

No Child Left Behind, Stereotype Threat, and the Standardized Testing of African American Third-Graders

Martin Wasserberg

Florida International University, USA

Abstract: *Stereotype threat* theory proposes that the possibility of being judged in terms of a negative stereotype in a particular domain negatively affects one's performance. The proposed mixed-methods research will investigate the influences of stereotype threat on African American third-graders in a post-No Child Left Behind environment.

Stereotype threat (Steele & Aronson, 1995) refers to the risk of confirming a negative stereotype about one's group in a particular performance domain. Stereotype threat theory assumes that underperformance is triggered by the possibility of being judged in terms of said stereotype. Given the possibility of positive intervention (Aronson, Fried, & Good, 2002; Cohen, Garcia, Apfel, & Master, 2006; Good, Aronson, & Inzlicht, 2003), a necessary next step is to examine how children experience stereotype threat effects. This is particularly important when addressing potential remedies for the racial achievement gap in standardized testing, which has become increasingly important at the elementary level (U.S. Department of Education, 2007). Therefore, the research questions that guide this proposal are: (a) How does stereotype threat influence the reading test performance of African American children in an urban elementary school? What mediates these stereotype threat effects? (b) How do African American children perceive these influences, and navigate through the school year approaching the standardized test?

African American children are targeted by a negative stereotype of intellectual inferiority in all academic areas (Steele, 1997). This negative racial stereotype is made salient in a standardized testing situation, thereby impeding performance and causing African American elementary students to perform more poorly than they would in a neutral context (McKown & Weinstein, 2003, Study 2). However, past research on stereotype threat effects on African American children failed to take into account *domain-identification*. That is, individuals most affected by stereotype threat are highly identified with the domain in question (Steele). Threatening situational pressure thus has greater effect on a subset of the stereotyped group that more strongly ties their identity to the standardized test results.

No Child Left Behind (NCLB) (2002) has mandated for schools that do not make *Adequate Yearly Progress* (based on standardized test scores) to adopt state sponsored tutoring programs (U.S. Department of Education, 2007). These schools house predominantly African American populations, and often adopt curricula centered on standardized test practice (Ahlquist, 2003; Cawelti, 2006). African American students at such schools tie their identity to their standardized test scores (Kozol, 2005). Negative stereotype threat performance effects are augmented when the performance domain is highly self-relevant for the individual (Aronson et al., 1999; Keller, 2007; Spencer, Steele, & Quinn, 1999). The main hypothesis of the proposed study is that the stereotype threat conditions evoked by diagnostic testing will negatively affect the standardized test performance of stereotype-aware African American children, but more so if they are domain-identified. Another hypothesis is that the effect of stereotype threat conditions

Wasserberg, M. (2008). No Child Left Behind, stereotype threat, and the standardized testing of African American third-graders. In M. S. Plakhotnik & S. M. Nielsen (Eds.), *Proceedings of the Seventh Annual College of Education Research Conference: Urban and International Education Section* (pp. 173-183). Miami: Florida International University.

http://coeweb.fiu.edu/research_conference/

on standardized test performance will be mediated by increased anxiety, decreased self-efficacy, and orientation towards a performance-avoidance achievement goal. The last hypothesis is that stereotype threat effects can be mitigated by presenting exams as nondiagnostic, and eliminating the threatening environment. The qualitative segment of the proposed study will investigate how African American third-graders perceive their stereotyped status, and will allow for a deeper understanding of the underlying processes.

Theoretical Framework

The conceptual model clarifies the hypotheses (see Figure 1). Awareness of the threatening stereotype is a necessary prerequisite of stereotype threat experiences (McKown & Weinstein, 2003, Study 1). Identity salience, or one's consciousness of one's stereotyped identity (Shih, Pittinsky, & Ambady, 1999), and domain-identification (Aronson et al., 1999; Keller, 2007; Spencer et al., 1999; Steele, 1997), augment negative stereotype threat performance effects. Diagnostic standardized testing situations have been enough to evoke negative stereotype threat performance effects in stigmatized populations (McKown & Weinstein, 2003; Steele & Aronson, 1995). Experiencing stereotype threat has been linked to orientation towards *performance-avoidance* achievement goals (a focus on avoiding negative judgments instead of on mastering the task) (Ryan & Ryan, 2005). Whereas stereotype threat research has failed to reliably pinpoint mediating processes (see Smith, 2004 for a review), performance-avoidance achievement goals have been directly linked to increased anxiety (Elliot & McGregor, 1999, 2001), decreased self-efficacy (Middleton & Midgley, 1997), and subsequent depressed test performance. In this manner, when stereotype-aware African American children take a standardized test presented as diagnostic of ability, it can be reasonably hypothesized that they will experience increased anxiety and decreased self-efficacy, accompanied by depressed test performance. Furthermore, students who are domain-identified will likely experience augmented negative performance effects. Additionally, relevant experiences of these children during the school year approaching the standardized test will be investigated.

Proposed Method – Phase 1

The proposed research will utilize a mixed method design, combining both quantitative and qualitative approaches. Specifically, a *sequential mixed method design* (Tashakkori & Teddlie, 1998) will be implemented; a quantitative phase of the study will be followed by a separate qualitative phase. Methods will be mixed for the purpose of *complementarity* (Greene, 2007), to gain a more comprehensive understanding by exploring different aspects of a complex phenomenon. The phenomenon being investigated is stereotype threat: the risk of confirming a negative stereotype about one's group in a particular performance domain.

The participants will be African American third grade students at an urban elementary school in a major metropolitan area in Florida. The student composition of the school is 80% African American, 19% Hispanic, and 1% White. Over 90% of the students qualify for free or reduced lunch (MDCPS, 2006). The school has never made *Adequate Yearly Progress* (U.S. Department of Education, 2007) by NCLB standards, and has implemented several test-preparation protocols mandated by the state as a result. These characteristics are typical of many schools in urban centers in the United States (Kozol, 2005). Third-graders were chosen because of the particularly high stakes of the third grade standardized reading test: low performance can result in mandatory retention based on NCLB standards (U.S. Department of Education, 2007).

Following an open-ended approach to assessing children's stereotype awareness (Biek, 2006; McKown & Weinstein, 2003), a written measure was developed which asks participants how a planet where green people did not think blue people were smart was like the real world.

The Domain Identification Measure (DIM) will be administered to stereotype-aware participants. A median split will be performed on the resulting scores to create domain-identified and non-domain-identified groups within each threat condition. Steele and Aronson's (1995) methods will then be applied, manipulating stereotype threat conditions by characterizing a practice Florida Comprehensive Assessment Test (FCAT) in Reading as either a practice standardized test diagnostic of ability, or as a nondiagnostic performance task. The practice FCAT will consist of two grade-level reading passages, each followed by 10 multiple choice questions from state-provided standardized test practice materials (for a total of 20 items). Performance will be measured by number of questions answered correctly in 30 min. The domain-identified and non-domain-identified participants will each be randomly assigned to either a diagnostic or nondiagnostic testing condition. The experimenter will prepare students for the activity using a script derived from McKown and Weinstein (2003) (see Appendix A). After reading comprehension activity directions are given, but before the participants begin working, the goal orientation, self-reported anxiety, and self-efficacy of participants will be assessed using the appropriate measures. Self-reported anxiety will be measured for all participants using the State Anxiety Scale from State Trait Anxiety Inventory for Children (STAIC) (Spielberger, Edwards, Montouri, & Lushene, 1973). Self-efficacy will be measured for all participants using the academic self-efficacy component of the Patterns of Adaptive Learning Scales (PALS) (Middleton & Midgley, 1997; Midgley et al., 1996). Goal orientation of participants will be measured using the performance-avoid goal orientation (revised) component of the PALS (Midgley et al., 2000). Additionally, although research has established that stigmatized individuals suffer impaired performance under stereotype threat conditions, the anxiety presumed to help mediate this effect has proven difficult to establish through self-reports. Therefore, following the model of Bosson, Haymovitz, and Pinel (2004), anxiety will also be assessed by a judge blind to all procedures directed to look for behaviors that communicate anxiety during the test. Ratings for each participant will be averaged to find a mean level of observed anxiety during the test.

Proposed Method – Phase 2

The participants will be a purposefully selected group of 4 of the African-American third grade students found to be domain-identified with reading in Phase 1 of the study.

The design of the study will include interviews supplemented by classroom observations. Participants will be interviewed on several occasions throughout the school year with questions related to identity salience, stereotype awareness, goal orientation, test anxiety, and domain-identification related to the FCAT. The interviews will be semistructured. Semistructured interviews allow the interviewer and the person being interviewed the flexibility to probe for details or discuss issues (Gall, Gall, & Borg, 2007). A framework for each interview will be developed beforehand. The initial framework will be adapted to create an outline of interview questions more specific to participants. The interviews will be recorded and transcribed and participants will read the transcripts and make any changes they deem necessary. Similar statements will then be categorized through a detailed line-by-line analysis. Codes will then be applied to transcripts, and coded passages will be reanalyzed to develop cogent categories related to the discussed mediators of stereotype threat. The passages will then be grouped and reorganized into themes. Conclusions on perceptions of influencing factors will be drawn from these themes. This data will be supplemented with fieldnotes from classroom observations. The classroom observations and member checking will serve as forms of triangulation.

Pilot Study

A pilot study was conducted to provide preliminary quantitative data on the effects of stereotype threat on African American third graders in an urban elementary school. The procedure was identical to that delineated in the Phase 1 methods, with two exceptions: (a) anxiety was assessed only through self-report, and (b) the final sample of stereotype-aware participants included 17 students from one classroom.

A 2 (domain-identification) x 2 (threat condition) MANOVA was conducted evaluate the effects of stereotype threat on reading test scores, anxiety, self-efficacy, and goal orientation. There was a significant main effect of threat condition on reading score $F(1,13) = 15.81, p < .01$, and a near significant main effect of threat condition on anxiety $F(1,13) = 4.32, p = .06$. As hypothesized, participants in the nondiagnostic condition scored significantly higher ($\bar{x} = 47.56\%$) than participants in the diagnostic condition ($\bar{x} = 15.75\%$). Also, participants in the nondiagnostic condition reported less anxiety ($\bar{x} = 28.65$) than participants in the diagnostic condition ($\bar{x} = 36.68$) (see Figures 2 and 3). There were no other significant main effects. This may be due to the small sample size utilized in the pilot study.

There was also a near significant interaction between domain-identification and threat condition on reading score $F(1,13) = 2.58, p = .13$. The simple main effects were further analyzed, revealing highly significant differences in reading test scores between threat conditions only for domain-identified participants $F(1,13) = 19.15, p < .01$. Domain-identified participants in the nondiagnostic condition scored significantly higher ($\bar{x} = 57.80\%$) than domain-identified participants in the diagnostic condition ($\bar{x} = 15.20\%$). Scores differed in the same direction for non-domain-identified participants; however, this difference was not statistically significant (see Figure 4). These results were in line with the primary research hypothesis. There were no other significant interactions. This is likely a factor of the small sample size utilized in the pilot study. The results of the pilot study provide preliminary data highlighting the importance of the proposed study.

Educational Implications of Potential Results

Stereotype threat theory posits that awareness of negative stereotypes induces members of stigmatized groups to become concerned that their performance will be judged in terms of these stereotypes, which depresses their performance within the stereotyped domain (McKown & Weinstein, 2003; Spencer et al., 1999; Steele & Aronson, 1995). Members of stigmatized groups who are more identified with the performance domain are particularly affected by stereotype threat manipulations (Steele, 1997). Results of the pilot study support the main hypothesis that the stereotype threat conditions evoked by diagnostic testing will negatively affect the standardized test performance of stereotype-aware African American children, but more so if they are domain-identified. If results of the larger proposed study support this hypothesis, this would indicate that racial stereotypes regarding academic performance become salient and have adverse performance effects at early ages. The subsequent qualitative investigation will lead to a deeper understanding of the situational and cognitive processes mediating these effects from student perspectives. These potential effects are disconcerting in a time when NCLB has drastically increased the significance of standardized testing in the elementary grades. NCLB supported testing regimens may have adverse cognitive and performance effects on certain populations of students. Attention to the situational presentation of testing is therefore critical. The proposed research could bolster the argument that high-stakes standardized testing situations make relevant social stereotypes salient, thereby causing depressed performance in stigmatized populations. Such information is of particular importance to educational policymakers involved in standardized achievement based legislation.

Additionally, if results support the hypotheses, this would suggest that test performance is sensitive to situational and cognitive processes amenable to teacher intervention. This is particularly important to teachers of African American children in that attention to the environmental details surrounding standardized testing situations can potentially prevent maladaptive consequences for their students. Positive intervention has already been demonstrated in middle school and college populations by teaching students to view intelligence as malleable rather than fixed (Aronson et al., 2002; Good et al., 2003), by having students reaffirm their sense of self-worth (Cohen et al., 2006; Frantz, Cuddy, Burnett, Ray, & Hart, 2004; Martens, Johns, Greenberg, & Schimel, 2006), and by increasing accessibility to positive in-group role models (Huguet & Regner, 2007; Marx & Roman, 2002). If educators are interested in ameliorating the racial achievement gap, implementation of interventions to help prevent the negative performance consequences evoked by stereotype threat is essential at earlier ages. The proposed research will highlight opportunities for possible intervention.

References

- Ahlquist, R. (2003). Challenges to academic freedom: California teacher educators mobilize to resist state-mandated control of the curriculum. *Teacher Education Quarterly*, 30(1), 57–64.
- Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, 38, 29-46.
- Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, 35, 29-46.
- Biek, D. M. (2006). *Stereotype threat and domain identification*. Unpublished Ph.D., Cornell University, United States -- New York.
- Bosson, J. K., Haymovitz, E. L., & Pintel, E. C. (2004). When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of Experimental Social Psychology*, 40, 247-255.
- Cawelti, G. (2006). The side effects of NCLB. *Educational Leadership*, 64(3), 64-685.
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, 313(5791), 1307.
- Elliot, A. J., & McGregor, H. A. (1999). Test anxiety and the hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, 76(4), 628-644.
- Elliot, A. J., & McGregor, H. A. (2001). A 2 X 2 achievement goal framework. *Journal of Personality and Social Psychology*, 80(3), 501-519.
- Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the computer: The race implicit association test as a stereotype threat experience. *Personality and Social Psychology Bulletin*, 30(12), 1611-1624.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). New York: Pearson.
- Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, 24(6), 645-662.
- Greene, J. C. (2007). *Mixed methods in social inquiry*. San Francisco: Jossey-Bass.

- Huguet, P., & Regner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology, 99*(3), 545-560.
- Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' math performance. *British Journal of Educational Psychology, 77*(2), 323-338.
- Kozol, J. (2005). *The shame of the nation: The restoration of apartheid schooling in America*. New York: Crown Publishers.
- Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology, 42*(2), 236-243.
- Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting women's math test performance. *Personality and Social Psychology Bulletin, 28*(9), 1183-1193.
- McKown, C., & Weinstein, R. S. (2003). The development and consequences of stereotype consciousness in middle childhood. *Child Development, 74*(2), 498-515.
- Miami-Dade County Public Schools. (2006). *2005-2006 District and School Profiles*. Retrieved November 27, 2007, from <http://oada.dadeschools.net/DSPfiles/0506Profiles.asp>
- Middleton, M. J., & Midgley, C. (1997). Avoiding the demonstration of lack of ability: An underexplored aspect of goal theory. *Journal of Educational Psychology, 89*(4), 710-718.
- Midgley, C., Maehr, M., Hicks, L., Roeser, R., Urdan, T., Anderman, E., et al. (1996). *Patterns of adaptive learning survey (PALS)*. Ann Arbor, MI: Center for Leadership and Learning, University of Michigan.
- Midgley, C., Maehr, M. L., Huda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., et al. (2000). *Manual for the Patterns of Adaptive Learning Scales*. Ann Arbor, MI: University of Michigan.
- Ryan, K. E., & Ryan, A. M. (2005). Psychological processes underlying stereotype threat and standardized math test performance. *Educational Psychologist, 40*(1), 53-63.
- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*(1), 80-83.
- Smith, J. L. (2004). Understanding the process of stereotype threat: A review of mediational variables and new performance goal directions. *Educational Psychology Review, 16*(3), 177-206.
- 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.
- Spielberger, C. D., Edwards, C. D., Montouri, J., & Lushene, R. (1973). *State Trait Anxiety Index for Children (STAIC) test manual*. Menlo Park, CA: Mind Garden.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613-629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage Publications.
- U. S. Department of Education. (2007). *No Child Left Behind*. Retrieved November 7, 2007, from <http://www.ed.gov/nclb/landing.jhtml>

Appendix A

Directions

Diagnostic Condition

Now we are going to complete some reading questions. Some are easy and some are hard. You probably will not get all of the questions correct. Let me tell you why we are doing these questions. The questions you are going to answer are practice for the FCAT. They are a very, very good way of finding out how well you will perform on the actual FCAT. The test is difficult so that I can really find out how well you will do on the FCAT. Please do your best so I can see what you are good at, and what you are not so good at.

Nondiagnostic Condition

You are about to complete a problem solving activity. This is not a test. The questions are difficult so that I can really see how children solve problems. Please try the best that you can.

Figure Captions

Figure 1. Conceptual model illustrating the mechanisms mediating stereotype threat as related to diagnostic test performance.

Figure 2. Effect of threat condition on the reading test performance of African American third graders.

Figure 3. Effect of threat condition on the self-reported anxiety of African American third graders.

Figure 4. Effect of domain-identification on stereotype threat reading test performance effects for African American third graders.

Figure 1

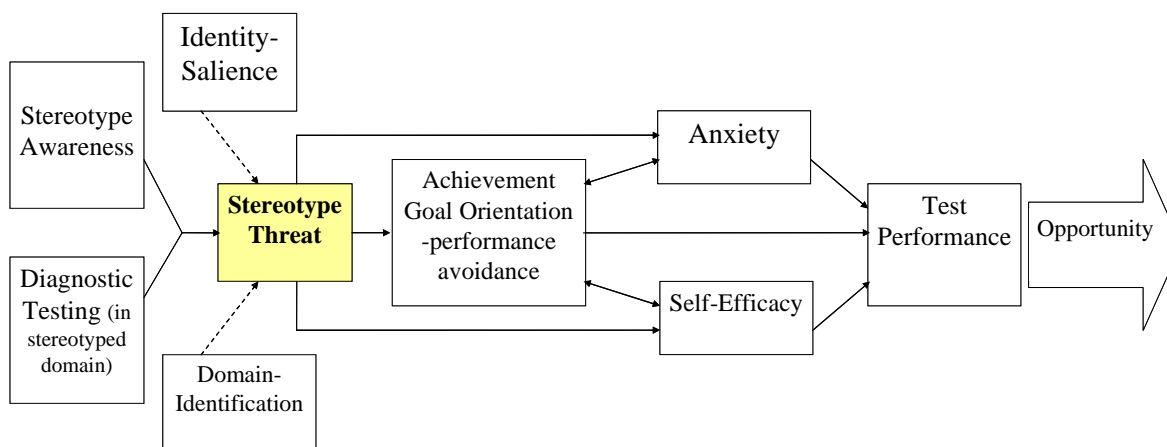


Figure 2

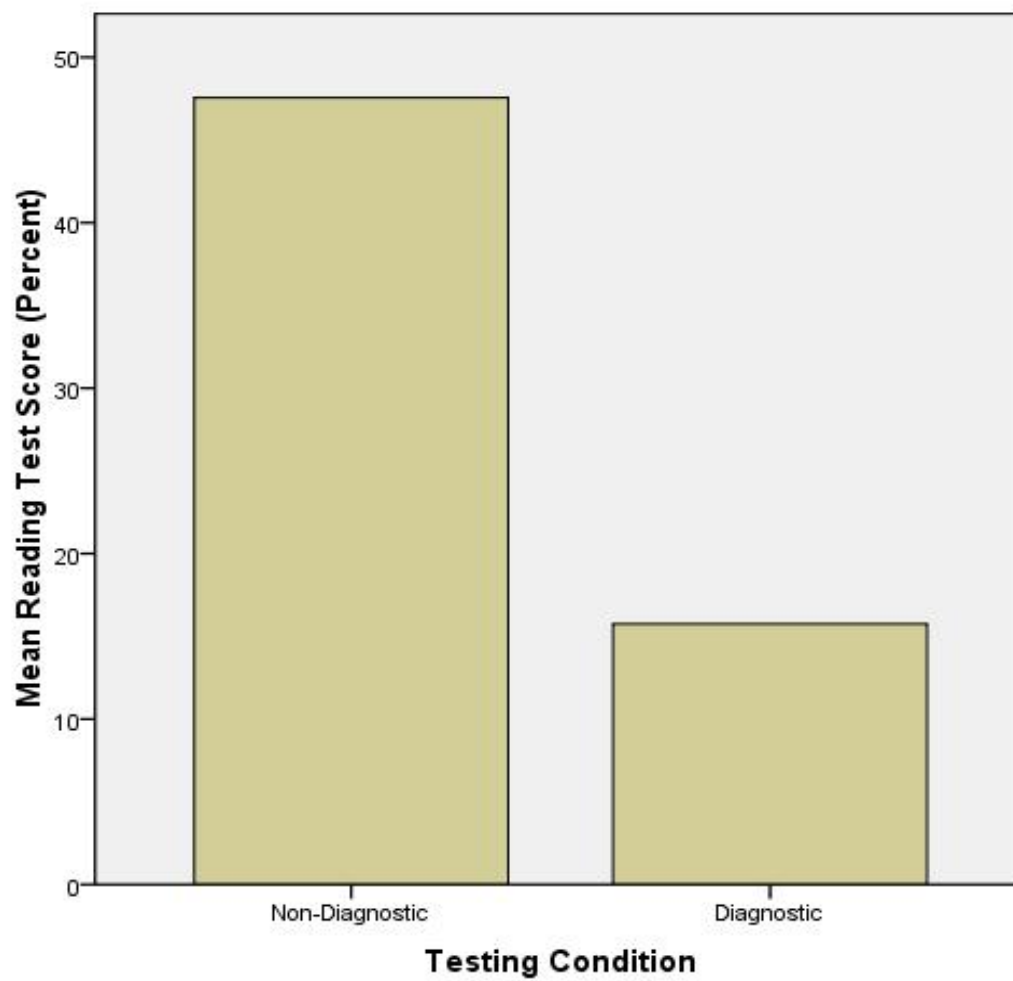


Figure 3

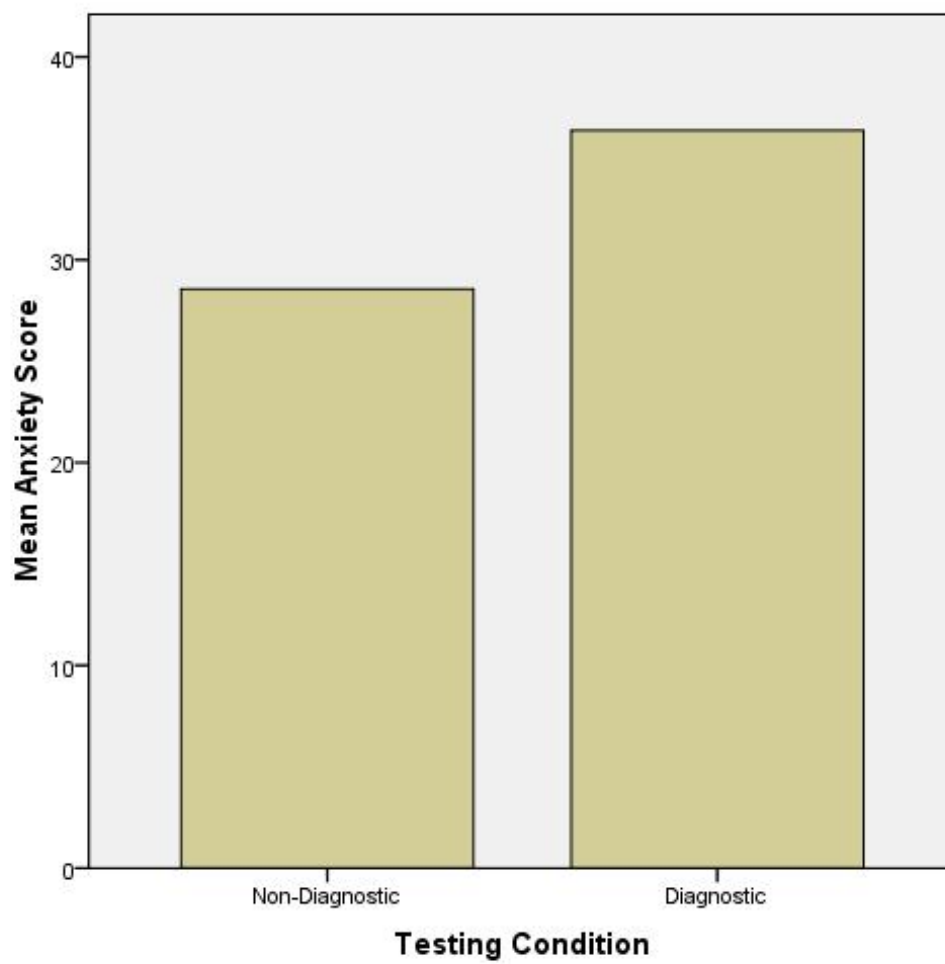


Figure 4

