

Volume 10, Number 8

Submitted: November 15, 2004

First Revision: January 25, 2005

Second Revision: January 31, 2005

Accepted: February 1, 2005

Published: February 1, 2005

REPUTATION OF INTELLECTUAL INFERIORITY UNDERMINES MEMORY EFFICIENCY AMONG COLLEGE STUDENTS

Marion Dutrevis

Universite Blaise Pascal, CNRS, France

Jean-Claude Croizet

Universite de Poitiers

ABSTRACT

Can a group reputation of intellectual inferiority interfere with memory efficiency? Students targeted by a stereotype of lower ability had to memorize letter-number pairs (C-49) in a task presented either as a test of learning ability (diagnostic condition) or as a laboratory exercise (non diagnostic condition). The difficulty of the task was manipulated through the time available to memorize the material. Results showed that when the task was difficult, students under the diagnostic condition performed worse than students under the non-diagnostic condition. When the task was easy, the reverse pattern was observed. The implications of this finding are discussed.

INTRODUCTION

There is now ample scientific evidence showing that stereotypes that depict certain social groups as intellectually inferior can disrupt academic achievement (Steele, 1992; Steele, Spencer & Aronson, 2002). This phenomenon, referred to as stereotype threat (Steele & Aronson, 1995), poses that under evaluative scrutiny, people of low status groups can be concerned about confirming their group reputation of inferiority. These worries create an additional burden susceptible to interfere with the resolution of the tasks during an exam or a test, which ultimately will result in underachievement (see Croizet, Despres, Gauzins, Huguët, Leyens & Meot, 2004; Schmader & Johns, 2003). Yet when the stake of confirming the stereotype is low (i.e., the task is not perceived as diagnostic of intelligence), such concerns are minimized, and performance is restored.

Stereotype threat has generated more than 100 publications in less than 9 years (Source: Psycinfo). The phenomenon has been replicated across a wide range of groups known for their higher risk of academic underachievement: African Americans (e.g., Steele & Aronson, 1995), students from low socioeconomic status (e.g., Croizet & Dutrevis, 2004; B. Spencer & Castano, 2004), girls in mathematics (e.g., Keller & Dauenheimer, 2003), Hispanic Americans (e.g., Gonzales, Blanton & Williams, 2002). Consistently, research shows that under standard testing situations, that is when the evaluative purpose of the test is made explicit, these groups show a performance decrement. However, their performance increases when the identical test is not presented as an evaluation of intellectual ability, contesting the claim that their underachievement is simply rooted in inherent lower ability. The debilitating nature of stereotype threat has been observed on a wide range of tasks, including standardized-like tests (e.g., Steele & Aronson, 1995) IQ like tests (e.g., Croizet et al., 2004; Lovaglia, Younggreen, Lucas, Nath, Ruström, & Willer, 2004), or arithmetic computation (e.g., Schmader & Johns, 2003). In other words, the generalizability of stereotype threat is well established.

Though there is accumulative evidence that this predicament disrupts reasoning, to our knowledge no research has examined whether the salience of a reputation of intellectual inferiority can by itself impair another basic academic skill: Memorization. Yet, memorization of material constitutes an important aspect of academic requirement and research has shown that memorization (and more properly here associative learning) skills do predict academic achievement even beyond IQ (see Mackintosh, 1997). Can stereotype threat impair specifically how students memorize? The study presented below was specifically designed to address this issue.

We asked undergraduate students to participate in an experiment where they had to learn associations of non-significant material (i.e., a letter and a two-digit number, for example "C 49") through several memorization sessions. Students were selected from a college major targeted by a reputation of intellectual inferiority. We chose French psychology undergraduates because Croizet et al. (2004) showed that these students are suspected of low intellectual ability, as compared to medical school students. Even though these students are aware of it, they do not endorse this reputation. Yet this study revealed that in an evaluative context they can suffer from it and experience a drop in performance (see Croizet et al., 2004). In the present study, the characterization of the memorization task was manipulated. For one half of the psychology students, it was presented as a measure of learning ability whose purpose was to better understand the cause of the poor achievement rate of psychology majors compared to other (medical school) majors (i.e., the diagnostic condition). For the other half, the task was simply introduced as a cognitive psychology pre-test of experimental material; it was stressed that ability measurement was not at stake (non diagnostic condition). We hypothesized that students who thought that the task measured their learning ability would experience fear of confirming the allegation of inferiority conveyed by their group reputation. This threat could then harm performance if the task demands are high (Steele, et al., 2002). However, as suggested by Spencer, Steele and Quinn (1999), when the odds of failing are low (i.e., the task is easy), the threat could actually heighten performance. To address specifically this issue, the task difficulty was manipulated through the time allotted to learn the pairs of non-significant material (10 vs. 20 seconds for each memorization session). It was expected that with the difficult task, the memorization performance of students reputed for their lower ability would decrease in the threatening context, that is when they perceive the task as diagnostic of their ability as compared to not diagnostic. However, when confronted with a less demanding task, students should do better when the task is introduced as diagnostic of learning ability than when it is not.

METHOD

Participants

Seventy-seven psychology undergraduates (62 females and 15 males, mean age 19.9) participated in exchange for course credit. Ultimately the data of two students was dropped from the analyses. One because she failed the manipulation check and the other because she expressed some suspicion about the cover story. Participants were randomly assigned to a 2 (intellectual characterization of the task: Diagnostic vs. non diagnostic) x 2 (level of task difficulty: Difficult vs. easy) between subjects participants.

Procedure

Participants arrived individually and were greeted by a female experimenter, and were seated in front of a computer. They were informed that they would perform a task that would involve learning some verbal material through several phases. For one half of the participants in the diagnostic condition, the task was introduced as a standardized measure of learning ability whose purpose was to inform the researchers on the reasons why the achievement rate for psychology majors was low compared to other majors, particularly medical school majors. It was then made explicit that their learning ability score was to be compared with that of other majors at the university. For the other half of the participants the task was presented as the fine-tuning of a software for further research on the cognitive psychology of memory. It was emphasized that individual ability was not at stake, instead that the software was being tested.

Participants were informed that the task consisted of a repetition of a two-step sequence. In the first step (memorization phase) participants would have to memorize random associations of letter and two-digit numbers (e.g., N 56, C 49). In the second step (judgement phase), they would be presented with pairs (e.g., N 35) and would have to indicate whether the pair was correct (i.e., presented in the memorization phase) or not. This sequence was repeated four times.

Participants were then provided with an example of the structured material to be learned (see Appendix 1). After controlling that the task was well understood, participants were left alone and the task as well as the instructions were administered through the computer using the Pyscope software (Cohen, MacWhinney, Flatt, & Provost, 1993). During the memorization sequence, a table containing the 6 pairs to be learned was displayed (see Appendix 2). Depending on the level of difficulty, participants had either 10 or 20 seconds for memorizing these pairs. They were then presented with a series of 36 pairs (see Appendix 3). Pairs were presented successively and randomly for each participant. The pairs were for one third correct and for the two other thirds incorrect. This sequence was repeated 5 times. Note however that in the first sequence, participants learned only one half of the material (i.e., the three first columns) and were tested on only 18 pairs (i.e., one half of the material to learn). The remaining sequences used the complete material (the six columns plus a test on 36 pairs). There were then 18 (sequence 1) + 144 (36 x 4) (sequences 2, 3, 4 and 5) pairs, which resulted altogether in 162 items. The number of pairs correctly identified and rejected constituted the main dependant measure.

After completing the task, participants were required to fill out the post-experimental questionnaire (see below) that would help the researchers in interpreting the data. They were then debriefed and thanked for their participation.

Post-Experimental Questionnaire

The post-experimental questionnaire included several sets of questions. Participants first filled out a self-handicapping scale adapted from Steele and Aronson (1995). They were required to indicate how many hours of sleep they got the night before, then on 7 point Likert-type scales how much stress they had been under in the past few days (from 1 "none" to 7 "a lot"), whether they were having personal worries (from 1 "none" to 7 "many"), and finally how their current physiological state was (for example "headache," "flu") (from 1 "very bad" to 7 "very good"). The purpose of these questions was to detect any differences in apprehension about performance among the participants. We expected that evaluation apprehension -- as measured by the degree to which participants claim impediments to performance (self-handicapping) -- would be higher for the participants in the diagnostic condition confronted to the difficult task, as a result of the risk of validating the negative stereotype through a lower performance. These self-handicapping measures (Jones & Berglas, 1978) provided information about the extent to which stereotype threat affects individuals' conscious experience of the challenge and difficulty of the test. The next set of scales (ranging from 1 "never" to 7 "often") measured participants' verbal report of threat: They were asked whether during the task they feared to make mistakes and whether they felt pressured. Again, we hypothesized that participants who thought the task aimed at assessing their learning efficiency and who would be confronted to the difficult task would report a higher sense of threat and pressure.

RESULTS

The condition-by-condition results are presented in Table 1.

Table 1. Means and Standard Deviations for Dependant Variables by Diagnosticity and Task Difficulty

Condition	Diagnostic/ easy task		Diagnostic/ difficult task		Non diagnostic/ easy task		Non diagnostic/ Difficult task	
	M	SD	M	SD	M	SD	M	SD
Performance	151.56	6.16	139.61	10.32	146.68	8.53	147.35	8.85
Hours of sleep	7.91	1.57	7.36	1.05	7.24	1.74	7.15	1.04
Stress	3.50	1.98	3.33	1.75	3.42	1.95	2.85	1.69
Personal worries	2.78	1.83	3.56	1.79	2.68	1.67	3.80	1.70
Physiological state	4.89	1.37	4.22	1.77	5.26	1.33	4.55	1.93
Fear of mistakes	3.72	1.56	4.11	1.37	4.05	1.18	3.85	1.66
Evaluative pressure	2.50	1.50	3.06	1.47	2.53	1.47	2.90	1.92

Memorization Performance

The number of correct answers out of 164 was submitted to an ANOVA with test description (diagnostic of memorization ability vs. not) and task difficulty (10 vs. 20 seconds to memorize the material) as between factors. The analysis first revealed a main effect of task difficulty, in that participants performing better in the easy task ($M = 149.05$, $SD = 7.77$) than those confronted to the difficult one ($M = 143.68$, $SD = 10.22$); $F(1, 71) = 8.04$, $p = .006$ (eta squared = .10). More interestingly, and as expected, the description of the task interacted with the level of difficulty in predicting performance, $F(1, 71) = 10.05$, $p = .002$ (Eta-squared = .12). In accordance with our hypothesis, a contrast analysis revealed that when the task was difficult, participants who thought the task was measuring their learning ability performed worse ($M = 139.61$, $SD = 10.32$) than the participants who thought the task was non diagnostic of their ability ($M = 147.35$, $SD = 8.85$), $t(71) = 2.77$, $p = .007$. On the contrary, when the task was easier, participant in the diagnostic condition tended to outperform ($M = 151.56$, $SD = 6.16$) those in the non diagnostic condition ($M = 146.68$, $SD = 8.53$), $t(71) = 1.72$, $p = .09$. Further analyses taking into account the different sequences as a repeated measure showed that these findings were observed for each sequence composing the memorization task. There was no interaction between task description, level of difficulty and sequence number. In other words, being targeted by a reputation of intellectual inferiority created a sense of threat that had important consequences on performance. When the task was easy, this threat enhanced memory efficiency. However, when it was more difficulty as we suspect it is often the case in school, this threat disrupted memory efficiency.

Self-Handicapping

We expected threatened participants who had been confronted with the difficult task to claim more impediments to performance because of their apprehension about validating the negative reputation of their group. However a series of 2 (task description) \times 2 (task difficulty) ANOVAs performed on each of the self-handicapping measures did not support our prediction. When the difficult task was presented as a measure of their memorization ability, participants did not claim more excuses for poor performance than the other participants (all p's for the interaction not significant). The analyses did yield some significant main effect of task difficulty. Thus, regardless of task description, participants confronted with the difficult version of the memorization task reported higher personal worries ($M = 3.68$, $SD = 1.73$) than those confronted with the easy version ($M = 2.73$, $SD = 1.73$), $F(1, 74) = 5.49$, $p = .02$ (Eta squared = .07). In the same vein, those confronted to the difficult task reported a poorer physiological state ($M = 4.39$, $SD = 1.84$) than those exposed to the easy version of the memorization task ($M = 5.08$, $SD = 1.34$), $F(1, 74) = 3.37$, $p = .07$ (Eta squared = .04). No other effect reached significance.

Self-Report of Threat

We expected threatened participants who had been confronted to the difficult task to report a fear of producing mistakes and an evaluative pressure during the task. However the 2 (task description) x 2 (task difficulty) ANOVAs performed on the reports of threat and pressure did not fit our prediction. When the difficult task was presented as a measure of their learning ability, participants did not report higher concern with producing errors or a higher evaluative pressure than other participants (all p's for the interaction not significant). No other effect reached significance.

DISCUSSION

Past literature has repeatedly demonstrated that being under the threat of reputation of lower ability can disrupt performance on standardized test-like exams (see Maass & Cadinu, 2003). The purpose of this research was to examine the role that stereotype threat and task difficulty may play on memory efficiency. We asked students to memorize associations of non-significant material. Results indicated that the fear of confirming allegations of inferiority brought on by a stereotype can, when the task is difficult, interfere with memory efficiency. Indeed, psychology majors, reputed for having low intellectual ability, had a lower memory performance when their reputation was made salient but only when the task was difficult. When the task was easier, being under the threat of a negative stereotype resulted in better performance. This finding has important implications at both theoretical and practical levels.

In accordance with the stereotype threat literature (see Steele et al., 2002), our finding shows that the difficulty of the task faced by individuals is crucial in determining whether the threat will impair performance. When the task is demanding, stereotype threat can compete with the task resolution and ultimately impair performance. However, when the task is less demanding, a threatening environment can result in heightened performance. Our finding replicates a recent study by O'Brien and Crandall (2003) that revealed that women under stereotype threat performed better than women in a control condition on an easy test but worse when the test was difficult. However, because the tasks used in this previous study differed on several dimensions with regards to the format (multiple choice vs. computational) and the timing (10 vs. 11 minutes), it is hard to know if it was difficulty alone or some intrinsic features distinguishing the two tasks that explained the findings. In our study, the format of the easy and difficult tasks was perfectly equated so that we can be confident that it is the level of difficulty per se that interacts with the task description to predict performance. This interaction is congruent with a conceptualization of stereotype threat in terms of arousal (see Easterbrook, 1959; O'Brien & Crandall, 2003). Arousal enhances performance on easy tasks but harms performance on difficult tasks (Brehm, 1999; Zajonc, 1965). Though the question of the mediation of stereotype threat remains open, the present study suggests that heightened arousal may be one of the mechanisms by which stereotype threat affects performance, as initially proposed by Steele and Aronson (1995).

Whereas the performance measures fit our expectations, self-reports on self-handicapping and perception of threat are more disappointing. Indeed, threatened participants confronted with the difficult version of the task did not claim more impediments for their poor performance nor did they report increased fear or pressure. The self-report results were mostly not significant except for some measures of self-handicapping that showed nevertheless a consistent pattern: participants confronted with the most difficult task reported having higher personal worries and a poorer physiological state than those confronted with the easier version of the task. This last finding shows that participants were able to use some cues provided by the situation to claim some impediments for their performance. However, whereas the level of task difficulty seemed to be consciously perceived and apprehended, test characterization did not show any effect. One possibility is that the psychological experience of stereotype threat is less conscious or less identifiable to lead to self-handicapping (Nisbett & Wilson, 1977). The other possibility is that self-report measures of threat, because they are open to social desirability bias, are simply poor indicators of people's feelings (Bosson, Haymovitz, & Pinel, 2004).

Our study is the first one to show that suspicion of inferiority can impair memorization among college students. In the present study, participants did not have to memorize complex information that requires intensive reasoning as is usually the case when students have to master the content of class before an exam. On the contrary, the material used in the present study was not significant and could not easily be connected to previous knowledge stored in memory. In other words, elaborate reasoning, like the kind involved when one takes a test, was not involved. Previous research has widely documented that elaborate reasoning can be disrupted by stereotype threat (Quinn & Spencer, 2001). We showed that it can harm another basic skill: memorization. It seems reasonable to assume that memorization shows some sensitivity to stereotype threat because it mobilizes the articulatory loop of working memory (Baddeley, 1986). Moreover several studies have shown that stereotype threat, because it drains cognitive resources, affects the efficiency of working memory (Schmader, 2004; Schmader & Johns, 2003).

At a more practical level, our study extends the range of academic skills that are vulnerable to stereotype threat. By showing that basic memorization can be disturbed when one is suspected of inferiority, our research reveals that the effects of stigmatization on academic achievement are probably largely underestimated. It suggests that students' attention and immediate retention during a class can be impaired because of stereotype threat. However further research will have to examine whether such a predicament can undermine the complex learning process that usually takes place within and outside the classroom.

REFERENCES

- Baddeley, A.D. (1986). *Working memory*. Oxford: OUP.
- Bosson, J.K., Haymovitz, E.L., & Pinel, 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.
- Brehm, J.W. (1999). The intensity of emotion. *Personality and Social Psychology Review*, 3, 2-22.

Cohen J.D., MacWhinney B., Flatt M., and Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25, 257-271.

Croizet, J.-C., Desprès, G., Gauzins, M.-E., Huguet, P., Leyens, J.-P., & Méot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin*, 30, 721-731.

Croizet, J.-C., & Dutrévis, M. (2004). Why test scores do not equal merit. *Journal of Poverty*, 8, 91-107.

Easterbrook, J.A. (1959). The effects of emotion on cue utilization and the organization of behavior. *Psychological Review*, 66, 183-201.

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, 659-670.

Jones, E.E., & Berglas, S. (1978). Control of attributions about the self through self-handicapping strategies : The appeal of alcohol and the role of underachievement. *Personality and Social Psychology Bulletin*, 42, 200-206.

Keller, J., & Dauenheimer, D. (2003). Stereotype threat in the classroom : Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin*, 29, 371-381.

Lovaglia, M. J., & Younggreen, R., Lucas, J. W., Nath, L. E., Rutstrom, E., & Willer, D. (2004). Stereotype threat and differential expected consequences : Explaining group differences in mental ability scores. *Sociological Focus*, 37, 107-125.

Maass, A., & Cadinu, M. (2003). Stereotype threat : When minority members underperform. *European Review of Social Psychology*, 14, 243-275.

Mackintosh, N. J. (1997). Has the wheel turned full circle? Fifty years of learning theory, 1946-1996. *The Quarterly Journal of Experimental Psychology*, 50A, 879-898.

Nisbett, R.E., & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.

O'Brien, L.T., & Crandall, C.S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, 29, 782-789.

Quinn, D.M., & Spencer, S.J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues*, 57, 55-71.

Schmader, T. (2004). Gender identification moderates stereotype threat effects on women’s math performance. *Journal of Experimental Social Psychology*, 38, 194-201.

Schmader, T., Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, 85, 440-452.

Spencer, B. & Castano, E. (2004, January). The threat of a "Classless society": Extending the concept of stereotype threat to social class in America. *Proceedings of Columbia University, Teachers college 21st Annual Winter Roundtable on Cultural Psychology and Education: Strategies for building cultural competence in psychology and education*, New York, NY.

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

Steele, C.M. (1992). Race and the schooling of Black Americans. *The Atlantic Monthly*, 269, 68-78.

Steele, C.M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of african americans. *Journal of Personality and Social Psychology*, 69, 797-811.

Steele, C.M., Spencer, S.J., & Aronson, J. (2002). Contending with group image : The psychology of stereotype threat and social identity threat. In M.P. Zanna, *Advances in experimental social psychology* (Vol 34). San Diego, CA : Academic Press.

Zajonc, R.B. (1965). The requirements and design of a standard group task. *Journal of Experimental Social Psychology*, 1, 71-88.

APPENDIX 1: EXAMPLE TABLE GIVEN TO PARTICIPANTS BEFORE TASK

A	B	C	D	E	F
1	2	3	4	5	6

APPENDIX 2: ACTUAL TABLE PRESENTED TO PARTICIPANTS FOR MEMORIZATION TASK

N	W	I	H	L	J
56	39	27	48	61	13

APPENDIX 3: PAIRS PRESENTED TO PARTICIPANTS TO TEST MEMORIZATION EFFICIENCY (EACH PAIR PRESENTED TWICE IN A GIVEN SEQUENCE)

- N 56
- W 39
- I 27
- H 48
- L 61
- J 13
- N 45
- W 17
- I 34
- H 63
- L 52
- J 29
- O 68
- M 42
- R 36
- C 21
- F 15
- K 54

APPENDIX 4: CORRELATION MATRIX (PEARSON CORRELATIONS)

	Mean	SD	1	2	3	4	5	6	7	8	9
1	1.52	.50	1.00	.01	.08	-.15	-.08	.03	.11	.01	-.02
2	1.51	.50	.01	1.00	-.29**	-.11	-.11	.27**	-.21	.03	.15
3	146.33	9.43	.08	-.29**	1.00	-.01	-.04	-.07	.19	-.20*	-.14
4	7.39	1.38	-.15	-.11	-.01	1.00	.05	-.16	.12	.25**	.24**
5	3.27	1.83	-.08	-.11	-.04	.05	1.00	.41***	-.27**	.35***	.13
6	3.21	1.78	.03	.27**	-.07	-.16	.41***	1.00	-.35***	.06	.06
7	4.73	1.64	.11	-.21	.19	.12	-.27**	-.35***	1.00	-.07	-.15
8	3.93	1.44	.01	.03	-.20*	.25**	.35***	.06	-.07	1.00	.72***
9	2.75	1.59	-.02	.15	-.14	.24**	.13	.06	-.15	.72***	1.00

*p < 0.10 (2-tailed).

**p < 0.05 (2-tailed).

***p < 0.01 level (2-tailed).

- 1 Task description
- 2 Level of difficulty
- 3 Performance
- 4 Hours of sleep
- 5 Stress
- 6 Personal worries
- 7 Physiological state
- 8 Fear of mistakes
- 9 Evaluative pressure

AUTHORS' BIOGRAPHIES

Marion Dutrevis is a PhD student at the Universite Blaise Pascal (Clermont-Ferrand, France). Her primary research interest is stereotype threat. dutrevis@srvpsy.univ-bpclermont.fr

Jean-Claude Croizet is professor at the University of Poitiers (France). His main research focuses on the impact of stigmatization on academic achievement. croizet@srvpsy.univ-bpclermont.fr