(6), 8–15.

Popham, W. J. (2016). Standardized tests: Purpose is the point. *Educational Leadership, 73*(7), 44–49.

Porter, A., McMaken, J., Hwang, J., & Yang, R. (2011). Common core standards: The new U.S. intended curriculum. *Educational Researcher, 40*(7), 103–116. doi:10.3102/0013189X11405038

Ravitch, D. (1985). *The schools we deserve*. New York, NY: Basic Books.

Rice, P. (2000). Gendered readings of a traditional "feminist" folktale by sixth-grade boys and girls. *Journal of Literacy Research, 32,* 211-236. doi:10.1080/10862960009548074

Rennels, J. L., & Langlois, J. H. (2014). Children's classification and lexicalization of attractiveness, gender, and race: Differential displays of these concepts and relatedness to bias and flexibility. *Journal of Experimental Child Psychology, 126,* 1–18. doi: 10.1016/j.jecp.2014.02.009

Stahlberg, D., Braun, F., Irmen, L., & Sczesny, S. (2007). Representations of the sexes in language. In K. Fiedler (Ed.), *Social communication* (pp. 163-187). New York, NY: Psychology Press.

Stringer, J. A. C. (2011). *Trans* and queer terms/LGBTQPIA terminology. Heartland Trans Wellness Group*. Retrieved from http://transwellness.org/resources/educational-materials/trans-and-queer-terms/

Tannen, D. (1990). *You just don't understand: Women and men in conversation.* New York, NY: Harper Collins.

Thomson, R., Murachver, T., & Green, J. (2001). Where is the gender in gendered language? *Psychological Science, 12*, 171–175. doi:10.1111/1467-9280.00329

Thorne, B., & Henley, N. (1975). *Language and sex: Difference and dominance.* Rowley, MA: Newbury House Publishers.

Van Fleet, D. D., & Atwater, L. (1997). Gender neutral names: Don't be so sure! *Sex Roles, 37*, 111–123. doi:10.1023/A:1025696905342

Vervecken, D., Hannover, B., & Wolter, I. (2013). Changing (S) expectations: How gender fair job descriptions impact children's perceptions and interest regarding traditionally male occupations. *Journal of Vocational Behavior, 82,* 208-220. doi:10.1016/j.jvb.2013.01.008

Waxman, S. R. (2013). Building a better bridge. In M. Baraji, S. Gelman, & S. Lehr (Eds.), *Navigating the social world: The early years* (pp. 292-296). New York, NY: Oxford University Press.

Welch-Ross, M. K., & Schmidt, C. R. (1996). *Gender-schema development and children's constructive story memory: Evidence for a developmental model. Child Development, 67,* 820-835. doi:10.1111/j.1467-8624.1996.tb01766.x