Sex differences in neural efficiency: Are they due to the stereotype threat effect? ☆

Beate Dunst [a,*], Mathias Benedek [a], Sabine Bergner [a,b], Ursula Athenstaedt [a], Aljoscha C. Neubauer [a]

[a] Department of Psychology, University of Graz, Austria
[b] Department of Leadership and Entrepreneurship, University of Graz, Austria

ABSTRACT

The neural efficiency hypothesis postulates a more efficient use of brain resources in more intelligent people as compared to less intelligent ones. However, this relationship was found to be moderated by sex and task content. While the phenomenon of neural efficiency was previously supported for men when performing visuo-spatial tasks it occurred for women only when performing verbal tasks. One possible explanation for this finding could be provided by the well-studied phenomenon called stereotype threat. Stereotype threat arises when a negative stereotype of one's own group is made salient and can result in behavior that confirms the stereotype. Overall, 32 boys and 31 girls of varying intellectual ability were tested with a mental rotation task, either under a stereotype exposure or a no-stereotype exposure condition while measuring their EEG. The behavioral results show that an activated negative stereotype not necessarily hampers the performance of girls. Physiologically, a confirmation of the neural efficiency phenomenon was only obtained for boys working under a no-stereotype exposure condition. This result pattern replicates previous findings without threat and thus suggests that sex differences in neural efficiency during visuo-spatial tasks may not be due to the stereotype threat effect.

© 2013 The Authors. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Individual differences in cognitive performance can be elucidated from different perspectives. The personality based approach takes an ability perspective in attributing performance differences to stable traits. This point of view differs from the social psychological approach, which acknowledges that cognitive performance can be affected by various state factors. As trait and state effects can be reflected in test scores (Wicherts, Dolan, & Hessen, 2005; cf. Cronbach, 1957) we will conjoin both perspectives when investigating sex differences in neural efficiency (negative IQ-brain activation relationship; e.g., Haier et al., 1992).

According to the individual differences/trait perspective, sex differences in neural efficiency would be attributed to sex differences in the underlying ability domain (women typically show higher verbal ability, men show higher spatial ability). The neural efficiency hypothesis is represented by an IQ-brain activation correlation, which can be moderated by sex and task content

(Jaušovec & Jaušovec, 2008; Lipp et al., 2012; Neubauer, Fink, & Schrausser, 2002; Neubauer, Grabner, Fink, & Neuper, 2005): Males and females showed the expected inverse IQ-brain activation relationship primarily in those tasks in which they usually perform better, i.e. males in visuo-spatial tasks and females in verbal and emotional intelligence tasks. This certainly holds true for the visuo-spatial domain where considerable evidence demonstrates that men usually outperform women (for a review cf. Halpern et al., 2007). However, with respect to the verbal domain it is more complex. While women stereotypically think to perform better in verbal tasks evidence for an actual performance difference is rather mixed (Halpern, 2004).

Sex differences in task performance can be explained by ability factors as well as by situational factors. Moreover, ability differences can have genetic causes but also long-term environmental causes (Halpern et al., 2007; Steele, 2010). In particular, performance can be influenced by implicitly activated stereotypes. A stereotype threat arises in a situation in which the stereotype is relevant and the situation strikes one as a test of stereotype-relevant qualities. Steele (1997) proposed that a negative stereotype about a group to which one belongs leads to fear, self-doubt, which in turn may impair working memory and hamper cognitive performance. Empirical evidence demonstrates for instance that, White males underperform in athletics (Stone, 2002) and women underperform in math and science domains (Good, Aronson, & Harder, 2008; Spencer, Steele, & Quinn, 1999). According to Steele

* Corresponding author. Address: Department of Psychology, University of Graz, Maiffredygasse 12b, 8010 Graz, Austria. Tel.: +43 316 380 5120; fax: +43 316 380 9811.

E-mail address: beate.dunst@uni-graz.at (B. Dunst).

(2010), negative stereotypes can have long term consequences, leading to a loss of interest and eventually diminishing math or spatial ability in the long run.

Although research has shown that activated negative stereotypes may impair performance and lead to fear and self-doubt, little is understood about the mechanism that accounts for these effects. Stereotype threat effects have been explained within different frameworks such as the mere effort account (Jamieson & Harkins, 2007), the disruptive mental load (Croizet et al., 2004), the attentional control theory (Eysenck, Derakshan, Santos, & Calvo, 2007) or the arousal-based theory (O'Brien & Crandall, 2003).

The integrated process model (Schmader, Johns, & Forbes, 2008) attempted to integrate existing frameworks for explaining stereotype threat effects. It assumes that interrelated cognitive, physiological and affective processes can impair executive resources thus hampering efficient processing. In an fMRI study by Wraga, Helt, Jacobs, and Sullivan (2007), the confrontation with a negative stereotype about one's own group resulted in impaired performance and in raised activation of amygdala as well as in reduced activity in brain regions associated with high performance in spatial ability (e.g., ventral and medial portions of anterior prefrontal cortex). Additionally, increased activation in the rostral-ventral anterior cingulate cortex (a region associated with emotional self-regulation) and the right orbital gyrus (a region associated with social knowledge) were found. Similar results were found by Krendl, Richeson, Kelley, and Heatherton (2008). These results largely support behavioral research showing that coping with negative stereotype related emotions seize cognitive resources that could otherwise be used for cognitive tasks (Schmader & Johns, 2003; Schmader et al., 2008). In other words, women may underperform under stereotype threat because valuable cognitive resources are spent on emotional regulation and thereby reducing working memory capacity.

1.1. Research question

The main aim of this study was to examine whether sex differences in neural efficiency could be attributed to the stereotype threat effect. In this study a visuo-spatial task is selected, since there exist robust sex differences and stereotypes regarding visuo-spatial performance, especially in mental rotation (for a review cf. Halpern et al., 2007). Furthermore, visuo-spatial skills are a fundamental element in STEM (Science, Technology, Engineering, and Mathematics) which indicates the practical significance (Lubinski, 2010) of this study. Lubinski (2010) even suggested that selecting students for advanced learning opportunities in STEM without considering spatial ability might be unprogressive. Therefore, several attempts have been made to discover the origins of sex differences in spatial ability. Women working on visuo-spatial tasks might be affected by implicitly activated stereotypes resulting in higher arousal (cf. O'Brien & Crandall, 2003). Moreover, higher arousal could lead to higher and more diffuse brain activation which then would oppose efficient processing.

We assume that the stereotype threat may affect brain activation differentially in women according to their individual level of intellectual ability. It can be hypothesized that high IQ women (who sense the task easier than low IQ women generally show lower brain activation according to the neural efficiency hypothesis) confronted with the stereotype show increased brain activation because they feel challenged to disprove this stereotype (cf. Jaušovec & Jaušovec, 2008). Low IQ women may also strive to disprove the stereotype, but their already high level of arousal (due to their perception of increased task difficulty) may limit a further increase of activation. As a consequence IQ and brain activation would be no longer correlated in women under stereotype threat, which would explain why neural efficiency in visuo-spatial tasks

has only been found for men but not for women. Therefore, this study aims at testing whether stereotype threat is partly responsible for sex differences in neural efficiency. To this end, neural efficiency during visuo-spatial processing shall be investigated under two experimental conditions, either involving an explicit stereotype threat or involving no stereotype threat. If behaviorally a stereotype threat can be elicited and if the above described sex difference in neural efficiency can be found only in the threat condition then it might be concluded that the particular threat is responsible for sex differences in neural efficiency.

2. Method

2.1. Participants

Out of a pool of 929 participants, 63 healthy Austrian adolescents (31 girls and 32 boys aged between 15 and 18 years) were selected to represent a large variability in figural intelligence participated in the study. All participants were IQ-matched between experimental groups in order to avoid a confounding. The sample showed an average IQ of 100.50 (SD = 15.52), and there were no differences in figural IQ, neither between sex groups ($F(1,54) = 0.04$, $p = .84$; $M_{girls} = 101.11$, $SD_{girls} = 17.59$; $M_{boys} = 100.26$, $SD_{boys} = 13.89$) nor between stereotype exposure conditions (stereotype exposure vs. no-stereotype exposure) ($F(1,54) = 0.17$, $p = .68$; $M_{non\_st} = 99.83$, $SD_{non\_st} = 17.55$; $M_{st} = 101.54$, $SD_{st} = 13.21$). Prior to the study, participants provided written informed consent (for underage students it was provided by their parents). Participation was voluntary and students received €20 for participation. The data of 5 persons were excluded from the analysis either because of excessive EEG artefacts or because they disagreed to one of the two following statements: (1) "I am good at math" and (2) "It is important to me that I am good at math", leaving a total of 58 participants (26 girls and 32 boys).

2.2. Experimental task

A mental rotation task was employed, in which participants were presented 48 pairs of Shepard-Metzler (SM) figures. Participants' task was to judge whether the figures were congruent or incongruent. In order to come to the correct solution, SM figures have to be rotated mentally until the main axis points in the same direction, before it can be decided whether the pair of figures is identical or not (i.e., mirror images). All SM-figures were presented in a 3D presentation mode. The 3D presentation mode was employed because a mental rotation task in a 3D presentation mode seems to create fair conditions for both sexes (Neubauer, Bergner, & Schatz, 2010).

2.3. Experimental design

A 2 × 2 design was employed using the between-subject factor SEX and STEREOTYPE EXPOSURE (stereotype exposure vs. no-stereotype exposure). Participants of both sexes were randomly assigned to one of the two experimental conditions. The experimental manipulation was part of the written task instruction, which was presented prior to working on the task. In the stereotype exposure condition, students received the message that boys perform better. ("This test measures your visuo-spatial ability. Recent studies demonstrated that in this task boys usually perform better than girls. That means that girls solve fewer items than boys.") This information reflects a stereotype threat for girls and a stereotype lift for boys. Participants working under the no-stereotype exposure condition were informed that in the particular task no sex differences exist. ("This test measures your vi-

suo-spatial ability. Recent studies demonstrated that girls perform equally well as boys in this test.") These instructions were adapted from prior studies which successfully investigated the stereotype threat effect (e.g., Moè & Pazzaglia, 2006).

2.4. EEG-recording/analyses

The EEG was measured by gold electrodes with 9 mm in diameter. Thirty-three electrodes were placed according to the international 10–20 system. A ground electrode was placed on the forehead, a reference electrode on the tip of the nose. To measure eye movements, an electrooculogram (EOG) was recorded bipolarly between two diagonally placed electrodes above and below the inner and the outer canthus of the right eye. EEG impedances were kept below 5 kΩ; EOG below 10 kΩ. All signals were sampled at a frequency of 512 Hz. During recording a bandpass (0.1–100 Hz) as well as a 50 Hz notch-filter in order to avoid power line contaminations were applied (all apparatus distributed by BrainProducts GmbH, Gliching/GER).

The raw EEG was corrected for ocular artefacts by means of a regression-based algorithm (Gratton, Coles, & Donchin, 1983) using the software Brain Vision Analyzer (1.05; BrainProducts Gmbh, Gliching/GER). Remaining artefacts were removed by visual inspection. Further analysis steps were performed by means of a set of Matlab scripts (R2011b; The MathWorks, Inc.). The bandpower of the EEG ($\mu V2$) was computed by means of a time–frequency analysis employing a Fast Fourier-transformation (FFT) with a window size of 1000 ms and an overlap of 900 ms. For each trial the EEG band power in the upper alpha band (10–12 Hz) was computed as this alpha frequency band is particularly sensitive to task- and ability-related effects (Grabner, Fink, Stipacek, Neuper, & Neubauer, 2004). We decided to use a fixed alpha band rather than an individually defined band in order to ensure comparability with previous studies. Changes in cortical activation were quantified by means of task-related power (TRP) changes between reference and activation phases for each electrode and trial (Pfurtscheller & Lopes da Silva, 1999).

In Fig. 1 a schematic display of one trial and the EEG measurement intervals are depicted. The presentation of the fixation cross (3 s) marked the beginning of each trial. After the 3 s, the stimulus presentation started (max. 8 s) and the participants had to respond as fast and accurately as possible. Each response was followed by an inter-trial interval of 4 s. The time during the presentation of the fixation cross served as reference interval (3 s) for the TRP calculation. As activation interval the time window from the stimulus onset until the reaction (max. 8 s) was defined. For the TRP calculation only correctly solved trials were used. Task-related power at an electrode i was obtained by subtracting the log-transformed power during the activation interval ($POW_{i,activation}$) from the log-transformed power during the reference interval ($POW_{i,reference}$) according to the formula: $TRP_{(i)} = \log(POW_{i,reference}) - \log(POW_{i,activation})$. Negative values therefore reflect increases in power from reference to activation (subsequently referred to as desynchronization), positive values reflect decreases (referred to as synchronization; cf. Pfurtscheller & Lopes da Silva, 1999).

For further analysis, the TRP data was aggregated from different electrode positions in the following way (cf. Neubauer et al., 2005): frontal left (FP1, AF3, F3, F7), frontal right (FP2, AF4, F4, F8), frontocentral left (FC1, FC5, C3), frontocentral right (FC2, FC6, C4), centroparietal left (CP1, CP5, P3), and centroparietal right (CP2, CP6, P4), parietooccipital left (PO3, PO5, O1), parietooccipital right (PO4, PO6, O2), temporal left (T3, T5), and temporal right (T4, T6). The midline electrodes (FZ, CZ, PZ) were not included in the analyses as the hemispheric differences were of interest.
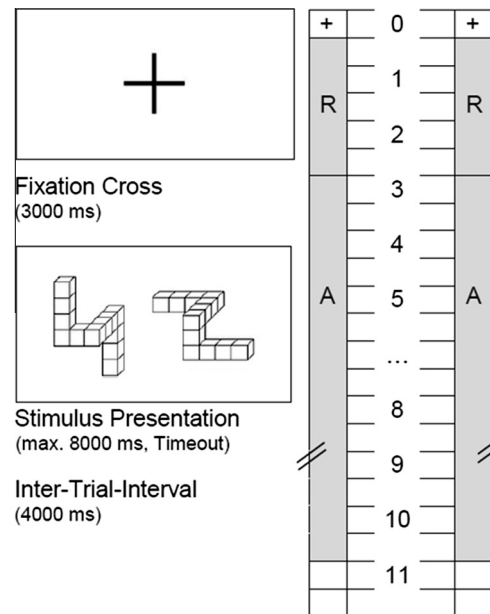

Fig. 1. Schematic time course and EEG measurement intervals for the mental rotation task. After a reference period (R, 3 s), participants worked on the task item (activation period, A; timeout after 8 s) until they gave their response by key press.

3. Results

3.1. Behavioral results

In order to examine possible group differences between girls and boys and between the stereotype exposure groups with respect to task performance, a two-way ANOVA with SEX and STEREOTYPE EXPOSURE as between-subjects variables was computed. The average response time (for correct trials) was 4.02 s (SD = 0.78). There were neither significant group mean differences for SEX ($F(1,54) = 1.20$, $p = .28$), nor for the STEREOTYPE EXPOSURE condition ($F(1,54) = .05$, $p = .82$); the two-way interaction was also not significant (SEX * STEREOTYPE EXPOSURE: $F(1,54) = .01$, $p = .95$; no-stereotype exposure condition: $M_{girls} = 4.04$, $SD_{girls} = 0.91$; $M_{boys} = 4.04$, $SD_{boys} = 0.84$; stereotype exposure condition: $M_{girls} = 3.86$, $SD_{girls} = 0.79$; $M_{boys} = 4.11$, $SD_{boys} = 0.63$).

For the analysis of solution rates similar results were found. There were neither significant group mean differences for SEX ($F(1,54) = 2.94$, $p = .09$, partial $\eta^2 = .05$), nor for the STEREOTYPE EXPOSURE condition ($F(1,54) = 0.15$, $p = .70$, partial $\eta^2 = .00$). Contrary to our hypothesis, the interaction of SEX * STEREOTYPE EXPOSURE remained insignificant ($F(1,54) = 2.43$, $p = .12$, partial $\eta^2 = .04$; no-stereotype exposure condition: $M_{girls} = 36.92$, $SD_{girls} = 5.55$; $M_{boys} = 37.12$, $SD_{boys} = 5.43$; stereotype exposure condition: $M_{girls} = 34.46$, $SD_{girls} = 4.68$; $M_{boys} = 38.60$, $SD_{boys} = 4.36$; see Fig. 2).

3.2. EEG results

In a first step, we analyzed the effect of stereotype exposure and sex on task-related power (TRP) changes in the upper alpha band. This was done by means of a four-way ANOVA, where STEREOTYPE EXPOSURE and SEX were treated as between-subjects factors, and HEMISPHERE and AREA were considered as within-subjects factors. A main effect STEREOTYPE EXPOSURE ($F(1,54) = 3.93$, $p = .05$, partial $\eta^2 = .07$) indicated that participants working in the stereotype exposure condition show higher cortical activation ($M = 0.07$, $SD = 0.03$) than participants working in the
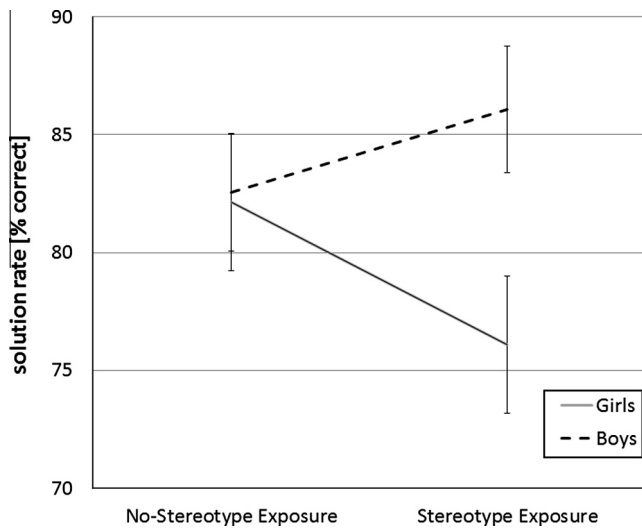
Fig. 2. Performance (solution rate) in the mental rotation task for experimental conditions either involving stereotype exposure or not. Error bars indicate ± 1 *SE* of the mean.

no-stereotype exposure condition ($M = -0.03$, SD = 0.03). No further TRP effects reached statistical significance.

We then analyzed the effect of stereotype exposure and sex on neural efficiency. In line with previous studies (Neubauer et al., 2005), the correlation between figural intelligence and brain activation (TRP) during performance of the mental rotation task was used as an inverse indicator of neural efficiency (i.e., a negative correlation would support the neural efficiency hypothesis). Correlations were computed separately for each experimental condition (factors STEREOTYPE EXPOSURE and SEX; i.e., girls and boys working under stereotype exposure or no-stereotype exposure condition, respectively) and each topographic area of both hemispheres (factors AREA and HEMISPHERE). The TRP was normally distributed in each topographic area for all groups.

As depicted in Fig. 3, the IQ-brain activation relationship differs considerably depending on sex and stereotype exposure condition. In the no-stereotype exposure condition, boys showed the expected negative IQ-brain activation relationships especially at centroparietal ($r = -.45$, $p = .05$) and temporal areas ($r = -.50$, $p = .04$) of the left hemisphere. Girls working under the no-stereotype exposure condition rather tended to show a positive IQ-brain activation relationship especially at frontal areas ($r = .48$, $p = .10$) in

the right hemisphere of the brain. In the stereotype exposure condition, no significant IQ-brain activation correlations were found, neither for boys nor girls. To sum up, in the no-stereotype exposure condition the neural efficiency hypothesis is supported only for boys, but not for girls. In the stereotype exposure condition no support for the neural efficiency hypothesis was obtained, neither for girls nor boys.

4. Discussion

This study aimed at further examining sex differences regarding the phenomenon of neural efficiency. Our hypothesis was based on research demonstrating that neural efficiency is primarily dominant in those sex groups that are working on tasks in which they commonly perform better (Neubauer & Fink, 2009). More specifically, it was tested whether these findings might be attributed to the social-psychological phenomenon of stereotype threat, as specific gender-stereotypes can affect task performance as well as brain activation (e.g., Wraga et al., 2007). The behavioral results of this EEG study are not in conformity with previous findings demonstrating that stigmatized groups underperform when the negative stereotype about their group seems relevant and when the situation strikes one as a test of stereotype-relevant qualities (e.g., Good et al., 2008; Spencer et al., 1999). Under stereotype exposure girls showed no significant decrease in mental rotation performance. Evidence exists, that participants do not necessarily perform poorly although confronted with a negative stereotype that increases the experience of stress, heightened vigilance and emotional suppression (Davies, Spencer, & Steele, 2005; Schmader et al., 2008). Under stereotype exposure there was an increase of cortical arousal which indicates that girls working under stereotype exposure have an increased stress arousal.

The main aim of this study was to examine whether sex differences in neural efficiency can be attributed to stereotype threat effects. When the mental rotation task was described as a task to produce sex differences (i.e., in the stereotype exposure condition), girls and boys did not show any negative IQ-brain activation relationship. When the task was described as being unaffected by sex (i.e., in the no stereotype exposure condition) the hypothesized neural efficiency findings occurred only for boys. The later condition represents a replication of findings reported previously by Neubauer et al. (2002, 2005). It hence could be concluded that those findings were not due to stereotype threat. In contrast, eliciting a stereotype threat seems to disrupt the neural efficiency phenomenon, likewise in boys and girls. This finding was somewhat surprising as we had originally hypothesized that sex differences
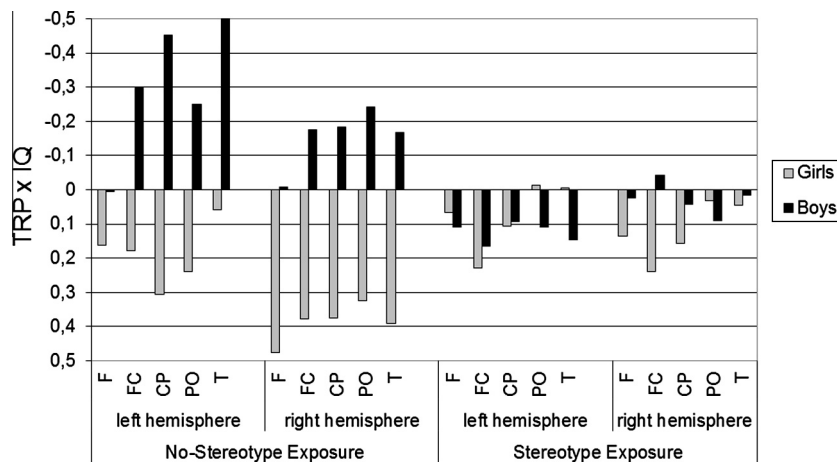

Fig. 3. Correlations of TRP and figural IQ by experimental conditions, sex and topographic region.

in neural efficiency might only occur in the stereotype threat condition.

Girls and boys working in the no-stereotype exposure condition showed equal task performance but nevertheless differed in the correlation between brain activation and intelligence. Only for boys the neural efficiency phenomenon was supported especially at parietal and temporal cortices. These areas, together with frontal brain areas, are assumed to constitute an important network involved in complex information processing (cf. the parieto-frontal integration theory by Jung and Haier (2007)). The finding that sex differences in brain activation do not concur with behavioral results has been reported frequently (e.g., Kober & Neuper, 2011). One reason for this incongruence between behavioral and neurophysiological results might be that sex differences in the cortical activation pattern can be attributed to fixed differences in the cerebral organization in men and women. For instance, Nopoulos, Flaum, O'Leary, and Andreasen (2000) reported that females have a smaller overall brain size than males. Haier, Jung, Yeo, Head, and Alkire (2005) found that men have more gray matter (neurons, synapses, dendrites) in fronto-parietal brain regions whereas women have more white matter (myelinated axons). Moreover, in males, intelligence is correlated more with gray matter areas whereas in females white matter areas are correlated higher with intelligence (for a review cf. Deary, Penke, & Johnson, 2010).

Remarkably, during explicit stereotype exposure the neural efficiency phenomenon could no longer be observed, neither for boys nor girls. In this condition boys received the message that they usually perform better than girls. Boys might have reframed this stereotype as a challenge. Considering a test situation as a challenge is known to lead to increased performance (Alter, Aronson, Darley, Rodriguez, & Ruble, 2010; Keller, 2007). The arousal associated with this challenge could also result in increased brain activation, especially in high IQ boys who typically show lower brain activation (Neubauer & Fink, 2009). This might explain why no neural efficiency was observed in this specific task condition.

In a similar vein, the reported brain activation pattern found for girls in the stereotype exposure condition might also be the consequence of the increased performance pressure. However, in contrast to boys the stereotypic expectancies for girls result in a threat experience, because of the possibility to confirm the stereotype. This argument appears to be supported by the finding that the stereotype exposure condition was associated with higher arousal in terms of higher TRP. Moreover, the selective increases in brain activation due to increased arousal could again have counteracted the general phenomenon of neural efficiency.

4.1. Conclusion

Our results provide preliminary evidence that the stereotype threat itself cannot explain sex differences in neural efficiency in visuo-spatial tasks. Results corroborate the neural efficiency hypothesis for men only when sex differences were described to be irrelevant. This suggests that visuo-spatial sex differences in brain activation patterns may be caused by biological but also by long term social factors like learned or socially determined interests and not only short-lived stressing effects of stereotype threat on performance. It still has to be acknowledged that activated stereotypes significantly affected brain activation, but they are probably not responsible for the reported sex differences in neural efficiency during visuo-spatial tasks. Therefore, it is still important to consider the phenomenon of stereotype threat in forthcoming studies. A replication of the present findings including a verbal task could be of particular interest for future investigations, as this would represent a stereotype threat for boys and a stereotype lift for girls.

References

Alter, A. L., Aronson, J., Darley, J. M., Rodriguez, C., & Ruble, D. N. (2010). Rising to the threat: Reducing stereotype threat by reframing the threat as a challenge. *Journal of Experimental Social Psychology, 46*, 166–171.

Croizet, J. C., Després, 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.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12*, 671–684.

Davies, P. G., Spencer, S. J., & Steele, C. M. (2005). Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of Personality and Social Psychology, 88*(2), 276–287.

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience, 11*, 201–211.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*, 336–353.

Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology, 29*, 17–28.

Grabner, R. H., Fink, A., Stipacek, A., Neuper, C., & Neubauer, A. C. (2004). Intelligence and working memory systems: Evidence of neural efficiency in alpha band ERD. *Cognitive Brain Research, 20*, 212–225.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55*, 468–484.

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., & Alkire, M. T. (2005). The neuroanatomy of general intelligence: Sex matters. *Neuroimage, 25*, 320–327.

Haier, R. J., Siegel, B. V., MacLachlan, A., Soderling, E., Lottenberg, S., & Buchsbaum, M. S. (1992). Regional glucose metabolic changes after learning a complex visuospatial/motor task: A positron emission tomographic study. *Brain Research, 570*, 134–143.

Halpern, D. F. (2004). A cognitive taxonomy for sex differences in cognitive abilities. *Current Directions in Psychological Science, 13*, 135–139.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science, 8*, 1–51.

Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology, 93*, 544–564.

Jaušovec, N., & Jaušovec, K. (2008). Spatial rotation and recognizing emotions: Gender related differences in brain activity. *Intelligence, 36*, 383–393.

Jung, R. E., & Haier, R. J. (2007). The parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Science, 30*, 135–187.

Keller, J. (2007). When negative stereotypic expectancies turn into challenge or threat: The moderating role of regulatory focus. *Swiss Journal of Psychology, 66*, 163–168.

Kober, S. E., & Neuper, C. (2011). Sex differences in human EEG theta oscillations during spatial navigation in virtual reality. *International Journal of Psychophysiology, 79*, 347–355.

Krendl, A. C., Richeson, J. A., Kelley, W. M., & Heatherton, T. F. (2008). The negative consequences of threat – A functional magnetic resonance imaging investigation of the neural mechanisms underlying women's underperformance in math. *Psychological Science, 19*, 168–175.

Lipp, I., Benedek, M., Fink, A., Koschutnig, K., Reishofer, G., Bergner, S., et al. (2012). Investigating neural efficiency in the visuo-spatial domain: An fMRI study. *PLoS One, 7*(12), e51316.

Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences, 49*, 344–351.

Moè, A., & Pazzaglia, F. (2006). Following the instructions! Effects of gender beliefs in mental rotation. *Learning and Individual Differences, 16*, 369–377.

Neubauer, A. C., Bergner, S., & Schatz, M. (2010). Two- vs. three-dimensional presentation of mental rotation tasks: Sex differences and effects of training on performance and brain activation. *Intelligence, 38*, 529–539.

Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience & Biobehavioral Reviews, 33*, 1004–1023.

Neubauer, A. C., Fink, A., & Schrausser, D. G. (2002). Intelligence and neural efficiency: The influence of task content and sex on the brain–IQ relationship. *Intelligence, 30*, 515–536.

Neubauer, A. C., Grabner, R. H., Fink, A., & Neuper, C. (2005). Intelligence and neural efficiency: Further evidence of the influence of task content and sex on the brain–IQ relationship. *Cognitive Brain Research, 25*, 217–225.

Nopoulos, P., Flaum, M., O'Leary, D., & Andreasen, N. C. (2000). Sexual dimorphism in the human brain: Evaluation of tissue volume, tissue composition and surface anatomy using magnetic resonance imaging. *Psychiatry Research, 98*, 1–13.

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.

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronisation and desynchronisation: Basic principles. *Clinical Neurophysiology, 110*, 1842–1857.

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

Schmader, T., Johns, M., & Forbes, C. E. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115*, 336–356.

Spencer, S., Steele, C., & Quinn, D. (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*, 613–629.

Steele, C. M. (2010). *Whistling Vivaldi and other clues to how stereotypes affect us.* New York: W.W. Norton & Company.

Stone, J. (2002). Battling Doubt by avoiding practice: The effects of stereotype threat on self-handicapping in white athletes. *Personality and Social Psychology Bulletin, 28*, 1667–1678.

Wicherts, J. M., Dolan, C. V., & Hessen, D. J. (2005). Stereotype threat and group differences in test performance: A question of measurement invariance. *Journal of Personality and Social Psychology, 89*, 696–716.

Wraga, M., Helt, M., Jacobs, E., & Sullivan, K. (2007). Neural basis of stereotype-induced shifts in women's mental rotation performance. *Social Cognitive & Affective Neuroscience, 2*, 12–19.