Stereotype Threat and OC Symptomatology: The Impact of Messy vs. Clean Environments on Cognitive Test Performance

Ellen R. Kendall
Butler University, ekendall@butler.edu

Sarah M. McRoberts
Butler University, smcrober@butler.edu

STEREOTYPE THREAT AND OBSESSIVE COMPULSIVE SYMPTOMATOLOGY: THE IMPACT OF MESSY VS. CLEAN ENVIRONMENTS ON COGNITIVE TEST PERFORMANCE

ELLEN R. KENDALL, BUTLER UNIVERSITY
SARAH M. McROBERTS, BUTLER UNIVERSITY
MENTOR: TARA LINEWEAVER

Abstract

Stereotype threat has been researched in a variety of contexts such as African Americans' intellect, older adults' memory, and women's performance in math. Despite this extensive research, little has been done in the domain of mental illness. This study examines whether stereotype threat can be induced in people high in obsessive compulsive (OC) symptoms. We hypothesized that, when given explicit information about their OC tendencies, individuals high in OC symptoms would perform less well on cognitive tests in a messy than a clean environment compared to those low in OC symptoms. Group testing sessions included a mix of college students high ($n = 25$) and low ($n = 22$) in OC symptomatology. The classroom and testing packets were either messy or clean. At the beginning of the session, participants were given confidential, accurate information about their OC tendencies before completing tests of concentration and immediate and delayed memory. Across the four tests, the High and Low OC groups performed similarly in a non-threat inducing clean environment. However, in a threat-inducing messy environment the High OC group showed a strong tendency to perform less well than the Low OC group on a test of auditory attention. Thus, our results suggest that individuals with OCD or related symptoms may be susceptible to stereotype threat, much like other vulnerable populations.

Stereotype threat was first studied and described by Steele and Aronson (1995). It involves a situation in which an individual is "at risk of confirming, as self-characteristic, a negative stereotype about one's group." Various groups can be prone to differing stereotypes, and stereotype threat can emerge when society holds stereotypes about the group that might transfer to the individual. Because one common stereotype held by society is that African

Americans are less intelligent than whites, Steele and Aronson looked at the performance of African Americans on intelligence tests in their 1995 study. They found that African Americans who were told that the test was diagnostic of their abilities performed worse than those who were not told it was aiming to measure their abilities, consistent with the effect of stereotype threat on performance. They also found that even when the test was not presented as diagnostic of abilities and participants were asked to simply indicate their race on a questionnaire prior to taking the test (a race prime), blacks performed worse than whites.

Other researchers have examined threats to other groups. Levy (1996) found both positive and negative effects on memory performance and memory self-efficacy as a result of exposure to varying stereotypes of older adults. This study was performed in a between-subject design in which participants aged 60 or older were given subliminal information about aging stereotypes through a computer that presented older adults as being either senile (negative) or wise (positive). In order to help them self-identify with the old age category, they were presented with subtle cues throughout the process such as being recruited through advertisements that asked for participants aged 60 or older, being asked to answer questions related to their own age, and being exposed to the words old or senior at the beginning of each word block. Results showed that older adults primed with positive views of aging performed better than those primed with negative views. This finding did not transfer to younger participants (ages 18-35), showing that only when the stereotype is relevant to the person does the priming have an effect.

Another study found that women's performance on mathematical tests was increased when they were presented with an intervention that nullified the stereotype of women's performance in mathematics (Good, Aronson & Harder, 2008). Participants in this study were in one of two conditions: stereotype threat or stereotype nullification. The intervention was presented through the use of math test instructions that described the test as either examining why differences in mathematical abilities exist between people (threat condition) or as examining how well the test can measure abilities, noting that the test has shown no differences between genders (stereotype nullification). Results of the study showed that women in the stereotype nullification condition actually outperformed both women and men in the threat condition.

Aronson, Lustina, Good, Keough, Steele, and Brown (1999) also examined stereotype threat in the domain of mathematics by studying white males' math abilities. In this study, they used a group who generally fit more positive stereotypes and looked to see whether stereotype threat could be

induced. Specifically, they gave a group of mathematically high achieving white males a math test in either a threat or no-threat condition. The threat condition involved having participants skim over articles about the high mathematical abilities of Asians and being told that the study was interested in finding out why Asians perform better than others on math tests. Participants in the control condition did not receive the stereotype information. These researchers demonstrated that stereotype threat results in lower performance, even with these mathematically talented individuals.

One particularly interesting study found that the same task can produce different outcomes depending on which racial stereotypes the threat activates (Stone, Lynch, Sjomeling, & Darley, 1999). Specifically, the way in which you describe a task can affect whether participants feel threatened or not. This study found that, when a test of golf performance was described as measuring "sports intelligence," blacks performed less well than when it was described as testing "psychological factors correlated with general sports performance." However, when it was described as testing natural athletic ability, the performance of blacks improved. In contrast, when it was described as testing natural athletic ability, whites performed less well than they did when it was described as testing sports intelligence.

Together, these studies have illustrated the effects of stereotype threat in many different domains of life. This literature indicates that stereotype threat can be affected by the highlighting of stereotype information, by increasing someone's self-identification with a particular stereotyped group, by nullifying stereotypes, and by manipulating the description of a task.

STEREOTYPE THREAT IN MENTAL ILLNESS

One out of five adults (Substance Abuse and Mental Health Services Administration, 2013) and one out of every four to five adolescents ages 13 to 18 (Merikangas *et al.*, 2010) experiences mental illness. Given these numbers, stereotype threat associated with mental illness is a concept that could affect many people. Research examining stereotype threat and individuals with mental illness has been somewhat limited. One study examined the impact of stereotype threat on people with schizophrenia (Henry, von Hippel, Shapiro, 2010). The common stereotype of those with schizophrenia is that they do not function well in social situations. In this study, participants were in one of two conditions: stereotype threat or non-stereotype threat. Those in the stereotype threat condition had a conversation with a confederate after being told that the confederate knew

about their diagnosis, and those in the non-stereotype threat condition were told that the confederate was unaware of their diagnosis. The confederates rated those in the no-threat condition as displaying better social skills than those in the threat condition, even though they were unaware of the participant's condition during the interaction. However, no differences existed between the two conditions in how the participants perceived their own social behavior during the interactions. Results of this study show that stereotype threat can have an effect in the domain of mental illness and that individuals may not even be aware of how the threat is influencing their own behavior.

OBSESSIVE-COMPULSIVE DISORDER

One particular mental illness that has received much attention in recent years is Obsessive Compulsive Disorder (OCD). The Mayo Clinic describes OCD as "an anxiety disorder characterized by unreasonable thoughts and fears (obsessions) that lead you to do repetitive behaviors (compulsions)." This is a serious disorder that can make life very difficult for those who suffer from it for many reasons, such as spending countless hours on obsessions and compulsions each day. Although many people with OCD realize that their obsessions are unreasonable, they may perform compulsions to relieve their stress, only causing their stress to rise. The latest Diagnostic and Statistical Manual of Mental Disorders (5th ed.; DSM–5; American Psychiatric Association, 2013) describes certain themes that are common in the obsessions and compulsions of different people such as cleanliness, symmetry, forbidden or taboo thoughts, and harm. The DSM-5 states that 1.2% of people have the disorder each year, and people first experience the disorder at an average age of 19.5 years.

Given the prevalence and life-altering nature of OCD, it is crucial to find ways to diminish its effects for those who suffer from it. Symptoms may be very distressing and disabling for those with the disorder even when no one is aware of their diagnosis. Additionally, recent evidence suggests that mental illnesses like OCD may not be categorical (individuals either meet criteria for the disorder or they do not), but rather may reflect symptoms that lie along a continuum (see Linscott & van Os for a review of the literature relating to the psychosis continuum). Multiple recent studies have focused on subclinical obsessive compulsive (OC) symptomatology and have related behavioral (Toffolo, van den Hout, Hooge, Engelhard & Cath, 2013; Zhu *et al*., 2014), neural (Kubota, et al., 2016; Zhu et al., 2014) and neurocognitive (Sternheim, van der Burgh, Berkhout, Dekker & Ruiter,2014) outcomes to

the presence of OC symptomatology. Adding stereotype threat either to a formal diagnosis of OCD or to the presence of OC symptoms in individuals who have not been formally diagnoses may exacerbate the difficulties they experience in their everyday lives. To our knowledge, no research has been conducted about stereotype threat and OCD or OC tendencies, and in fact little has even been conducted about common stereotypes of the disorder. However, given the diagnostic criteria, it is possible that one common believable stereotype is that individuals with OCD or OC symptoms may struggle to function in a messy environment.

RESEARCH QUESTIONS & HYPOTHESES

Many studies have investigated and confirmed the existence and effect of stereotype threat in a variety of domains. Although some research has been done in the domain of mental illness, it was previously unknown whether stereotype threat can affect individuals with Obsessive Compulsive symptoms. The present study addresses this question by examining the possible effect of stereotype threat on individuals who are high in OC symptoms. Because of the wide range of settings in which the literature has shown stereotype threat to occur, we expected that individuals high in OC symptoms would also be prone to stereotype threat. Although no prominent stereotypes about OC exist in the literature, we predicted that a believable stereotype to present to participants would be that individuals with OC symptomatology do not function well in messy environments. We hypothesized that, with presentation of this stereotype, individuals high in OC symptoms placed in a messy environment would experience stereotype threat and perform less well on tests of cognitive functioning compared to those low in OC symptoms. We also hypothesized that, in a clean environment, the performance of individuals high in OC symptoms would not differ from those low in OC symptoms on the same test measures because no stereotype threat would occur.

Methods

PARTICIPANTS

We recruited 48 Butler students, 26 high in OC symptoms and 22 low in OC symptoms, to participate in our study. To recruit participants we contacted the Student Disabilities listserv and posted entries in the Butler Connection with information about how to complete a questionnaire to determine

eligibility for the study. Additionally, we included a questionnaire within the pretesting packets completed by students enrolled in psychology courses at the beginning of the fall and spring semesters. This questionnaire, OCI-R (OCI-R; Foa et al., 2002) helped identify students who were high or low in OC symptoms. This 18-item, 5-point Likert scale questionnaire asked respondents to indicate how much each experience had distressed or bothered him or her during the past month on a scale from "Not at all" to "Extremely." The items combine to form six subscales: hoarding, checking, order, germs, thoughts, and counting. Students with all six subscale scores at or below 2.00 were invited to participate in the study as part of the Low OC group. To be included in the High OC group, participants needed an Order subscale score of 4.00 or above. Students who were enrolled in psychology classes had the choice of either extra credit or a $10 gift card as a thank you for their participation. Those who were not in psychology courses were each given a $10 gift card at the completion of the session.

PROCEDURE

Testing sessions took place in a classroom in Jordan Hall. Participants were typically tested in small groups, with the number of participants per session ranging from 1 to 11. Group sessions included a mix of participants who were high and low in OC symptoms, and the classroom was set up in either a messy or clean fashion (see **Appendix**). The messy classroom was arranged so that the desks were out of alignment, there were old bottles and cups strewn around the room, there was partially erased writing on the white board, and the testing packets were stained with coffee. On the other hand, the clean condition classroom was set up in a meticulous fashion so that the desks were in perfect alignment and there was not a single piece of trash lying around. Participants were welcomed and informed of the study's procedures. They then gave their consent to participate. To highlight the stereotype threat, the study was described as examining "why messy environments interrupt cognitive performance for people with OCD." Then, each participant was provided with confidential, accurate information about his or her level of OC symptomatology based on the OCI-R. This feedback was as follows:

HIGH OC GROUP

> "Based on your responses to one of the questionnaires you completed about yourself, you qualify for participation in our study. Your endorsement of certain personal characteristics is above average

compared to all of the undergraduate students who completed the questionnaire. This does not mean that you have a diagnosable disorder. It only means that you have been assigned to a HIGH OBSESSIVE COMPULSIVE GROUP for the purposes of our research study.

This feedback is confidential. We will collect this sheet back from you before any test measures are administered, and we will shred it immediately following the testing session. You will answer one question about your group assignment on the questionnaires during the study. This is to make sure you know which group you are a part of for the purposes of our study. Your name will not be on any of the study questionnaires, which are only coded with your participant ID number in order to continue to keep this information confidential.

If you have any questions or concerns or would like referral information for a more formal clinical evaluation, you may contact one of the researchers listed on your consent form at any time in the future."

LOW OC GROUP

"Based on your responses to one of the questionnaires you completed about yourself, you qualify for participation in our study. Your endorsement of certain personal characteristics is below average compared to all of the undergraduate students who completed the questionnaire. This does not mean that you have a diagnosable disorder   It only means that you have been assigned to a LOW OBSESSIVE COMPULSIVE GROUP for the purposes of our research study.

This feedback is confidential. We will collect this sheet back from you before any test measures are administered, and we will shred it immediately following the testing session. You will answer one question about your group assignment on the questionnaires during the study. This is to make sure you know which group you are a part of for the purposes of our study. Your name will not be on any of the study questionnaires, which are only coded with your participant ID number in order to continue to keep this information confidential.

If you have any questions or concerns or would like referral information for a more formal clinical evaluation, you may contact one of the researchers listed on your consent form at any time in the future."

Participants then filled out a demographic questionnaire and a mood questionnaire before completing several tests of intelligence, concentration, and memory.

MATERIALS

Demographic questionnaire. Participants completed a brief questionnaire that asked about age, gender, class rank, race, and their OC classification. This assured that participants understood their feedback regarding their level of OC symptomatology correctly.

Center for Epidemiological Studies-Depression Scale (CES-D: Radloff, 1977). The CES-D included 20 questions that asked participants how often they had experienced specific symptoms of depression (e.g., "I felt lonely," "I felt that everything I did was an effort," or "I talked less than usual") during the prior week. Participants answered using a 4-point Likert scale ranging from 0 "rarely or none of the time" to 3 "most or all of the time." Possible scores ranged from 0-60, with higher scores indicating more depressive symptoms.

Shipley Institute of Living Scale (Zachary, 1986). This test required participants to identify synonyms of given words (Vocabulary Subscale) and to recognize and complete logical sequences (Abstraction Subscale). Possible scores ranged from 0-40 for Vocabulary and 0-20 for Abstraction, with higher scores indicating higher levels of verbal and non-verbal intelligence.

Written Digit Span. Adapted from a subtest of the Wechsler Memory Scale—Third Edition (Wechsler, 1997), participants heard a series of numbers read aloud. After they heard the numbers, participants attempted to write the numbers in the same order that they heard them. The number of digits in each sequence increased as the test progressed. Possible scores ranged from 0-16, with higher scores indicating greater concentration.

Sentence Construction (Hultsch, 1990). In this test, participants viewed individual sentences displayed on PowerPoint slides. As a group, they read each sentence aloud. In each sentence, one word was underlined, and participants attempted to remember those underlined words. After a certain number of slides, participants recalled all of the underlined words by writing them down in order, forming a sentence. Scores ranged from 0-54, with higher scores indicating greater concentration.

Immediate and Delayed Story Recall (Wechsler, 1945). During this subtest from the original Wechsler Memory Scale, participants heard a story and then wrote down word-for-word as many of the details of the story as they could recall. After approximately 25 minutes (during which time they completed other measures for the study), they again wrote down as much of the story as they could remember. Possible scores for each type of recall (immediate and delayed) ranged from 0-22 points, with higher scores representing greater memory of the story.

Results

SELF-CATEGORIZATION

Before running other analyses, we examined participants' self-identification of group on the demographic questionnaire to determine whether their self-identification after receiving their feedback matched with the actual feedback they were given (low or high OC). One participant in the High OC group incorrectly identified this classification and was excluded from all analyses; all other participants identified their classification correctly.

PARTICIPANT CHARACTERISTICS

Next, we examined the demographic characteristics of the remaining 25 High and 22 Low OC participants assigned to either the non-threatening clean or the threatening messy condition. **Table 1** summarizes the demographic characteristics of these four groups.

A 2 (OC: high versus low) × 2 (Condition: clean versus messy) analysis of variance showed that the four groups of participants were similar in age and class year, all main effect and interaction effect Fs (1, 43) < 1.4, n.s. Chi Square analyses demonstrated that the four groups were also similar in their distribution of gender ($\chi2$ (n = 47) = 2.2, p = .53) and race, $\chi2$ (n = 47) = 3.78, p = .29. A 2 (OC) × 2 (Condition) MANOVA with subscale scores from the Shipley Institute of Living Scale as dependent variables revealed that neither the OC groups (OC F (2, 42) <1), the participants assigned to the two conditions (condition F (2, 42) = 2.04, p = .14) nor the four groups created by their interaction (OC x condition F (2, 42) <1) differed in intelligence. However, a 2

Demographic Trait	Clean Environment		Messy Environment	
	Low OC ($n = 10$)	High OC ($n = 12$)	Low OC ($n = 12$)	High OC ($n = 13$)
Age	19.50 (1.35)	19.08 (1.17)	19.83 (1.12)	19.38 (1.39)
Year in College	2.20 (1.23)	2.00 (0.95)	2.58 (1.17)	2.15 (1.21)
Gender (% Female)	90%	83.34%	100%	84.62%
Race (% White)	90%	100%	100%	100%
Depression*	13.50 (7.91)	22.92 (11.94)	12.08 (5.95)	17.80 (9.69)
Shipley Vocabulary	29.00 (2.36)	29.08 (3.75)	31.08 (2.15)	30.46 (3.05)
Shipley Abstraction	17.10 (1.60)	17.58 (1.44)	17.33 (2.15)	17.23 (2.01)

* Participants high in OC symptoms had significantly higher levels of depressive affect than those low in OC symptoms, (F (1, 43) = 14.37, $p < .001$). Otherwise, the four groups were statistically equivalent in their demographic characteristics.

Table 1. Demographic characteristics of the sample.

× 2 ANOVA did indicate that the two OC groups differed significantly in their depressive affect (OC F (1, 43) = 14.37, $p < .001$), with high OC students endorsing more depressive symptoms than low OC students. Because of this, we used depression as a covariate in all subsequent analyses.

PRIMARY ANALYSIS

Our primary objective in this study was to determine whether the groups differed on the measures of neurocognitive functioning in a manner consistent with the effects of stereotype threat. To determine this, we ran a 2 (OC: high versus low) × 2 (Condition: messy versus clean environment) MANCOVA that included raw scores on digit span, sentence span, and both immediate and delayed story recall as outcome variables. Table 2 summarizes the scores of the four groups on each test measure. We included depression as a covariate since the two OC groups differed in their depressive affect.

Neither the main effect of OC (F (4, 39 < 1) nor the main effect of condition (F (4, 39) < 1) reached significance in the MANCOVA. However, a trend towards significance emerged in the OC by condition interaction, F (4, 39) =

Test Measure	Clean Environment		Messy Environment	
	Low OC ($n = 10$)	High OC ($n = 12$)	Low OC ($n = 12$)	High OC ($n = 13$)
Digit Span*	11.20 (1.75)	12.58 (1.88)	12.58 (2.39)	11.62 (2.06)
Sentence Span	50.40 (4.22)	48.67 (6.04)	49.50 (3.71)	52.08 (2.84)
Immediate Recall	5.85 (1.83)	6.67 (2.90)	7.13 (2.59)	6.23 (2.06)
Delayed Recall	5.45 (2.63)	5.75 (2.35)	7.00 (2.70)	5.89 (2.41)

* Participants high in OC performed less well than those low in OC in the messy environment on the digit span subtest, ($F (1, 22) = 6.65$, p =.02).

Table 2. Performance on neuropsychological tests by group and condition.
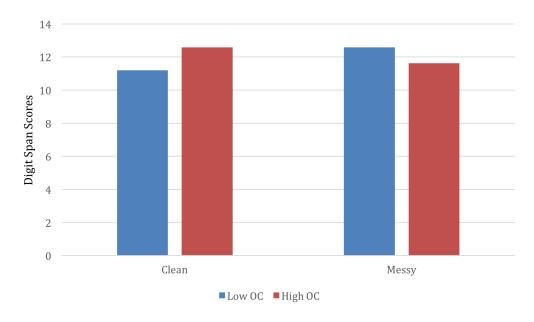

Figure 1. Scores on the Digit Span Test by group and condition.

2.29, $p = .08$, $\eta p2 = .19$. Because the effect size associated with this effect was moderate to large, we examined the univariate analyses to see if there were differences on individual tests. The univariate interaction effect reached significance for scores on the digit span subtest, $F (1, 42) = 4.66$, $p = .037$, $\eta p2 = .10$. Follow-up simple main effect analyses indicated that, consistent with

our hypothesis, participants high in OC symptoms performed less well than those low in OC symptoms in the threatening messy environment on the digit span subtest, $F(1, 22) = 6.65$, $p = .02$, $\eta p2 = .23$. In contrast, the two OC groups performed equivalently on the digit span subtest in the non-threatening clean environment, $F(1, 19) = 1.55$, $p = .23$, $\eta p2 = .08$. Figure 1 depicts this result.

Discussion

This study examined stereotype threat in regards to Obsessive Compulsive Disorder and its associated symptoms. Stereotype threat is a concept that has been studied in various contexts including: African Americans' intellect (Steele & Aronson, 1995), older adults' memory (Levy, 1996), and women's performance in math (Good, *et al.*, 2008). However, there is little research about stereotype threat in mental illness, and to our knowledge there have been no studies examining stereotype threat in individuals with OC symptoms. Our goal was to induce stereotype threat in college students with OC symptomatology to determine whether they perform less well on neurocognitive tests when they experience stereotype threat. Although only true for one of our outcome measures, consistent with our expectations, we found that participants high in OC symptoms tended to perform worse than participants low in OC symptoms in a threat inducing messy environment, but the two groups performed more similarly when in a non-threat inducing clean environment. These results suggest that individuals with OC symptomatology may be susceptible to stereotype threat, much like other vulnerable populations.

These findings are consistent with past research examining stereotype threat and mental illness. A previous study found that stereotype threat interfered with social skills of people with schizophrenia (Henry *et al.*, 2010) when they were told that a confederate was either aware or unaware of their diagnosis. The current study expands on this previous work by applying the phenomenon of stereotype threat to a new population—individuals high in OC symptoms. Additionally, Henry *et al.* measured social skills under threat and no-threat conditions, whereas we examined the impact of stereotype threat on neurocognitive abilities. We found that students with OC tendencies scored lower on the digit span subtest than their non-OC peers in a threatening (messy), but not a non-threatening (clean) environment.

Although our study supports the possible presence of stereotype threat in OC populations, we only found a significant OC by condition effect on one

of the four dependent variables included in our study in follow up to a near significant multivariate trend. Thus, the findings were not robust. Stereotype threat did not appear to impact immediate or delayed memory, and it only had an effect on one of the two concentration tests. The fact that digit span was vulnerable to threat but sentence construction was not could be due to the fact that the information was presented differently in the two tests. During the sentence construction test, the information was presented visually on a screen at the front of the room. Therefore, participants may have been focused on the screen rather than on the condition of the room. During the digit span test, however, the information was presented only orally. Thus, during this test of concentration participants may have been more likely to notice and feel threatened by the room condition.

There are some limitations to this study. The stereotype about individuals with OC tendencies that we presented to participants has not been documented in the literature and may not be a commonly held belief about OCD or its symptoms. Thus, it is possible that people may not have accepted the stereotype as true. Additionally, the sample size in our study was fairly small, especially because the sample was divided into four groups. We were pleased to identify 26 participants with OC tendencies, but a future study with larger samples might yield better evidence for stereotype threat in this population. Participants in this study were not required to have a formal OCD diagnosis. Therefore, it is possible that with a clinical sample, stronger results would emerge. Another limitation is that the two OC groups differed in their levels of depressive affect. Although we accounted for these differences by covarying depression, using groups matched in depression in future studies would better control for depression's potential influence on both cognitive performance and susceptibility to stereotype threat. Lastly, because there was only a strong trend in the MANCOVA, it is difficult to determine whether it reflects a true underlying difference between the groups based on the threatening versus non-threatening environment.

In addition to addressing the limitations already discussed, future studies could implement some additional features to help sort out potential explanations of differences between the groups. One possibility would be to tell some participants that the researcher is aware that they are high or low in OC symptoms and tell other participants that the researcher is unaware of their OC classification. Another possibility would be to create a more neutral environment that is neither messy nor clean but is described as messy in some conditions and as clean in other conditions. Yet another option would be to introduce a stereotype nullification condition to the study like that used by

Good, *et al.* (2008). Due to the anticipated small sample size of our study, we did not include these additional factors. However, expanding the study design in one or more of these ways would help determine whether stereotype threat accounts for the decreased concentration we documented in students with OC symptoms in the threatening messy environment.

Although it is not possible for the results of this study to completely separate the effect of stereotype threat from the effect of a messy environment on concentration, these results do indicate that messy environments can directly disrupt cognitive performance in students with OC symptoms. This has implications for educational settings in that these students may learn less well in messy environments compared to clean environments. Additionally, the discovery that stereotype threat could be a factor in OCD is important because OCD is a debilitating illness for individuals who experience it. If stereotype threat is exacerbating the symptoms of people with this disorder, it is important to find ways to diminish the effects of these stereotypes in order to help individuals with OCD to reach their full potential despite stereotypes other people may hold about them.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association. http://dx.doi.org/10.1176/appi.books.9780890425596

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. *J Exp Soc Psychol*, *35*(1), 29-46. http://dx.doi.org/10.1006/jesp.1998.1371

Foa, E.B., Huppert, J.D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G. & Salkovskis, P.M. (2002). The Obsessive Compulsive Inventory: Development and validation of a short version. *Psychol Assessment*, *14*(4), 485-96. http://dx.doi.org/10.1037/1040-3590.14.4.485

Good, C., Aronson, J. & Harder, J.A., (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *J Appl Dev Psychol*, *29*(1), 17-28. http://dx.doi.org/10.1016/j.appdev.2007.10.004

Henry, J. D., von Hippel, C., Shapiro, L. (2010). Stereotype threat contributes to social difficulties in people with schizophrenia. *Brit J Clin Psychol*, *49*(1), 31-41. http://dx.doi.org/10.1348/014466509X421963

Hultsch, D. F., Hertzog, C., Dixon, R. A. (1990). Ability correlates of memory performance in adulthood and aging. *Psychol Aging*, *5*(3), 356-68. http://dx.doi.org/10.1037/0882-7974.5.3.356

Kubota, Y., Sato, W., Kochiyama, T., Uono, S., Yoshinura, S., Sawada, R., Sakihama, M. & Toichi, M. (2016). Putamen volume correlates with obsessive compulsive characteristics in health population. *Psychiat Res-Neuroim*, *249*, 97-104. http://dx.doi.org/10.1016/j.pscychresns.2016.01.014.

Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *J Pers Soc Psychol*, *71*(6), 1092-1107. http://dx.doi.org/10.1037/0022-3514.71.6.1092

Linscott, R.J. & van Os, J. (2010). Systematic reviews of categorical versus continuum models in psychosis: Evidence for discontinuous subpopulations underlying a psychometric continuum. Implications for DSM-V, DSM-VI and DSM-VII. *Annu Rev Clin Psycho*, *6*, 391-419. http://dx.doi.org/10.1146/annurev.clinpsy.032408.153506

Merikangas, K.R., He, J., Burstein, M., Swanson, S.A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., and Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: Results from the National Cormorbidity Study-Adolescent Supplement (NCS-A). *J Am Acad Child Psyc*, *49*(10), 980-9. http://dx.doi.org/10.1016/j.jaac.2010.05.017

Obsessive-compulsive disorder (OCD). Diseases and Conditions. Retrieved December 23, 2013, from http://www.mayoclinic.org/diseases-conditions/ocd/basics/definition/con-20027827

Radloff, L.S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Appl Psych Meas*, *1*(3), 385-401. http://dx.doi.org/10.1177/014662167700100306

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *J Pers Soc Psychol*, *69*(5), 797-811. http://dx.doi.org/10.1037/0022-3514.69.5.797

Sternheim, L., van der Burgh, M., Berkhout, L.J., Dekker, M.R. & Ruiter, C. (2014). Poor cognitive flexibility, and the experience thereof, in a subclinical sample of female students with obsessive-compulsive symptoms. *Scand J Psychol*, *55*(6), 573-7. http://dx.doi.org/10.1111/sjop.12163

Stone J., Lynch C., Sjomeling, M., & Darley, J. (1999). Stereotype threat effects on black and white athletic performance. *J Pers Soc Psychol*, *77*(6), 1213-27. http://dx.doi.org/10.1037/0022-3514.77.6.1213

Substance Abuse and Mental Health Services Administration (2013). Results from the 2012 National Survey on Drug Use and Health: Mental Health Findings. NSDUH Series H-47, HHS Publication No. (SMA) 13-4805. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Toffolo, M.B.J., van den Hout, M.A., Hooge, I.T.C., Engelhard, I.M. & Cath, D.C. (2013). Mild uncertainty promotes checking behavior in subclinical Obsessive-Compulsive Disorder, *Clin Psychol Sci*, *1*(2), 103-9. http://dx.doi.org/10.1177/2167702612472487

Wechsler, D. (1945). *Wechsler Memory Scale*. San Antonio: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Memory Scale* (3rd ed.). San Antonio: The Psychological Corporation.

Zachary, R. A. (1986). *Shipley Institute of Living Scale: Revised Manual*. Los Angeles: Western Psychological Services.

Zhu, C., Yu, F., Ye, R., Chen, X., Dong, Y., Li, D....Wang, K. (2014). External error monitoring in subclinical obsessive-compulsive subjects: Electrophysiological evidence from a gambling task. *PLOS One, 9*(3): e90874. http://dx.doi.org/10.1371/journal.pone.0090874.

Appendix A