FOLIA 273

Annales Universitatis Paedagogicae Cracoviensis

Studia ad Didacticam Mathematicae Pertinentia 10(2018)

ISSN 2080-9751
DOI 10.24917/20809751.10.10

Anna Rędzio

Gender stereotypes concerning mathematical abilities: threat or challenge?*

Abstract. Stereotypes concerning women's mathematical abilities are widespread and can become an obstacle for some women to succeed in this domain, thus, they can be one of the possible reasons for the gap between men and women in mathematics. There is significant empirical evidence confirming that women in the situation of stereotype threat (i.e. when they are afraid of confirming that they are less gifted in mathematics than men are) perform worse than their colleagues who are not threatened in this way. There is also empirical evidence that another psychological phenomenon, intellectual helplessness in mathematics, is a predictor of school achievement in mathematics as well as general intellectual capabilities are. An experimental study with women active in STEM (graduates at university of engineering) as participants was conducted to test the relationships between intellectual helplessness experienced in mathematics classes and vulnerability to stereotype threat. All participants completed Intellectual Helplessness Inventory and afterwards they all received the same tasks to perform. One group was informed those tasks were diagnostic of mathematical abilities (experimental condition, stereotype threat induced) and the other group received the information that those tasks were testing their perceptiveness (control condition, no threat induced). The results indicated that there was an interesting correlation: Women who exhibited lack of intellectual helplessness in mathematics reacted paradoxically to stereotype threat (improved instead of impairing their performance). However, those who reported higher intellectual helplessness in mathematics, reacted typically to stereotype threat, i.e. performed worse than the control group. These results suggest that intellectual helplessness can be a factor impacting vulnerability or immunity to stereotype threat, but further research confirming this dependency is necessary.

*2010 Mathematics Subject Classification: Primary: 97C60

Keywords and phrases: stereotypes, gender stereotypes, intellectual helplessness in math, stereotype threat, women in STEM

1. Introduction

One could argue that mathematics is the most 'democratic' domain: That everyone, regardless of their ethnicity, skin color, gender, socio-economic status or place of residence have similar opportunity to become a good mathematician, at least at school level. After all, to achieve this goal, you don't need to have the best and the most expensive books, nor education tools. You don't have to live in a big city, nor to have the best teachers (self-teaching of mathematics skills is possible). You don't have to remember historical dates, difficult names, nor complicated spelling. You don't need any special gifts like a good ear for music or physical abilities. All you need to be a good mathematician, is understanding. Finding the rules which already exist in your mind, in everybody's mind. One could believe that.

Thus, why does the reality look completely different? Why is it that lots of students don't understand mathematics and experience anxiety just thinking about mathematics lessons or even imagining that they are performing a mathematical task? Why do some children (and adults too) believe that they can't solve simple equations? Last, but not least, why do the gaps in STEM (Science, Technology, Engineering, Math) between women and men or between people living in the country and those living in big cities, still exist? To find the answers, we should transgress mathematics and look at it from socio-psychological point of view. In this chapter, a few possible psychological causes for existing inequalities in STEM are to be presented. They are: intellectual helplessness in mathematics and stereotype threat. Additionally, the results of own research concerning the relationship between these phenomena and mathematics performance are to be presented.

Stereotype threat. Let's try to imagine the following situations: 1) Bill, an African American boy living in the USA, is taking an intelligence test, and 2) Amanda is taking part in a chess tournament and she is the only woman there among about 30 men. There are widespread stereotypes that the African Americans are less intelligent than the anglo/European Americans (this belief is especially common in the USA) and that women are worse chess players than men. What could Bill and Amanda think and feel in such situations? We could imagine that Bill would think: "If I do this task worse than my white classmates, everyone would believe that African Americans are stupid". Amanda might be anxious: "They will all look at me and if I play poorly, they would think: of course, females are not able to think logically and play chess as well as their male counterparts are. So, their beliefs concerning male and female talents would be confirmed". They could both feel apprehension and if they really performed poorly (which in such situation happens very often) Bill would start to doubt his own intelligence and Amanda her chess abilities. This is a phenomenon called stereotype threat (ST). Steele and Aronson (1995) have defined it as a situation when one is afraid of confirming a negative stereotype about one's social group (e.g. women, people of colour, or people with low socioeconomic status).

It is worthy of note here that ST can happen only in situations that demand using someone's competences in the counter-stereotypic domain. It is impossible when, for example, a Pole is abroad and he might be perceived thru the lens of stereotype that Poles are hard drinkers. It is enough for him to drink moderately to disconfirm this stereotype. After all, everyone is able to drink moderately, so he

does not experience ST. But when the stereotype concerns competences of one's social group, one feels that the only way to disconfirm it is a good performance in the counter-stereotypic domain: a domain, that, according to existing stereotypes, is difficult for one's social group, for example, mathematics, chess playing and logic for women or intelligence and verbal activities for African Americans. For some people, good performance in such domains can seem very difficult or even impossible.

The direct effect of ST is performance deterioration in the counter-stereotypic domain. That means, people under ST pressure, perform worse than they might in a normal situation, according to their normal intellectual possibilities. It happens because anxiety, self-doubt and intrusive thoughts linked to negative stereotypes usually disturb individuals from performing well. The first study concerning ST was conducted in 1995 by Joshua Aronson and Claude M. Steele. They asked the participants (black and white undergraduates) to solve a task. A group of the participants were told that this was a diagnostic task assessing their intellectual ability (diagnostic condition - threat situation for the black participants) and the other undergraduates were informed that this was a laboratory tool for studying problem solving (non-diagnostic condition - no threat for anyone). In both groups, about half of the participants were black. In the diagnostic condition (threat for the African Americans), significant performance differences between the black and the white subjects occurred: The white participants outperformed the African-American participants. But in the non-diagnostic condition (no threat) the African Americans performed as well as the white participants (and as the white participants in the diagnostic condition). That was the effect of ST and that is the way this phenomenon usually works: people perform worse in the stereotypic domain.

The results of other studies indicate that ST may concern many different social groups and domains. For example, people with low socio-economic status under ST exhibited performance impairment in intelligence tests (Croizet, Claire, 1998; Croizet, Millet, 2011) or elderly people in such condition performed worse in short memory tasks (Hess, Hinson, Statham, 2004; Levy, Leifheit-Limson, 2009). Of specific importance here is that females in the threat condition (ST activated) usually underperform females in the control condition (no ST) and male participants regardless of condition in mathematics tests (Spencer, Steele, Quinn, 1999; McJunkin, 2009; Steinberg, Okun, Aiken, 2016).

In the study by Spencer, Steele, and Quinn the participants (female and male undergraduates) in the experimental condition (threat for women) were taking mathematics test described as producing gender differences. In the other condition, the participants were told the test didn't produce gender differences (threat alleviation condition). In the other experiment, McJunkin (2009) divided the participants (only female undergraduate students) into three groups. All of them had to do a mathematics test, however, the participants in the first group were told the test indicated differences between men and women (threat induced, $n = 18$), the ones in the second group weren't told anything about that test ($n = 16$), and the participants in the last group were told this was gender-fair test (exhibited no differences between sexes; threat alleviated; $n = 17$). As predicted, in both

studies (by Spencer et al. and by McJunkin) the participants in the threat condition underperformed the participants in the threat alleviation condition. But what is maybe even more important in the study by McJunkin is that the differences between the threat and the baseline condition were not significant, however, the differences between them and the last group (threat alleviated) were large and significant (for the means and standard deviations see Table 1).

Table 1. Tasks solved correctly in different conditions in McJunkin's (2009) study: means and standard deviations

Condition	M	SD
Threat induced	5.94	3.10
No information concerning test	7.25	2.02
Threat alleviated	11.41	3.76

These results indicate that ST can exist permanently and it might be enough for some females just to know that they are taking a difficult math test to perform worse. It is not even necessary to suggest them that the test is gender-biased or that their scores would be compared to men's scores. Of course, such information could reinforce the effect of ST, but these results indicate that this effect can appear not just in laboratories, in artificially created situations, but that ST is an every day phenomenon which can be induced in every mathematics lesson, every mathematics test, exam, or other mathematics task performance. That's why it is very important to improve the existing knowledge about this phenomenon: only if we know its mechanisms, its reasons, and who is especially vulnerable or immune to ST, we can overcome its negative effects on the mathematics performance of females and reduce the gap between women and men in STEM (Science, Technology, Engineering, Math). It is also important to convince mathematics teachers and schoolgirls that such phenomena exist. The results of some studies indicate that individuals who are aware that it exists are more immune to the negative effects of ST (Johns, Schmader, Martens, 2005).

Intellectual helplessness. Classical model of learned helplessness. Intellectual helplessness (IH) is a kind of learned helplessness. The first research on learned helplessness (Overmier, Seligman, 1967) concerned behavior of dogs that were given electric shocks. Some of them could avoid the shocks and learned it quickly but the others could not, though they were attempting, so they learned that nothing they did had any effect on experienced pain. In the next part of the experiment all the dogs were given a chance to avoid painful shocks, but only the group that could avoid the shocks before, learned the new way to do it. The dogs that weren't given any chance to avoid pain before, did not even try anymore – they were lying passively, waiting for the shock. This is exactly what we call learned helplessness. Studies concerning learned helplessness were conducted on animals (Overmier, Seligman, Overmier, Seligman; Maier, Seligman, 1976) and on humans too (Hiroto, Seligman, 1975; Seligman, 1975).

Informational model of learned helplessness. A similar mental state is possible not just after receiving painful shocks but in every situation where one makes multiple attempts to do something without any effect. In their experiments, Sędek and Kofta (1990; see also Kofta, Sędek, 1989) indicated that informational

helplessness training (e.g. many unsuccessful trials to solve unsolvable tasks) generated performance impairment on subsequent tasks. A lot of data show that such deficits were linked to minimizing of effort, withdrawal, or self-handicapping as an ego-protective maneuver (von Hecker, Sędek, 1999). The crucial mechanism of informational helplessness training is the same as that in the classical model of learned helplessness: an individual learns that nothing s/he does can bring any effects (Sędek, 1995). The further analyses indicated that cognitive exhaustion was the process accounting for performance deterioration after informational helplessness training. Normally, people dealing with problem-solving situations are likely to engage in systematic mental activity: They attempt to understand the task's demands and meaning, they notice information important for task solution, they draw conclusions, detect regularities and inconsistencies (Rydzewska et al., 2016). In the beginning of an uncontrollable situation (e.g. when they start to deal with unsolvable tasks) they proceed the same way but the inability to achieve task solutions despite long and continuous mental effort brings them to a state of cognitive exhaustion. Their ability to build mental models, to systematize knowledge, to generate hypothesis and to form new ideas is diminished (Kofta, 1993; Kofta, Sędek, 1989; Sędek, Kofta, 1990; Sędek et al., 1993; Rydzewska et al., 2016) and they tend to avoid effortful information-gathering strategies (Sędek et al., 1993). That's why they cannot perform well on the next tasks, even if these tasks are feasible. Cognitive exhaustion states are especially disruptive to more complex, cognitive-demanding tasks, which require non-routine and flexible thinking (Rydzewska et al., 2016).

IH at school. IH is a long-term state described as permanent lack of engagement in one domain, minimizing of effort linked to tasks in this domain, cognitive exhaustion and avoiding challenges in this domain (Sędek, 1995). It appears often in school-age and it is related mainly to a teaching style and teachers' competencies to communicate the knowledge (Rydzewska et al., 2016; Sędek, 1995). But, of course, at school no one wants pupils to perform unsolvable tasks, so how can IH appear? Even without informational helplessness training, the students can experience prolonged cognitive efforts without cognitive gain (the necessary condition for learned helplessness): if they conscientiously make multiple attempts to understand new concepts or to solve a task without being successful, without progress. The repeated inabilities to understand new materials during the same lessons (e.g. during math classes) can block active problem thinking, for example, involving comparison, reasoning or building knowledge schemas (von Hecker, Sędek, 1999) and, as the consequence, it can evoke a permanent IH state. It is worthy of mention here, that IH is not just a general lack of abilities: it can happen even for gifted students. It is not the lack of engagement in learning in general, either: IH can concern just one or a few domains, for example, some students may be helpless in mathematics but masterly in their natural language or vice-versa, and, as results of studies by Sędek (1995) indicated, IH in mathematics is not correlated to intellectual helplessness, for example, in the Polish language.

IH in mathematics and mathematics anxiety as predictors of school achievement in mathematics. IH in mathematics in positively and strongly correlated to mathematics anxiety, a feeling of tension, apprehension, or fear that interferes with math performance (Ashcraft, 2002). But they are quite different phenomena. Mathematics anxiety is an emotional state while IH has mainly the cognitive component, but it also has motivational and emotional components. Emotions experienced in mathematics anxiety (fear, tension) are also different than the emotions experienced in the state of IH (discouragement, disengagement, weariness). Mathematics anxiety and IH in math are significant predictors of school grades in mathematics as well as intellectual capabilities are (Rydzewska et al., 2016). Thus, it was hypothesized that they are also predictors of mathematics performance, especially under ST: low levels of them would reduce or even eliminate the cognitive deterioration, higher would increase it. IH in mathematics is also strongly and negatively correlated to self-evaluation of methematical abilities and mathematics engagement, thus, it was hypothesized that they would be predictors of mathematics performance under ST.

2. Study

Assumptions. The purpose of the present study was to test the relations between vulnerability to stereotype threat (ST) and intellectual helplessness (IH) in the counter-stereotypic domain and some variables correlated to it: psychological threat linked to the same domain, self-evaluation of abilities and engagement in this domain. The most common and direct effect of ST is deterioration of performance in the stereotypic domain. The long-term effects of this state are disengagement (Major, Spencer, Schmader, Wolfe, Crocker, 1998) or disidentification (Spencer, Steele, Quinn, 1999; Steele, 1997; Steele, Aronson, 1995) from the domain in question. That is, the individual excludes performance in that domain as a basis for self-esteem. The other consequence of multiple exposure to negative stereotype might be disidentification from the group targeted by negative stereotype (Bedyńska, 2013; Davies et al., 2002; Steele, 2007; Woodcock et al., 2012). The existing empirical evidence concerning ST indicates that this phenomenon can have mainly negative effects on performance. There are hardly few studies confirming that it can have positive direct effects, too (Ben-Zeev et al., 2005; Jamieson, Harkins, 2007, 2009; Crisp et al., 2009) and their results tell us only about positive effects of ST emerging under specific conditions (e.g., the tasks to perform must be easy, demanding prepotent response or at best moderately difficult).

However, it is well known in psychology that sometimes threat situation (e.g. ST) can evoke paradoxical effects. According to the Transactional Model of Stress Coping (Lazarus, Folkman, 1991) or to Challenge and Threat Theory (Blascovich, Mendes, 2000, 2010), an individual in the threat situation evaluates the situation's demands and his or her possibilities to cope with them. If these possibilities are (regarding to the subjective evaluation, not objectively) sufficient to cope with the demands of the situation, the individual can experience an extraordinary mobilization of cognitive, emotional and motivational resources. As a consequence, s/he would be able to proceed better (e.g. run faster, or think more creatively) than

usual, in situations without threat. However, if one evaluates his or her resources as not sufficient to cope with the situation, one would perform worse under threat than normally. According to these theories, ST should sometimes have positive (if one evaluates highly one's possibilities to perform in the counter-stereotypic domain) and sometimes negative effects (if one evaluates those possibilities as low). This statement was the basis for the current study. It was hypothesized that ST can have positive effects on performance as well as negative. It was also hypothesized that IH in the stereotypic domain, psychological threat linked to this domain, self-evaluation of skills and engagement in this domain would be the factors responsible for typical (performance decrement) vs. paradoxical (performance improvement) response to ST. Since our participants were women active in STEM, our operationalized hypotheses were:

1. The participants who report low levels of IH in mathematics and psychological threat experienced during mathematics lessons (PT), would perform better under ST than the participants in no-threat (NT) condition.

2. The participants who report relatively high levels of IH in mathematics and PT, would perform worse under ST than their counterparts in NT condition.

3. IH in mathematics and PT would be negatively correlated to performance in ST condition but they would not in NT condition.

4. Self-evaluation of mathematical abilities (SEMA) and mathematics engagement (ME) would be positively correlated to performance in ST but not in NT condition.

3. Method

Participants. Twenty-nine female students of Rzeszów University of Technology, aged 22-24 (M = 22.72, SD = 0.75) participated in the experimental study which had been conducted during their class, with the consent of the lecturer. The major of all the participants was mathematics. It was important for the study, because, as it was assumpted according to Lazarus' and Folkman's theory of stress coping, only individuals who every day have to do with the stereotypic domain and feel confident in it are likely to response to ST in an adaptive way and improve, instead or impair, their performance.

Measures. The measure for IH, psychological threat experienced during mathematics lessons, self-evaluation of mathematical abilities and mathematics engagement was the Intellectual Helplessness Inventory (Krejtz, 2012; for the full version see Sędek, McIntosh, 1998), including intellectual helplessness scale, psychological threat scale, self-evaluation of one's abilities scale and engagement scale. This inventory can be used not just for measuring IH and other variables regarding to mathematics, but to every other school domain – it is enough to change "Mathematics" into "Polish" or "Physics".

The intellectual helplessness scale consists of 10 statements (the full version has 20 statements) concerning feelings and thoughts that had accompanied the

participants during mathematics classes in high school, e.g. *I found I didn't understand what I was writing in my notes; I felt empty-headed; I felt like it was all Greek for me.* The answers were given on a 5-point Likert scale (1 = *never*, 5 = *every time*).

The psychological threat scale consists of 8 statements about apprehension and tension that had been experienced by the participants during mathematics classes in high school, e.g. *I felt tension from the very beginning of the lesson; I was especially afraid of tests, I was afraid that I can get a poor grade.* The answers were given on a 5-point Likert scale (1 = *never*, 5 = *every time*).

The self-evaluation of one's own mathematical abilities scale is a 2-item scale, including statements: *I think I'm good in math* and *I am mathematically talented.* The answers were given on a 5-point Likert scale (1 = *I completely disagree*, 5 = *I completely agree*).

The math-engagement scale is a 2-item scale, including statements: *Being good in math is very important to me* and *Math was a crucial domain for me.* The answers were given on a 5-point Likert scale (1 = *I completely disagree*, 5 = *I completely agree*).

The means of scores in each of the four scales were basis for establishing four independent variables: intellectual helplessness experienced during mathematics classes (IH), psychological threat experienced during mathematics classes (PT), self-evaluation of mathematical abilities (SEMA) and math engagement (ME).

The dependent variable's measure. The dependent variable in our study was vulnerability to ST, and its measure was performance in two conditions: under ST and without ST. Fifteen tasks from Raven's Progressive Matrices, Version Standard Plus (Raven, Styles, Raven, 1998; Polish Version by Jaworowska, Szustrowa, 2000), all Series C and 3 tasks of Series D performance were used in this experiment. The measure for dependent variable was the number of tasks solved correctly in 10 minutes. The Raven's Progressive Matrices was convenient for the study for a few reasons. First, it could be presented as test diagnostic of mathematical abilities as well as perceptiveness task and this was necessary for the manipulation task. In fact, this test is usually described as general intelligence task, but it demands inductive thinking and finding hidden schemas which are the basis of mathematical abilities. Second, the problems in Raven's Progressive Matrices are cognitively demanding and it was important for this study, since there is already empirical evidence for paradoxical response to ST if the tasks are easy (Ben-Zeev et al., 2005; Crisp et al., 2009) or demanding prepotent response (Jamieson, Harkins, 2007, 2009). Last, but not least, it was normalized on big samples, and its reliability is high.

Manipulation task. After completing the Intellectual Helplessness Inventory, the participants received tasks with instruction. The tasks were the same for everyone, but the instructions differed. Thirteen participants received instructions describing the test as "diagnostic of mathematical innate abilities and part of a bigger research concerning differences between men and women in this domain" (ST condition). It induced threat on three different ways. First, participants' gender was made salient and as many studies indicate (see e.g. Steele, Aronson, 1995) making the stereotypic identity salient is enough to induce ST. Second, the par-

ticipants were all mathematics students, thus it was important to them to exhibit mathematical abilities, however, the possibility to do the tasks wrong was threatening for them. Third, the instruction suggested (considering existing stereotypes) that females do these tasks worse than males do. The other participants ($n = 16$) received instruction describing this test as the task testing their perceptiveness (no-threat, NT condition: perceptiveness is not an important trait for the mathematicians, thus the possibility of exhibiting its lack is not threatening, the gender was not made salient and there are no stereotypes that women are less perceptive than men are). After 10 minutes, the participants were thanked and given the debriefing notes.

4. Results

Intellectual helplessness in math (IH), psychological threat experienced during math lessons (PT), mathematical abilities self-evaluation (SEMA) and math engagement (ME). The participants in general reported low level of IH ($M = 2.01$; $SD = 0.42$; spread: $1.27-3.00$). They also reported low level of PT ($M = 2.03$, $SD = 0.63$; spread: $1.14 - 3.50$). They exhibited high levels of MAE ($M = 4.10$, $SD = 0.49$; spread: $3.00 - 5.00$) and ME: $M = 4.40$, $SD = 0.57$; spread: $3.50 - 5.00$). There were no significant differences in any of these variables between the conditions (all $ps > .13$).

Task performance. On average, the participants correctly solved 9.25 tasks ($SD = 2.25$). The highest result was 14 and the lowest was 6. There was no difference in performance between ST and NT condition ($p > .89$).

Correlations. The independent variables (IH, PT, MAE, ME) were strongly and significantly correlated to Raven's Progressive Matrices Performance and to each other (see Table 2).

Table 2. Correlations between independent variables and Raven Test performance ($N = 29$)

	IH	PT	SEMA	ME	Raven Test
IH		$r = .656^{**}$ $p = .000$	$r = -.546^{**}$ $p = .002$	$r = -.712^{**}$ $p = .000$	$r = -.542^{**}$ $p = .002$
MA	$r = .656^{**}$ $p = .000$		$r = -.399^{*}$ $p = .032$	$r = -.436^{*}$ $p = .018$	$r = -.510^{**}$ $p = .005$
SEMA	$r = -.546^{**}$ $p = .002$	$r = -.399^{*}$ $p = .032$		$r = .677^{**}$ $p = .000$	$r = .418^{*}$ $p = .024$
ME	$r = -.712^{**}$ $p = .000$	$r = -.436^{*}$ $p = .018$	$r = .677^{**}$ $p = .000$		$r = .490^{**}$ $p = .007$
Raven Test	$r = -.542^{**}$ $p = .002$	$r = -.510^{**}$ $p = .005$	$r = .418^{*}$ $p = .024$	$r = .490^{**}$ $p = .007$	

* p < .05 (2-way significance); ** p < .01 (2-way significance); N = 29

However, when separated analyses for each of the two conditions had been conducted, we could see that only in ST condition IH, PT, SEMA, and ME were significantly correlated to Raven Test performance. There were no significant correlations in NT condition (see Table 3 and Table 4).

Table 3. Correlations between Raven Test performance and other variables in ST condition (N = 13)

	IH	PT	SEMA	ME
Pearson r	$-.773^{**}$	$-.706^{**}$	$.650^{*}$	$.802^{**}$
2-way significance	.002	.007	.016	.001

Table 4. Correlations between Raven Test performance and other variables in NT condition ($N = 16$)

	IH	PT	SEMA	ME
Pearson r	$-.281$	$-.376$	.028	.169
2-way significance	.292	.152	.919	.530

IH as well as PT were strongly and negatively correlated to performance in ST condition. In the same condition, SEMA was moderately and positively correlated to performance. ME and performance was also positively correlated, but this correlation was strong. However, in NT condition none of these correlations was significant.

When the independent variables have been dichotomized (below average = 0, above average = 1), significant differences in performance in ST condition appeared for each of them but PT. The participants who reported IH lower than average in the whole group, SEMA and ME higher than average, outperformed these participants who reported IH higher, SEMA and ME lower than average. They also scored higher in Raven's Matrices Test than participants in NT condition (both who reported lower and higher than average levels of IH, SEMA and ME). In NT condition none of those differences was significant nor even marginally significant (see Figure 1 and Table 5).

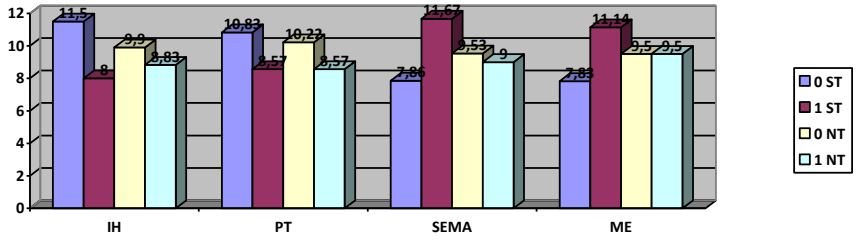

Fig. 1.

Figure 1. Means in performance (number of tasks solved correctly) for each condition and each value of dichotomized variables. 0 ST: reported variable below average, ST condition. 1 ST: reported variable above average, ST condition. 0 NT and 1 NT, respectively.

Table 5. Test t values and ps for dichotomized variables in both conditions

Condition		IH	PT	SEMA	ME
NT	Test t value	1.233	1.649	0.238	0.000
	p	.243	.121	.815	1.000
ST	Test t value	3.502	1.762	−4.277	−3.132
	p	.005	.106	.001	.010

5. Conclusions

There was no main effect of experimental condition: participants in ST condition scored on the same level that participants in NT condition did. One possible reason for it could be that the experimental manipulation did not work, did not evoke ST. However, the correlations observed in ST condition and absent in NT condition disconfirm such explanation. There were strong and negative correlations of IH and PT with performance in ST condition and positive correlations of SEMA, ME, and performance. All of them were significant in spite of small groups. None of them was significant nor even marginally significant in NT condition. Moreover, when the independent variables have been dichotomized (below vs. above average), significant differences in performance occured in ST condition, but not in NT condition. In ST condition the participants who reported IH lower than average, SEMA and ME higher than average outperformed the participants in ST who reported higher IH, lower SEMA and ME. They also scored higher than the participants in NT condition.

These correlations and differences suggest that the manipulation was successful but a part of participants exhibited immunity, or even paradoxical response to stereotype threat. The participants who reported low IH and PT, high SEMA and ME were more likely to perform under ST as well as their counterparts in NT condition or even better (to react paradoxically). However, the participants who reported higher IH and PT and lower SEMA, were likely to perform worse under ST (the typical reaction). Thus, the results of the current study indicate that 1) The paradoxical response to ST is more common than the existing evidence indicates, and 2) The susceptibility to ST among women active in STEM might be linked to IH in mathematics, and a few factors correlated to IH: psychological threat experienced during mathematics lessons in high school, mathematical abilities self-evaluation, and mathematics engagement.

6. Limitations

The first serious limitation of this study was the sample size ($N = 29$). This was a rather small group, even for an experimental study, so its results should be considered warily until they are replicated on a larger group of participants. Further, the results confirming that some individuals are immune to ST or can even paradoxically respond to it (i.e. perform better under threat than without it) cannot be generalized to other social groups. For example, African Americans are affected by ST differently than women are (see e.g. Logel, Peach, Spencer, 2011). Thus, the results could be different if we conducted an experiment with ST

manipulation among participants of colour or other stigmatized group, for example elderly people or people with low social-economic status. The results obtained in our study cannot even be generalized to females in another situation (e.g. female drivers: in such situation, there might be other factors impacting the susceptibility to ST than learned helplessness or psychological threat) or to females in other countries (perhaps there are some important cultural factors that contribute to vulnerability to ST).

Finally, these results cannot be considered as evidence for an argument that some individuals overcame ST and became immune to all its effects. It is only the evidence that some people right after exposure to ST can improve their performance. These results do not show that such reaction is possible (and dependably on the same factors) over a long-term period. We also do not know the emotional and cognitive costs of the paradoxical response to ST. Perhaps, in the long term, the costs of disconfirming the negative stereotypes are so significant that they discourage even the most talented females in the STEM domain from pursuing careers in this domain, despite the fact that they seem to perform very well on tests. These results cannot be seen as a "proof" that there are individuals who efficiently overcame ST or, that there are ways to overcome this phenomenon. This study is just a small step in the research concerning overcoming ST.

7. Future directions

The results obtained in this study might be an inspiration to many other studies. ST is a phenomenon, which has been investigated for more than two decades, but there are still many unresearched aspects of this state. For example, its relations to IH in the stereotypic domain has been tested in just a few studies (e.g. Bedyńska, 2013). It is an important and very interesting topic and it should be taken into consideration by psychologists. The subject of paradoxical response to ST is also a subject that demands many more studies. For example, the mediators of paradoxical effect of ST should be explored for, i.e. what are the processes responsible for the paradoxical effect of ST? Perhaps it is the extraordinary mobilization of cognitive, emotional and motivational resources and this hypothesis should be tested in future research. But it is also possible that these processes are different and they should be found in some explorative studies.

Of course, it is necessary to replicate the results of the present study on other social groups, in other countries and in different circumstances, for example, after a different manipulation task. It would also be worth investigating whether the paradoxical response to ST can occur when the individuals are asked to perform the tasks in a prescribed order, e.g. starting from the difficult one followed by easier tasks (in the presented experiment, particiapnts did the test starting from the easiest tasks and, perhaps, solving them correctly boosted their later performance). The response to ST can depend on the type of tasks and their difficulty level. Thus, it could be interesting to conduct similar experiments but with other kind of tasks or with extremely difficult tasks (the tasks used in this study had different difficulty levels: some of them were easy, some moderately difficult, and some difficult but none of them were extremely difficult).

ST is one of the possible explanations for the gap between males and females in STEM. But, even if some women are able to overcome this state evoked by single stereotype activation and perform very well on tests or exams, the more important question remains: Do they, in spite of widespread stereotypes, really believe in their possibilities in STEM so strongly that they seriously consider pursuing a career in the STEM field? Have they overcome not just direct effects of ST, but also become immune to its long-term consequences? This is a very important topic for research. If they did not, the overcoming of temporary ST cannot bring much benefit for women active in mathematics nor for any other social groups targeted by different negative stereotypes. In such situation traits, moderators and other factors for overcoming the long-term effects of ST should be sought.

References

Ashcraft, M. H.: 2002, Math anxiety: Personal, educational, and cognitive consequences, *Current Directions in Psychological Science* **11**(5), 181–185.

Bedyńska, S.: 2013, Długofalowe skutki zagrożenia stereotypem: Moderacyjna rola identyfikacji z własną grupą i z dziedziną, *Studia Psychologiczne* **51**, 53–61.

Ben-Zeev, T., Fein, S., Inzlicht, M.: 2005, Arousal and stereotype threat, *Journal of Experimental Social Psychology* **41**, 174–181.

Blascovich, J., Mendes, W. B.: 2000, Challenge and threat appraisals: The role of affective cues, in: J. P. Forgas (ed.), *Feeling and thinking: The role of affect in social cognition*, Cambridge University Press, Paris, 59–82.

Blascovich, J., Mendes, W. B.: 2010, Social psychophysiology and embodiment, in: S. T. Fiske, D. T. Gilber (ed.), *The handbook of social psychology, Ed. 5*, Wiley, New York.

Crisp, R. J., Bache, L. M., Maitner, A. T.: 2009, Dynamics of social comparison in counter-stereotypic domains: Stereotype boost, not stereotype threat, for women engineering majors, *Social Influence* **4**, 171–184.

Croizet, J. C., Claire, T.: 1998, Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic background, *Journal of Personality and Social Psychology* **24**, 588–594.

Croizet, J. C., Millet, M.: 2011, Social class and test performance. From stereotype threat to symbolic violence and vice versa, in: M. Inzlicht, T. Schmader (ed.), *Stereotype Threat: Theory, Process, and Application*, Oxford University Press, New York, 188–201.

Davies, P. G., Spencer, S. J., Quinn, D. M., Gerhardstein, R.: 2002, Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally, *Personality and Social Psychology Bulletin* **28**, 1615–1628.

Hess, T. M., Hinson, J. T., Statham, J. A.: 2004, Explicit and implicit stereotype activation effects on memory: Do age and awareness moderate the impact of priming?, *Psychology and Aging* **19**, 495–505.

Hiroto, D. S., Seligman, M. E. P.: 1975, Generality of learned helplessness in man, *Journal of Personality and Social Psychology* **31**, 311–327.

Jamieson, J. P., Harkins, S. G.: 2007, Mere effort and stereotype threat performance effects, *Journal of Personality and Social Psychology* **93**, 554–564.

Jamieson, J. P., Harkins, S. G.: 2009, The effects of stereotype threat in the solving of quantitative GRE problems: A mere effort interpretation, *Personality and Social Psychology Bulletin* **35**, 1301–1314.

Jaworowska, A., Szustrowa, T.: 2000, *TMK - Test Matryc Ravena w wersji Standard – forma Plus*, Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego, Warszawa.

Johns, M., Schmader, T., Martens, A.: 2005, Knowing is half the battle. Teaching stereotype threat as a means of improving women's math performance, *Psychological Science* **16**, 175–179.

Kofta, M.: 1993, Uncertainty, mental models, and learned helplessness: An anatomy of control loss, w: G. Weary, F. Gleicher, K. L. Marsh (red.), *Control motivation and social cognition*, Springer, New York, 122–153.

Kofta, M., Sędek, G.: 1989, Repeated failure: A source of helplessness, or a factor irrelevant to its emergence?, *Journal of Experimental Psychology: General* **118**, 3–12.

Krejtz, I.: 2012, *Korepetycje poznawcze. Rola pamięci roboczej i kontroli uwagi w przewidywaniu osiągnięć szkolnych*, Sedno, Warszawa.

Lazarus, R., Folkman, S.: 1991, *The concept of coping. Stress and coping: An anthology (3rd Ed.)*, Columbia University Press, New York.

Levy, B. R., Leifheit-Limson, E.: 2009, The stereotype-matching effect: greater influence on functioning when age stereotypes correspond to outcomes, *Psychology and Aging* **24**, 230–233.

Logel, C., Peach, J., Spencer, S. J.: 2011, Threatening gender and race: Different manifestations of stereotype threat, in: M. Inzlicht, T. Schmader (ed.), *Stereotype Threat: Theory, Process, and Application*, Oxford University Press, New York, 159–172.

Maier, S. F., Seligman, M. E. P.: 1976, Learned helplessness: Theory and evidence, *Journal of Experimental Psychology: General* **105**, 3–46.

Major, B., Spencer, S., Schmader, T., Wolfe, C., Crocker, J.: 1998, Coping with negative stereotypes about intellectual performance: The role of psychological disengagement, *Personality and Social Psychology Bulletin* **24**, 34–50.

McJunkin, L.: 2009, Effects of stereotype threat on undergraduate women's math performance: Participant pool vs. classroom situations, *Emporia State Research Studies* **45**, 27–31.

Overmier, J. B., Seligman, M. E. P.: 1967, Effects of inescapable shock upon subsequent escape and avoidance learning, *Journal of Comparative and Physiological Psychology* **63**, 23–33.

Raven, J. C., Styles, I., Raven, M. A.: 1998, *Raven's Progressive Matrices: SPM plus test booklet*, Oxford Psychologists Press/San Antonio, TX: The Psychological Corporation, Oxford, England.

Rydzewska, K., Rusanowska, M., Krejtz, I., Sędek, G.: 2016, Uncontrollability in the classroom: The intellectual helplessness perspective, in: M. Bukowski, I. Fritsche, A. Guinote, M. Kofta (ed.), *Coping with Lack of Control in a Social World*, Routledge, Abingdon, 62–80.

Seligman, M. E. P.: 1975, *Helplessness: On depression, development, and death*, Freeman, San Francisco.

Sędek, G.: 1995, *Bezradność intelektualna w szkole*, IPPAN, Warszawa.

Sędek, G., Kofta, M.: 1990, When cognitive exertion does not yield cognitive gain: Toward an informational explanation of learned helplessness, *Journal of Personality and Social Psychology* **58**, 729–743.

Sędek, G., Kofta, M., Tyszka, T.: 1993, Effects of uncontrollability on subsequent decision making: Testing the cognitive exhaustion hypothesis, *Journal of Personality and Social Psychology* **65**, 1270–1281.

Sędek, G., McIntosh, D. N.: 1998, Intellectual helplessness: Domain specificity, teaching styles, and school achievement, in: M. Kofta, G. Weary, G. Sędek (ed.), *Personal Control in Action: Cognitive and Motivational Mechanisms*, Plenum Press, New York, 391–418.

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.: 1997, A threat in the air: How stereotypes shape intellectual identity and performance, *American Psychologist* **52**, 825–838.

Steele, C. M., Aronson, J.: 1995, Stereotype threat and the intellectual test performance of African American, *Journal of Personality and Social Psychology* **69**, 797–811.

Steinberg, J. R., Okun, M. A., Aiken, L. S.: 2016, Calculus GPA and math identification as moderators of stereotype threat in highly persistent women, *Basic and Applied Social Psychology* **34**, 534–543.

von Hecker, U., Sędek, G.: 1999, Uncontrollability, depression and the construction of mental models, *Journal of Personality and Social Psychology* **77**, 833–850.

Woodcock, A., Hernandez, P. R., Estrada, M., Schultz, P. W.: 2012, The consequences of chronic stereotype threat: Domain identification and abandonment, *Journal of Personality and Social Psychology* **103**, 635–646.

Robert B. Zajonc Institute for Social Studies
University of Warsaw
ul. Stawki 5/7
PL-00-183 Warszawa
e-mail: anna.redzio@psych.uw.edu.pl