8-2022

Social Comparison and Shifting in Quantitative Performance

Ashlee Pardo
CSUSAN BERNARDINO

SOCIAL COMPARISON AND SHIFTING IN QUANTITATIVE PERFORMANCE

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

in

Psychological Science

---

by

Ashlee Pardo

August 2022

SOCIAL COMPARISON AND SHIFTING IN QUANTITATIVE PERFORMANCE

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

---

by

Ashlee Pardo

August 2022

Approved by:

Donna Garcia, Committee Chair, Psychology

Joseph Wellman, Committee Member

Jodie Ullman, Committee Member

ABSTRACT

Research has produced conflicting results in terms of whether positive stereotypes enhance or impair the performance of individuals who have both positively- and negatively-stereotyped identities. Shih, Pittinsky, and Ambady (1999) found that relative to a no identity control group, Asian women performed better on a mathematics test when their positively-stereotyped identity was salient (i.e., ethnicity) but worse when their negatively-stereotyped identity was salient (i.e., gender). In contrast, Cheryan and Bodenhausen (2000) found that both positive and negative identity salience in Asian women led to reduced performance relative to a no identity control group. In the current study, I aimed to examine a potential reason for this discrepant finding. In a sample of Latino men, I further examined how cues that make people aware of their gender or ethnic identities might influence their performance on a mathematics test, depending on whether the salient identity was positively or negatively stereotyped in mathematics. I argued that the discrepancy in findings between Shih et al. and Cheryan and Bodenhausen was due to differences in the manipulation they used, that is, the former induced intergroup comparisons and the latter induced intragroup comparisons. I expected that gender salience using the “Shih” method would lead to better performance on a mathematics test whereas gender salience with the “Cheryan” method would lead to worse performance. Results can have important implications for understanding the

consequences of positive stereotypes and their effect, depending on the direction of the comparison.

Keywords: positive stereotypes, intergroup comparison, intragroup comparison, identities

TABLE OF CONTENTS

CHAPTER ONE

SOCIAL COMPARISON AND SHIFTING IN QUANTITATIVE PERFORMANCE

Stereotypes allow people to swiftly form beliefs or make generalizations based on traits and characteristics that social groups are believed to have (Lambert, Khan, Lickel, & Fricke, 1997). Stereotypes become culturally pervasive over time because people use them as a heuristic that makes it possible to quickly process new information, especially during a state of unfamiliarity (Croizet & Claire, 1998). For example, blondes are unintelligent, women are bad drivers, and Muslims are terrorists are all examples of pervasive stereotypes. Despite their prevalence, stereotypes are often based on culturally-shared assumptions, and more often than not, these assumptions are inaccurate (Judd & Park, 1993). The pervasiveness of unfavorable stereotypes about one's group can have negative consequences. For example, Belmi, Barragan, Neale, and Cohen (2015) reported that when people felt they were devalued simply because they belonged to particular groups, they were more likely to act in a socially deviant way such as lie, cheat, and steal. The researchers also found that the more college students were worried about being seen negatively because of their ethnicity, the more likely they were to report participating in delinquent behavior such as verbally abusing someone, skipping classes, or vandalizing school property.

Not only do negative stereotypes affect behavior, they can negatively impair intellectual performance as well (see Spencer, Logel, & Davies, 2016). Steele (1997) referred to this phenomenon as stereotype threat. In the first published study on this phenomenon, Steel and Aronson (1995) reported that when African and European American students were led to believe that a verbal test was diagnostic of ability, the African American students performed much poorer relative to the European American students. This performance difference was not found when the students were told the test was not diagnostic of ability. According to the researchers, the diagnostic (but not the non-diagnostic condition) reminded African American students about the negative stereotypes concerning their ability. Similarly, Spencer, Steele, and Quinn (1999) found that women scored much lower on a math test relative to men when they were led to believe that the test had previously shown gender differences. Women and men scored the same, however, when they were told the test was gender neutral. Results from these past studies suggest that when people know of a negative stereotype about their group and become aware of their group membership, they become burdened by fear of fulfilling the negative stereotype (Schmader, Johns, & Forbes, 2008).

Although negative stereotypes can impair intellectual performance, there is evidence that positive stereotypes can actually enhance it (Walton & Cohen, 2003; Shih, Pittinsky, & Ambady, 1999). According to Walton and Cohen (2003), members of non-stereotyped groups may benefit from “stereotype lift” when the

ability or worth of a particular negatively stereotyped out-group is specifically called into question (see also Inzlicht & Schmader, 2012, p. 143). In a meta-analytic review, Walton and Cohen (2003) found evidence supporting the effect of stereotype lift. Across 43 studies, members of non-stereotyped groups (e.g., White Americans and men) were found to perform better when a negative stereotype about a particular out-group (e.g., Black Americans and women) was tied to an intellectual test than when it was not. Walton and Cohen argue that this lift effect happens when downward comparisons are made with out-groups that are negatively stereotyped in the relevant performance domain. The researchers contend that because downward social comparisons with socially devalued out-group members can enhance self-esteem (Morse & Gergen, 1970) and self-efficacy (Bandura, 1986), these comparisons can improve performance by reducing performance-related stress/anxiety (Wills, 1981). Walton and Cohen's explanation presumes that only negative stereotypes about an out-group rather than positive stereotypes of one's in-group are relevant to performance; however, other research suggests that positive stereotypes can have performance benefits.

Shih and colleagues (1999) examined whether mathematical performance differed among Asian-American women, depending upon whether they were in a no-identity salient control condition, gender salient condition, or ethnicity salient condition. The researchers found that participants in the gender salient condition performed worse on the mathematics test than those in the control condition. In

other words, when gender was salient, participants performed in line with negative stereotypes about women's quantitative ability. However, participants in the ethnicity salient condition performed better on the mathematics test than those in the control condition. That is, when ethnicity was salient, participants performed consistently with positive stereotypes of Asians' quantitative ability. This effect is known as stereotype boost (see Shih, Pittinsky & Ho, 2012), which is similar to stereotype lift in terms of performance enhancement. Boost occurs, however, when individuals are exposed to positive stereotypes about their social in-group (rather than negative stereotypes of an out-group). In the current study, I will be focusing on stereotype boost and not stereotype lift because Cheryan and Bodenhausen (2000) found that both positive and negative identity salience in Asian American women led to reduced performance relative to a no identity control group.

Shih, Pittinsky, and Ho (2012) offer some possible explanations for why research on the effects of positive stereotypes have produced different findings. One proposition is that whether positive stereotypes impair or boost performance might depend on whether the positive identity is subtly or blatantly made salient. Shih and colleagues (2012) argue that Cheryan and Bodenhausen's (2000) manipulation of the positively-stereotyped identity was blatant, resulting in negative performance outcomes; whereas, Shih et al.'s manipulation was subtle, resulting in positive performance outcomes. This proposition, however, is only speculative because as far as I am aware, no study has directly manipulated

whether the positive identity was blatantly or subtly made salient. Without that manipulation in a single study, it is impossible to assess whether the nature of the salience manipulation mattered.

I propose to examine one other potential reason for the discrepancy between the Shih et al. (1999) and Cheryan and Bodenhausen (2000) findings: differences in the comparison group used by participants in response to the different identity salient manipulations in the two different papers. According to social comparison theory, we compare ourselves to others that share similar characteristics to us so that we can evaluate ourselves as people (Guimond, 2006). There are two types of social comparison: downward comparison and upward comparison. Downward comparison can be defined as a comparison to someone we find less than or inferior to ourselves. By looking downward at someone else, we can evaluate ourselves at a higher social ranking. Therefore, if the comparison group target is supposed to be less capable than our own group, then the issue of the comparison may be positive and our own performance may increase. Konan, Chatard, Selimbegović, Mugn, and Moraru (2011) found that female European students performed better on a math test when they were led to compare with a female immigrant rather than with another female European student. In contrast to downward comparison, upward comparison can be defined as comparing ourselves to someone we believe is better than or superior to us. Therefore, if the comparison group target is supposed to be more capable than our own group, then the consequence of the comparison may be negative

and our performance may decrease. Many previous stereotype threat studies have shown a decrease in performance of low-status group members when they are comparing upward with high-status group members, whereas they perform just as well as these latter individuals when the upward comparison piece is removed (for a review, see Steele, Spencer, & Aronson, 2002). I argue that because the Shih et al. (1999) ethnic salient manipulation required Asian American women to think of how their ethnic group differed from the dominant ethnic group (i.e., White Americans), participants were led to focus on how their ethnic group differed from White Americans. Thus, participants were likely using intergroup comparisons when thinking about their mathematics capability. This comparison would be downward because Asian Americans are perceived as superior in mathematics ability to White Americans (Wills, 1981). Thus, Asian participants would have had high relative performance expectations, which could then enhance performance (see Wills, 1981). In contrast, because the Cheryan and Bodenhausen (2000) manipulation had Asian American women think of themselves relative to other members of their ethnic group, their manipulation likely led participants to make intragroup comparison and compare themselves to other Asians. Because Asians' mathematics ability is positively stereotyped, the comparison would be either lateral or upward in terms of self-performance expectations. Upward social comparisons tend to have the opposite effect of downward comparisons; rather than increasing performance expectations and actual performance, they undermine relative performance expectation and impair

actual performance (Blanton, Buunk, Gibbons, & Kuyper, 1999). In sum, when positive stereotypes are made salient in the context of a downward social comparison, they can lead to performance enhancement. In contrast, when positive stereotypes are made salient in the context of an upward social comparison, they can lead to performance impairment.

Study Overview

In the current study, I aim to examine how cues that make people aware of their gender or ethnic identities might influence their performance on a mathematics test depending on whether their identity is positively or negatively stereotyped in mathematics, and whether they are induced to make intergroup or intragroup comparisons. Because of the low Asian American women enrollment at California State University, San Bernardino (CSUSB), I will use Latino men as my participants. Similar to Asian American women, Latino men have both positive and negative stereotypes about their mathematics ability. The stereotype for men's ability is positive (Nosek, Banaji, & Greenwald, 2002) but the stereotype for Latino is negative (Gonzales, Blanton, & Williams, 2002). Thus, Latino men should experience stereotype boost and do better on a mathematics test when they are led to think in terms of their gender rather than their ethnic identity, but only if they are thinking in terms of intergroup comparisons with women (which is consistent with the Shih et al., 1999 manipulation). Latino men should show the tendency to choke under pressure when they think in terms of

themselves relative to other men (which is consistent with the Cheryan & Bodenhausen, 2000 manipulation). The Shih method should also induce stereotype threat (i.e., lowered performance) in Latino men when they are led to think in terms of their ethnicity rather than their gender. Thus, Latino men exposed to the Shih method should do significantly worse when ethnicity rather than gender is salient. On the other hand, because the Cheryan method fosters intragroup comparisons, stereotype threat effects should not be evident in the ethnicity salience condition. In fact, making intragroup comparisons with other Latino in-group members could even benefit performance. It is possible that Latino men exposed to the Cheryan method could perform best on a mathematics test when ethnicity rather than gender is salient.

To further test my proposition that the predicted results are due to social comparison, I will include a measure of relative performance expectations. After the manipulations, but before the mathematics test, participants will be asked to estimate how well they will do on a mathematics test relative to other student participants. If differences in the nature of social comparisons (downward versus upward) explain the disparate findings between Shih et al. (1999) and Cheryan & Bodenhausen (2000), then relative performance expectations should follow the same predicted pattern as actual performance. In addition, I should find that relative performance expectations will mediate the relationship between the identity salience manipulations and actual performance. If my predictions

regarding the manipulations are met, then the mediation analyses will allow me to be more confident that these effects are due to social comparison processes.

Design and Hypotheses

Participants in my study will be randomly assigned to one of four conditions: Shih gender salient, Shih ethnicity salient, Cheryan gender salient, or Cheryan ethnicity salient. Thus, my study will use a 2 (Method: Shih versus Cheryan) x 2 (Identity: gender versus ethnicity) between-groups design. In line with my argument that the Shih method stimulates intergroup comparisons and the Cheryan method stimulates intragroup comparisons, I predict that Latino men in the gender salient condition will expect to score higher on a mathematics test relative to other participants when they are exposed to the Shih rather than Cheryan method. I expect that actual performance will follow the same pattern. Because the Shih method encourages between-group comparisons, I also expect relative performance expectations and actual performance will be higher in the Shih-gender compared to the Shih-ethnicity condition. In contrast, relative performance expectations and actual performance should be lower in the Cheryan-gender than the Cheryan-ethnicity conditions.

CHAPTER TWO

METHOD

Participants

Participants were 123 male students who identify as Latino. They were recruited through the SONA Research Management System. Participants would receive course credit upon completion of the study in exchange for their participation. I used G-Power (power = .80 and α = .05) to determine sample size and then inflated that number by 20 percent to compensate for careless responders.

Procedure

Participants were invited to participate in a study concerning predictors of academic performance. Participants reported to the laboratory individually, be seated in front of a computer, and be randomly assigned to one of the four conditions. After agreeing to the informed consent (see Appendix B), participants completed a three-item scale to measure their level of identification with mathematics (Wininger, Adkins, Inman, & Roberts, 2014). This scale includes items such as “I really don’t care what tests say about my math ability” and “How I do in math has little relation to who I really am” (see Appendix C). The data for participants who scored above the midpoint on this scale were excluded from analyses. Brown and Pinel (2003) found that stereotype threat effects only occur for individuals who highly identify with the domain. Similarly, Shih, Pittinsky, and

Ambady (1999) found that stereotype boost effects occur only among those highly identified with the domain.

Manipulation

Next, participants completed a questionnaire designed to manipulate gender and ethnic identities through the methods of Shih versus Cheryan (see Appendix D & E). In the Shih-gender condition, participants were asked about their preferences of roommates in terms of gender. In the Shih-ethnicity condition, participants were asked about speaking more than one language. In the Cheryan-gender condition, the questions were adopted from the Collective Self-Esteem Scale (Luhtanen & Crocker, 1992) and focused on gender (e.g., "I feel good about the gender group I belong to," "Overall, my gender group is considered good by others"). In the Cheryan-ethnicity condition, the questions were also adopted from the same scale but focused on ethnicity (e.g., "I feel good about the ethnic group I belong to," "Overall, my ethnic group is considered good by others").

Measures

After the manipulation, participants completed a three-item scale about their performance expectations on the mathematics test (see Appendix F). A sample item from this scale is "Compared to other University students in this study, how well do you expect to perform on the mathematics test?" Participants were then given 20 minutes to complete the test consisting of 20 multiple-choice items taken from example Graduate Record Examinations (see Appendix G).

Scratch papers and a calculator were provided for this task. Performance of the mathematics test was analyzed by computing the ratio of the number of problems solved correctly to the number of problems attempted (Steele & Aronson, 1995). After completing the mathematics test, participants were asked to answer a final questionnaire (see Appendix H) assessing their attitudes towards the test and to provide their demographic information (see Appendix I). After the completion of the final questionnaire, participants were probed for suspicion and debriefed (see Appendix J). All surveys and the mathematics test were administered via Qualtrics.

CHAPTER THREE

RESULTS

Data from 123 male participants who identified as Latino students from the University were used to conduct a 2 x 2 between-groups analysis of variance (ANOVA) in IBM SPSS software version 24 to test if there was a significant mean difference in mathematics performance. The first independent variable was method with two levels: Shih method and Cheryan method. The second independent variable was identity with two levels: gender and ethnicity. The dependent variable were mathematic scores, attempted answers, expected number of correct answers, and relative performance expectation. Please refer to Table 1 for means and standard deviations across conditions.

Mathematics Test Score

The assumption of normality was evaluated and determined to be satisfied using the z-score criteria of ±3.3, $p < .001$. No outlier, skewness, or kurtosis were identified. The assumption of homogeneity of variance was examined and satisfied using Levene's test of equality of error variances, F (3, 119) = 1.212, p = .309. The samples were randomly selected and independent of one another; therefore, satisfying the assumption of independence. There were no missing data in the sample.

There were no significant mean differences in mathematic scores as a result of method (Shih vs. Cheryan), F(3, 119) = .108, p = .743, η_p^2 = .001. There were no significant mean differences in mathematic scores as a result of identity

(gender vs. ethnicity), $F(3, 119) = .013$, $p = .910$, $\eta_p^2 = .001$. There was no significant interaction between method and identity, $F(3, 119) = .337$, $p = .563$, $\eta_p^2 = .003$.

Attempted Answers

The assumption of normality was evaluated and determined to be satisfied using the z-score criteria of ±3.3, $p < .001$. No outlier, skewness, or kurtosis were identified. The assumption of homogeneity of variance was examined and satisfied using Levene's test of equality of error variances, $F(3, 119) = .055$, $p = .983$. The samples were randomly selected and independent of one another; therefore, satisfying the assumption of independence. There were no missing data in the sample.

There were no significant mean differences in attempt as a result of method (Shih vs. Cheryan), $F(3, 119) = .437$, $p = .510$, $\eta_p^2 = .004$. There were no significant mean differences in attempt as a result of identity (gender vs. ethnicity), $F(3, 119) = .527$, $p = .469$, $\eta_p^2 = .004$. There was no significant interaction between method and identity, $F(3, 119) = .005$, $p = .944$, $\eta_p^2 = .001$.

Performance Expectation 1

The assumption of normality was evaluated and determined to be satisfied using the z-score criteria of ±3.3, $p < .001$. No outlier, skewness, or kurtosis were identified. The assumption of homogeneity of variance was examined and satisfied using Levene's test of equality of error variances, $F(3, 119) = .862$, $p =$

.463. The samples were randomly selected and independent of one another; therefore, satisfying the assumption of independence. There were no missing data in the sample.

There were no significant mean differences in the expected number of correct answers (“How many items do you expect to get correct?”) as a result of method (Shih vs. Cheryan), $F(3, 119) = .536$, $p = .466$, $\eta_p^2 = .004$. There were no significant mean differences in performance expectation as a result of identity (gender vs. ethnicity), $F(3, 119) = .001$, $p = .971$, $\eta_p^2 = .001$. There was no significant interaction between method and identity, $F(3, 119) = .1.899$, $p = .171$, $\eta_p^2 = .016$.

Performance Expectation 2

The assumption of normality was evaluated and determined to be satisfied using the z-score criteria of ±3.3, $p < .001$. No outlier, skewness, or kurtosis were identified. The assumption of homogeneity of variance was examined and satisfied using Levene’s test of equality of error variances, $F(3, 119) = 1.264$, $p = .290$. The samples were randomly selected and independent of one another; therefore, satisfying the assumption of independence. There were no missing data in the sample.

There were no significant mean differences in the relative performance expectation (“Compared to other CSUSB students in this study, how well do you expect to perform on the mathematics test?”) as a result of method (Shih vs.

Cheryan), $F(3, 119) = 1.037$, $p = .311$, $\eta_p^2 = .009$. There were no significant mean differences in performance expectation as a result of identity (gender vs. ethnicity), $F(3, 119) = .099$, $p = .753$, $\eta_p^2 = .001$. There was no significant interaction between method and identity, $F(3, 119) = .002$, $p = .964$, $\eta_p^2 = .001$.

CHAPTER FOUR

DISCUSSION

In the current study, I aimed to examine how cues that make people aware of their gender or ethnic identities might influence their performance on a mathematics test depending on whether their identity is positively or negatively stereotyped in mathematics, and whether they are induced to make intergroup or intragroup comparisons. In line with my argument that the Shih method stimulates intergroup comparisons and the Cheryan method stimulates intragroup comparisons, I predicted that Latino men in the gender salient condition would expect to score higher on a mathematics test relative to other participants when they are exposed to the Shih rather than Cheryan method. I expected that actual performance will follow the same pattern. Because the Shih method encourages between-group comparisons, I expected relative performance expectations and actual performance would be higher in the Shih-gender compared to the Shih-ethnicity condition. Furthermore, relative performance expectations and actual performance would be lower in the Cheryan-gender than the Cheryan-ethnicity conditions. In contrast to my predictions, there were no significant effects found in mathematic scores, attempted answers, expected number of correct answers, and relative performance expectation as a result of both method (Shih vs. Cheryan) and identity (gender vs. ethnicity). Moreover, the highest effect size (partial eta-

squared) across all main and interactive effects was .016 and all others were less than .01. In sum, my hypotheses were not supported across all measures.

One possible explanation for the null results in my study is the stereotype that men are better at mathematics than women might be less pertinent than it was several years ago, when Cheryan and Bodenhausen (2000) and Shih et al. (1999) conducted their studies. At that time, the prevalence of gendered math stereotypes was long-established in the literature. In 1975, Ernest observed that students, ranging from high school freshmen level to seniors, viewed males as being more proficient in mathematics than females. In 1981, Weinreich-Haste reported that both school children and university students who assessed the masculinity versus femininity of various academic subjects rated mathematics as highly masculine and strongly associated with stereotypically masculine traits. In addition to evidence of an explicit association between gender and math, Nosek, Banaji, and Greenwald (2002) found that college students, especially women, had a strong implicit math=male association. More recent studies, however, indicate that both explicit and implicit gender stereotypes about math ability are starting to shift as a function of cultural changes. For example, Baker and Jones (1993) found evidence that the gender differences in mathematics ability is decreasing over time as women gain more access to jobs and higher education that foster the development of mathematic skills. These enhancements in women's demonstrated ability have contributed to reductions in the stereotype of men being superior at mathematics. Similarly, in 1996, Forgasz and Leder

showed that high school freshmen believed that both males and females were equal in terms of mathematics ability. In addition, recent large-scale replication studies have failed to produce gender stereotype threat effects (e.g., Agnoli, Melchiorre, Zandonella Callegher, & Altoè, 2021; Flore, Mulder, & Wicherts, 2018), which further suggests that both the explicit and implicit associations between male and math are fading. Without awareness of a positive stereotype about men's superior mathematics ability, then activation of gender identity is irrelevant to gender performance, which could explain the null effects in my study concerning gender.

The potential absence of gender math stereotypes could explain the null gender effects, but does not offer insight into why my study failed to produce stereotype threat effects in the Shih-ethnicity condition. If the gender identity condition was indeed stereotype-neutral, then performance expectations and math scores should have been lower when ethnicity was salient and the Shih manipulation was used. In addition, expectations and scores should have been lower when ethnicity was salient in the Shih rather than Cheryan conditions because the former evoked intergroup comparisons (with White students) whereas the latter evoked in-group comparisons (with other Latino students). Stereotype threat effects are essentially contingent on negatively-stereotyped group members making intergroup comparisons with outgroups that do not share the negative stereotypes (Walton & Spencer, 2009; Walton & Cohen, 2003). I can only speculate on my failure to produce the expected stereotype threat

effects. One contributing factor could be the nature of the institution at which I conducted my study. CSUSB is a primarily Hispanic-serving institution with over 66 percent of the student population identifying as Hispanic. Perhaps, Hispanic male students are less threatened by or less prone to making intergroup comparisons with White students who constitute only 12% of the student body and therefore are not highly visible. Moreover, CSUSB has fostered substantial (and highly publicized) academic achievement among Hispanic students, which is widely celebrated. An institutional culture that promotes and embraces the academic success of Hispanic students could plausibly buffer these students from threats that typically are associated with negative stereotypes in the greater culture.

Low math identification might explain both the non-significant gender and ethnic results in my study, regardless of the current state of gender math stereotypes. Research shows that stereotype threat effects (Brown & Pinel, 2003) and stereotype boost effects (Aronson et al., 1999; Saad, Meyer, Dhindsa, Zane, 2015; Shih, Pittinsky, & Ambady, 1999) only occur for individuals who highly identify with the testing domain. If mathematics was not important to my participants' academic identity, then their awareness of math stereotypes would be insignificant. That is, even if they had knowledge of positive stereotypes regarding men's mathematics ability and negative stereotypes about Latino's mathematics ability, either of the stereotypes would be personally irrelevant. Thus, participants would not have shown stereotype boost when led to think in

terms of their gender relative to women or choked under pressure when they were led to think in terms of themselves relative to other men. They also would not have exhibited stereotype threat when they engaged in ethnic-based intergroup comparisons. The performance expectations and test score of the men in my study provide evidence that the majority might not have been highly-math identified (see Table 2). Overall, they expected to answer less than 70 percent (*M* = 13.58) of the 20 questions correctly and actually only answered 31 percent (*M* = 6.26, ranging from 1 to 13) of the questions right. The highest number of correct answers was 13 (65%), indicating that even the top performers did not do particularly well. To the extent that expected and actual performance are indicators of math identification (Saad et al., 2015), these findings suggest that my participants might not have met the criteria to exhibit stereotype threat and boost effects (Aronson et al., 1999; Saad et al., 2015; Shih, Pittinsky, & Ambady, 1999).

A final possibility for failure to replicate the choking under pressure effect is that the disparate findings between the research by Shih and Colleagues (1999) and Cheryan and Bodenhausen (2000) might be due to differences in the overtness of their identity manipulations rather than the comparisons the manipulations evoked, as I proposed. Shih and colleagues (2012) argue that Cheryan and Bodenhausen's (2000) manipulation of the positively-stereotyped identity was blatant, resulting in negative performance outcomes; whereas, Shih et al.'s manipulation was subtle, resulting in positive performance outcomes.

Because this proposition was only speculation, future research should examine the type of manipulation in order to confirm whether the manipulation being blatant or subtle matters.

A major limitation was the low statistical power in this study. Because of the unprecedented nature of the COVID-19 pandemic, this situation limited the amount of participants who were recruited. Originally, 280 participants were to be recruited. However, only 123 participants' responses were recorded because of the unprecedented nature of the COVID-19 pandemic. According to Cashen and Geiger (2004), whenever researchers find non-significant effects, they must ensure that there is sufficient statistical power to confidently state whether the hypothesis was supported or not supported. Because low power was a problem in my study, a follow-up study should be conducted with at least the original proposed number of participants before any meaningful conclusions are reached.

Researchers mostly concur that negative stereotypes are detrimental to human performance (Spencer et al., 2016). In contrast, the effects of positive stereotypes are mixed, sometimes they can boost or lift performance (e.g., Shih et al., 1999; Walton & Cohen, 2003) and other times they can impair it (e.g., Cheryan & Bodenhausen, 2000). Although my results were not as predicted, the prospective role of social comparisons could still be an important avenue to pursue in order to understand when positive stereotypes are beneficial versus harmful. The resulting knowledge would not only be important to the theoretical understanding of positive stereotypes but would also be relevant to many real-

world environments, including academic contexts. It is a human tendency to compare to others who share similar characteristics as a means of self-evaluation and self-understanding (Festinger, 1954). Stereotypes can provide the basis for comparisons, but the effects of stereotypes on self-perception will vary depending upon whether they evoke intragroup or intergroup comparisons. Because the source of the comparison can affect self-perceptions, more research is needed to understand the consequences for both negative and positive stereotypes and how these consequences are contextually-dependent.

APPENDIX A:

TABLES

Table 1
Means and Standard Deviations across Conditions

	Cheryan Ethnicity		Cheryan Gender		Shih Ethnicity		Shih Gender	
DV	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Score	6.04_a	1.99	6.26_a	2.59	6.41_a	2.88	6.14_a	2.68
Attempt	15.78_a	4.11	16.28_a	4.06	15.38_a	4.56	15.83_a	4.21
PE 1	12.63_a	3.64	13.41_a	2.98	13.45_a	3.85	13.03_a	2.99
PE 2	4.41_a	1.08	4.33_a	1.06	4.55_a	1.38	4.55_a	0.99

Note. PE 1 = "How many items do you expect to get correct?" PE 2 = "Compared to other CSUSB students in this study, how well do you expect to perform on the mathematics test?" Cells in the same row that do not share subscripts reliably differ from each other at $p < .05$

Table 2

Descriptive Statistics

	N	Min	Max	*M*	*SD*
Score	123	1	13	6.26	2.54
Attempt	123	4	20	15.85	4.19
PE 1	123	2	20	13.58	3.33
PE 2	123	1	7	4.45	1.12

Note. PE 1 = "How many items do you expect to get correct?" PE 2 = "Compared to other University students in this study, how well do you expect to perform on the mathematics test?"

APPENDIX B:

INFORMED CONSENT

Introduction/Purpose: The purpose of this research is to examine the predictors of math ability.

Procedures: By choosing to participate in this study, you will be asked to complete an initial questionnaire about yourself, a mathematics test, and a posttest questionnaire assessing your attitudes of the test. This study has been approved by the Institutional Review Board Sub-Committee of the California State University, San Bernardino.

Duration and Compensation: The study should not take more than 45 minutes to complete. For those taking the study through SONA, you will receive 3.5 units of extra credit toward a psychology course of your choice as compensation for your participation. For those taking the study for payment, you will receive a $10 gift certificate from Amazon.

Confidentiality: The information that you give us will remain confidential. Your name will not be associated with your data in any way and will not appear on any data reports. Any reports of the data will present the information in aggregate form so no individual participant will be identifiable. The research might be presented at professional conferences or submitted to scientific journals for publication. The data will be stored indefinitely on a password secured survey management system and will potentially be made available to other researchers via Open Science Framework. Again, your name will not be contained in this data and your responses will not be identifiable. Your signed consent form will be stored in a locked cabinet and destroyed 7 years after the research is published.

Risks and Benefits: Participation in this study does not pose any foreseeable risks beyond those of everyday life. You might feel some discomfort if the math problems are challenging, but this discomfort should be no more than you normally experience when you take a test in your daily life as a student. There are no personal benefits for your participation in this study; however, you will be contributing to knowledge about factors that influence math ability.

Participant's Rights: You have the right to refuse to participate in this study, refuse to answer any questions, or to terminate your participation at any time without losing any rights to which you are otherwise entitled.

If you have any complaints or comments regarding this study, you can contact Dr. Donna Garcia at dmgarcia@csusb.edu. You can also contact Dr. Garcia for a copy of the study results after December 2019. If you feel you have been treated unfairly or differently as defined in this consent form, you may contact CSUSB's IRB mgillespie@csusb.edu.

I understand that any information about me obtained from this research will be held strictly confidential. I acknowledge that I am of at least 18 years old. I understand and agree with the terms described above.

Agree: yes _____ no _____

APPENDIX C:

MATH IDENTIFICATION

Math Identification Items [Participants' mean score on the below measure must be below 4 for them to qualify]

(1=strongly disagree; 4= neither agree nor disagree; 7=strongly agree)

____I really don't care what tests say about my math ability.

____No math test will ever change my opinion on how intelligent I am.

____How I do in math has little relation to who I really am.

Math Scores

1. If you completed the ACT, what was your mathematics score _____? NA____
2. If you completed the SAT, what was your mathematics score _____? NA____
3. What was your average overall grade in mathematics during high school _____ and what is your average mathematics grade so far in college _____?

(Wininger, Adkins, Inman, & Roberts, 2014)

APPENDIX D:

CHERYAN'S METHOD OF GENDER AND ETHNICITY SALIENCE

Cheryan's Method of Gender and Ethnicity Salience Condition [participants in the "Cheryan" condition will receive one or the other version of the CSE below]

PART 1A:

Strongly disagree	1	2	3	4	5	6	7 Strongly agree

1. I am a worthy member of the gender group to which I belong.
2. I am a cooperative participant in the gender group to which I belong.
3. In general, I'm glad to be a member of the gender group I belong to.
4. I feel good about the gender group I belong to.
5. Overall, my gender group is considered good by others.
6. In general, others respect the gender group that I am a member of.
7. The gender group I belong to is an important reflection of who I am.
8. In general, belonging to my gender group is an important part of my self-image.

Strongly disagree	1	2	3	4	5	6	7 Strongly agree

1. I am a worthy member of the ethnic group to which I belong.
2. I am a cooperative participant in the ethnic group to which I belong.
3. In general, I'm glad to be a member of the ethnic group I belong to.
4. I feel good about the ethnic group I belong to.
5. Overall, my ethnic group is considered good by others.
6. In general, others respect the ethnic group that I am a member of.
7. The ethnic group I belong to are an important reflection of who I am.
8. In general, belonging to my ethnic group is an important part of my self-image.

(Luhtanen &Crocker, 1992)

APPENDIX E:

SHIH'S METHOD OF GENDER AND ETHNICITY SALIENCE

Shih's Method of Gender and Ethnicity Salience Condition [participants in the "Shih" condition will receive either the gender or ethnicity version below]

PART 1B:

Gender

a) Do you have a female roommate?
b) Do you have a male roommate?
c) Do you prefer male or female roommates?
d) List 3 reasons why you prefer having both male and female roommates.
e) List 3 reasons why you prefer having only male roommates.
f) Who is cleaner male or female roommates?

Ethnicity

a) Do your parents/grandparents speak any other languages other than English?
b) What languages do you know?
c) What languages do you speak at home?
d) What opportunities allow you to speak other languages on campus?
e) What percentage of these opportunities are found in your residence halls?
f) How many generations of your family had lived in America?

APPENDIX F:

PERFORMANCE EXPECTATIONS

Performance Expectations

1. There are 20 items on the upcoming mathematics test. In the space provided, please indicate how many items that you think the average student in this study will get correct. _____

2. How many items do you expect to get correct? _____

3. Compared to other University students in this study, how well do you expect to perform on the mathematics test?

1	2	3	4	5	6	7
Extremely Poor						Extremely Well

Survey created by Ashlee Pardo and Donna Garcia.

APPENDIX G:

MATHEMATICS TEST

PART 2:

For each question, select and indicate one of the answer choices given.

1. A car covered 130 miles using 4 gallons of diesel. What distance would the same car cover, under similar conditions, on 6.7 gallons?
A) 77
B) 260.25
C) 520
D) 217.75
E) 871

2. If plotted in the same rectangular system of axis, the graphs of $f(x) = | |x| - 4 |$ and $g(x) = 2$ will have
A) no points of intersection
B) 1 point of intersection
C) 2 points of intersection
D) 3 points of intersection
E) 4 points of intersection

3. Which of the following is the largest?
A) 125%
B) 1.25
C) 1 + 1/3
D) 4/3
E) 0.0015/0.001

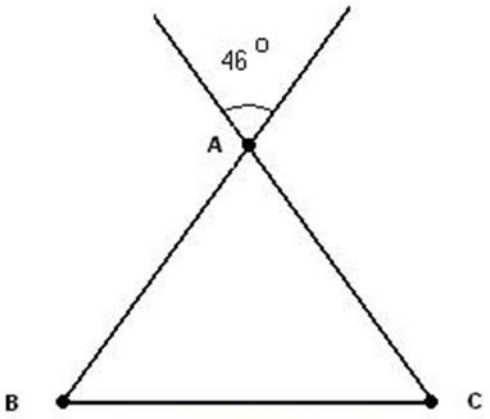
4. $| |- 10 - 19| - 20 | =$
A) 9
B) 49
C) - 49
D) 19
E) 11

5. The algebraic expression $x / (x + 2)$ is undefined if $x =$
A) 0
B) 1
C) 2
D) - 2
E) - 1

6. $(\sqrt{5} - \sqrt{7})(\sqrt{5} + \sqrt{7}) =$
A) 12
B) 2

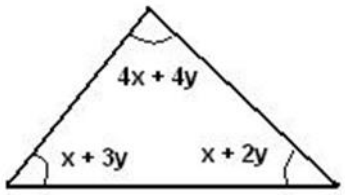
C) 24
D) - 24
E) - 2

7. In the figure below, the sides AB and AC of triangle ABC have equal lengths. Find the size of angle ABC.



A) 46°
B) 67°
C) 134°
D) 136°
E) Cannot be calculated using the given information

8. If y= 10° in the figure below, what is the value of x?



A) 20°
B) 30°
C) 10°
D) 5°
E) 15°

9. By what percent will the volume of a rectangular solid increase if its length, width, and height increase by 25% each?
A) 25%
B) 95%
C) 74%
D) 15.6%

E) 50%

10. What is the average of all prime numbers between 20 and 40?
A) 20
B) 25
C) 30
D) 35
E) 40

11. If w is the average of the numbers a, b, c and d, then the average of m(a + k), m(b + k), m(c + k), and m(d + k) is given by
A) m (w * k)
B) m (w + k)
C) m w + k
D) w + m k
E) w + m + k

12. What is the ratio of the area of the larger circle to the area of the smaller circle such that the radius of the larger circle is three times the radius of the smaller circle?
A) 3
B) 6
C) 12
D) 15
E) 9

13. A group of 20 employees in a company have an average (arithmetic mean) salary of $35,000 while a second group of 30 employees in the same company have an average salary of $40,000. What is the average salary of the 50 employees making the two groups?
A) $37,500
B) $38,000
C) $40,500
D) $39,000
E) $38,500

14. Which of the following is equal to $\sqrt{48}$
A) 16
B) $3\sqrt{4}$
C) $4\sqrt{3}$
D) $18\sqrt{3}$
E) 24

15. The sizes of the interior angles A, B, and C of a triangle are in the ratio 2:4:3. What is the measure, in degrees, of the smallest angle?

A) 40
B) 45
C) 32
D) 70
E) 30

16. If n is even and m is odd, then which of the following is true?
A) n + m is even
B) n - m is even
C) n * m is odd
D) n2 + m2 + 1 is even
E) 2n + 3m + 1 is odd

17. $5^{100} + 25^{50} + 3 (125^{34} / 25) =$
A) $33 * 5^{50}$
B) 405^{34}
C) 33^{100}
D) 5^{100}
E) 5^{101}

18. $[6x^{10} - 2x^9] / (9x^2 - 1) =$
A) $6x^9 - 2x^8 / (9x - 1)$
B) $2x^9 / (3x + 1)$
C) $(2/3)\ x^8 - 2x^9$
D) $(2/3)\ x^8 + 2x^9$
E) $2x^9 / (3x - 1)$

19. $(- 2x + 6)^2 =$
A) $4\ x^2 + 36$
B) $- 4\ x^2 + 36$
C) $4\ x^2 + 24\ x + 36$
D) $4\ x^2 - 24\ x + 36$
E) $4\ x^2 - 24\ x – 36$

20. The sum of all interior angle of a regular polygon is 1800°. How many sides does this polygon have?
A) 10
B) 11
C) 12
D) 13
E) 14

Retrieved from the example Graduate Record Examination (GRE).

APPENDIX H:

FEEDBACK SURVEY

Feedback Survey

1. Generally speaking, how well do you normally do on mathematics tests like the one you just completed?

1	2	3	4	5	6	7
Very Very Poorly						*Well*

2. Generally speaking, how well do you think that men do on mathematics tests like the one you just completed?

1	2	3	4	5	6	7
Very Very Poorly						*Well*

3. Generally speaking, how well do you think that women do on mathematics tests like the one you just completed?

1	2	3	4	5	6	7
Very Very Poorly						*Well*

4. I thought that the mathematics test was a good measure of intelligence.

1	2	3	4	5	6	7
Strongly Strongly Disagree						Agree

5. I thought that the mathematics test was a good measure of arithmetic ability.

1	2	3	4	5	6	7
Strongly Strongly Disagree						Agree

Survey created by Ashlee Pardo and Donna Garcia.

APPENDIX I:

DEMOGRAPHICS

Demographics

Please provide us with some information about YOURSELF in the space provided below.

Your Age: _____

Your Gender: _____

___Male ___Female ___Other

Your political affiliation:

1 = Liberal to 7 = Conservative

Your Race or Ethnicity:

____ Hispanic, Latina/o, or Chicano
____ European or White American
____ African or Black American
____ Asian American
____ Multi-Ethnic/Racial
____ Other: please indicate _____

Survey created by Ashlee Pardo and Donna Garcia.

APPENDIX J:

DEBRIEFING STATEMENT

Debriefing Statement

Now that the study is complete, we would like to provide you with some information about the purposes of this research.

We are interested in how cues that make people aware of their gender or ethnic identities might influence their performance on math test in ways that are consistent with common stereotypes. For example, there are both positive and negative stereotypes about Latino men's math ability because the stereotype for men's ability is positive but the stereotype for Latino is negative. Thus, Latino men might do better on a math test when they have been led to think in terms of their gender rather than their ethnic identity.

Please accept our apology for not telling you all of the details about the purposes of our study right from the beginning. If participants know the purpose of the study, they would behave very differently than they naturally would. For this reason, we ask that you not tell other students who might be participating in our research this quarter what the specific hypotheses are and the true purpose of this study.

If you would like additional information concerning this study or a copy of the results, please feel free to contact Donna Garcia (dmgarcia@csusb.edu). Thank you very much for participating in our study.

APPENDIX K:

IRB APPROVAL

Date: 1-27-2019

IRB #: IRB-FY2019-86
Title: Social Comparison and Shifting in Quantitative Performance
Creation Date: 11-28-2018
End Date: 1-14-2020
Status: Approved
Principal Investigator: Ashlee Pardo
Review Board: Main IRB Designated Reviewers for Department of Psychology
Sponsor:

Study History

Submission Type Initial	Review Type Exempt	Decision Exempt
Submission Type Modification	Review Type Exempt	Decision Approved

Key Study Contacts

Member Ashlee Pardo	Role Principal Investigator	Contact
Member Donna Garcia	Role Co-Principal Investigator	Contact dmgarcia@csusb.edu
Member Donna Garcia	Role Co-Principal Investigator	Contact dmgarcia@csusb.edu
Member Ashlee Pardo	Role Primary Contact	Contact
Member Ashlee Pardo	Role Investigator	Contact

REFERENCES

Agnoli, F., Melchiorre, F., Zandonella Callegher, C., & Altoè, G. (2021). Stereotype threat effects on Italian girls' mathematics performance: A failure to replicate. *Developmental Psychology*, *57*(6), 940.

Aronson J, Lustina MJ, Good C, Keough K, Steele CM, Brown J. 1999. When white men can't do math: necessary and sufficient factors in stereotype threat. *J. Exp. Soc. Psychol. 35*(1):29–46.

Baker, D. P., & Jones, D. P. (1993). Creating gender equality: Cross-national gender stratification and mathematical performance. *Sociology of Education*, 91-103.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ, USA: Prentice-Hall.

Belmi, P., Barragan, R. C., Neale, M. A., & Cohen, G. L. (2015). Threats to social identity can trigger social deviance. *Personality and Social Psychology Bulletin*, *41*(4), 467-484.

Blanton, H., Buunk, B. P., Gibbons, F. X., & Kuyper, H. (1999). When better-than-others compare upward: Choice of comparison and comparative evaluation as independent predictors of academic performance. *Journal of Personality and Social Psychology*, *76*(3), 420.

Brewer, M. B., & Weber, J. G. (1994). Self-evaluation effects of interpersonal versus intergroup social comparison. *Journal of Personality and Social Psychology*, *66*(2), 268.

Brown, R. P., & Pinel, E. C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology*, *39*(6), 626-633.

Cashen, L. H., & Geiger, S. W. (2004). Statistical power and the testing of null hypotheses: A review of contemporary management research and recommendations for future studies. *Organizational Research Methods*, *7*(2), 151-167.

Cheryan, S., & Bodenhausen, G. V. (2000). When positive stereotypes threaten intellectual performance: The psychological hazards of 'model minority' status. *Psychological Science*, *11*(5), 399-402.

Croizet, J. C., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, *24*(6), 588-594.

Czopp, A. M., Kay, A. C., & Cheryan, S. (2015). Positive stereotypes are pervasive and powerful. *Perspectives on Psychological Science*, *10*(4), 451-463.

Ernest, J. (1975). Mathematics and sex. Santa Barbara: California University, Department of Mathematics. (ERIC Document Reproduction Service No. ED107535)

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117-140.

Flore, P. C., Mulder, J., & Wicherts, J. M. (2018). The influence of gender stereotype threat on mathematics test scores of Dutch high school students: A registered report. *Comprehensive Results in Social Psychology*, *3*(2), 140-174.

Forgasz, H. J., & Leder, G. C. (1996). Mathematics and English: Stereotyped domains? *Focus on Learning Problems in Mathematics*, *18*(1, 2 & 3), 129–137.

Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin*, *28*(5), 659-670.

Guimond, S. (Ed.). (2006). *Social comparison and social psychology: Understanding cognition, intergroup relations, and culture*. Cambridge University Press.

Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research*, *50*(1), 1-22.

Hayes, A. F. (2019). The PROCESS macro for SPSS and SAS. Retrieved May 20, 2019, from http://processmacro.org/version-history.html

Inzlicht, M., McKay, L., & Aronson, J. (2006). Stigma as ego depletion: How being the target of prejudice affects self-control. *Psychological Science*, *17*(3), 262-269.

Inzlicht, M., & Schmader, T. (2012). *Stereotype threat: Theory, process, and application*. New York, NY: Oxford University Press, Inc.

Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, *100*(1), 109.

Konan, P. N. D., Chatard, A., Selimbegović, L., Mugny, G., & Moraru, A. (2011). Deflecting stereotype threat through downward comparison: When comparison with immigrants boosts the performance of stigmatized native students. *Social Justice Research*, *24*(2), 191-205.

Kray, L. J., Thompson, L., & Galinsky, A. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology*, *80*, 942–958.

Lambert, A. J., Khan, S. R., Lickel, B. A., & Fricke, K. (1997). Mood and the correction of positive versus negative stereotypes. *Journal of Personality and Social Psychology*, *72*(5), 1002.

Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, *18*(3), 302-318.

Morse, S., & Gergen, K. J. (1970). Social comparison, self-consistency, and the concept of self. *Journal of Personality and Social Psychology*, *16*(1), 148-156.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*(1), 44.

Saad, C. S., Meyer, O. L., Dhindsa, M., & Zane, N. (2015). Domain identification moderates the effect of positive stereotypes on Chinese American women's math performance. *Cultural Diversity and Ethnic Minority Psychology*, *21*(1), 162.

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*(2), 336.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*(1), 80-83.

Shih, M. J., Pittinsky, T. L., & Ho, G. C. (2012). Stereotype boost: Positive outcomes from the activation of positive stereotypes.

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, 67, 415-437.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 4-28.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(6), 613.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In *Advances in Experimental Social Psychology* (Vol. 34, pp. 379-440). Academic Press.

Tagler, M. J. (2012). Choking under the pressure of a positive stereotype: Gender identification and self-consciousness moderate men's math test performance. *The Journal of Social Psychology*, *152*(4), 401-416.

Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, *39*(5), 456-46.

Walton, G. M., & Spencer, S. J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science*, *20*(9), 1132-1139.

Weinreich-Haste, H. (1981). “The image of science,” in The Missing Half: Girls and Science Education, ed A. Kelly (Manchester, UK: Manchester University Press), 216–229.

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, *90*(2), 245.7.

Wininger, S. R., Adkins, O., Inman, T. F., & Roberts, J. (2014). Development of a student interest in mathematics scale for gifted and talented programming identification. *Journal of Advanced Academics*, *25*(4), 403-421.

Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. I., & O'connell, K. M. (1992). The role of categorization in visual search for orientation. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 34.

Wyer, R. S., & Srull, T. K. (1986). Human cognition in its social context. *Psychological Review*, *93*(3), 322.