Running Head: Task difficulty and stereotype threat in a spatial orientation task

Interaction of task difficulty and gender stereotype threat with a spatial orientation task in a virtual nested environment

Craig Allison

Edward S. Redhead

Wai Chan

University of Southampton

Address for correspondence

Dr Ed Redhead

School of Psychology

University of Southampton

Highfield

Southampton SO17 1BJ

Email: er2@soton.ac.uk

Tel: 02380592307

Fax: 02380584597

Abstract

Two experiments examined the interaction of task difficulty and stereotype threat in a spatial orientation task. Having explored the exterior and interior of a virtual building, participants were placed in a room with an external or internal view and asked to face a previously seen but occluded external target cue. In the internal room participants could use spatial updating to track their position in terms of the target cue, and in the external room they could also use the allocentric spatial relationship between the target cue and a visible external cue. Participants performed better in the external room, illustrating spatial updating is more difficult than allocentric array learning. In Experiment 1, participants were informed that they were likely to perform better, worse or the same as members of the opposite sex. Overall males performed better than females, but males given the threat statement performed worst. There was no difference between female groups. Experiment 2a, reduced the difficulty of the task by including internal orienting cues. Females with the orientation cues performed better than females without orientation cues and the same as males. In Experiment 2b, with orientation cues present, there was a significant effect of stereotype threat for both males and females but only in the more difficult internal room trial. The results suggest gender stereotype threats affect spatial orientation but only at an appropriate level of task difficulty.

Key words

Gender Stereotypic threat; Task difficulty; spatial updating; allocentric array learning; computer generated environment; gender differences; orientation cues

Stereotype threat comprises two elements; an opinion widely held by society that a certain group will perform poorly on a particular task, and the belief of the group in question that they will conform to this opinion (Steele, 1999). If people are presented with a stereotype suggesting that they are deficient in a particular domain then their performance will suffer (Aronson et al., 1999). Research into stereotype threat typically involves tasks with well established gender differences. Previous examples include studies of mathematical ability (Shih, Pittinsky, & Ambady, 1999) and tests of political knowledge (McGlone, Aronson, & Kobrynowicz, 2006). The particular interest in maths ability is that stereotype threat may play a role in dissuading females from entering STEM based courses at University and thus lead to the gender difference in STEM based careers (Schmader, Johns, & Barquissau, 2004).

A third area often investigated is spatial ability, usually measured by the Mental Rotation Task (MRT) (Vandenberg & Kuse, 1978; McGlone & Aronson, 2006; Heil, Jansen, Quaiser-Pohl, & Neuberger, 2012). Many studies have shown a robust gender difference in spatial ability (e.g. Lawton, Charleston, & Zieles, 1996), particularly when demonstrated with the MRT. With such a common perception that males are better than females at spatial tasks, it would seem that stereotype threat would emerge with spatial ability. Moè and Pazzaglia (2006) explored stereotype threat using the MRT. The researchers attempted to manipulate performance by offering opposing stereotypes to different groups. Participants were either informed that males should surpass females on the MRT or that females would surpass males. All participants that were told that they were expected to perform worse than the participants of the opposite sex did so. Heil et al. (2012) provided support for this finding in females but found that the stereotype manipulations only affected the males in their guessing behavior. The authors concluded that stereotype threat affects men and women differently. Stereotype threat in spatial ability has also been found to affect older populations (Meneghetti, Muffato, Suitner, De Beni, & Borella, 2015), which may exacerbate any true decline in spatial ability with age.

In an expansion of the work by Moè and Pazzaglia (2006), the stereotype threat manipulation has been shown to be effective on visuospatial tasks other than mental rotation (e.g. Campbell & Collaer, 2009; Rosenthal, Norman, Smith, & McGregor, 2012). Rosenthal, et al. (2012) combined stereotype threat with an active navigation task within a virtual environment. They required participants to search for a hidden platform within a digital watermaze after making them aware of a stereotype threat manipulation. Although females' performance did not deteriorate as a result of the stereotype threat, Rosenthal et al. found that males' performance improved, relative to controls, within the stereotype condition, demonstrating a stereotype lift. Stereotype lift (Walton & Cohen, 2004), facilitating, rather than impeding, performance, has also been identified within mental rotation tasks (Hausmann, Schoofs, Rosenthal, & Jordan, 2009). Hausmann et al., found that making males aware of the gender stereotype improved performance. Female performance was not impacted by a corresponding stereotype threat.

Not all studies have found an effect of stereotype threat and lifts with spatial tasks, indeed a meta-analysis by Doyle and Voyer (2016) found that although there was a small but significant effect of threat on performance in maths problems, in the studies analysed the effect of stereotype manipulation on spatial tasks was not significant. Flore and Wicherts (2015) have also suggested that a publication bias for reporting positive results may have led to the effect of stereotype threat being overstated. A possible reason for the variance in finding an effect of stereotype threat manipulation might be the number of proposed mediators, for example, the salience of the stereotype (Nguyen & Ryan, 2008), the gender of the person delivering the stereotype statement (Nichols & Maner, 2008) and the gender split within the group (Inzlitch & Ben-Zeev, 2000)

One mediator put forward for the stereotype effect is task difficulty. In their studies of stereotype threat, Quinn and Spencer (2001) and Steele (1997), discuss the importance of pitching the task at the right level; it needs to be difficult enough so people feel they have reached their limit of ability, but not so difficult that everybody finds it impossible and stereotypes become irrelevant. Keller (2007) demonstrated that the stereotype threat only decreased performance on difficult maths problems. The current study will systematically change the difficulty of a spatial task to examine the interaction between task difficulty and stereotype threat and lift.

The current study used a more naturalistic spatial orientation task than the typical mental rotation task. The current study required participants to maintain orientation within the wider environment while moving through a nested environment. Here we define nested as an internal environment (a building) contained within a larger environment (a university campus). Carlson, Holscher, Shipley, and Conroy (2010) have examined the factors which are important while navigating through a building. They have highlighted the difficulty when participants create a cognitive map based on incomplete or inconsistent information. Hegarty, Montello, Richardson, Ishikawa and Lovelace (2006) have also examined the association between psychometric tests of spatial ability such as Mental rotation tasks and an environmental spatial task such as wayfinding or learning the layout of a large building. In terms of gender differences for large scale environments, Lawton, Charleston, and Zieles (1996) examined navigational ability in an unfamiliar indoor nested environment. It was found that males were significantly better than females at indicating their direction of origin. Lawton and Morrin (1999) similarly found males performed better than females on a pointing task within a digitally created maze. When men and women were led through several turns of a maze and asked to point at the occluded starting position, pointing accuracy decreased with increasing maze complexity, but across all conditions males outperformed females.

As we move through an environment, our egocentric relationship with the environment constantly changes. If individuals experience or perceive motion, they are required to reorient themselves via the use of constant static cues relative to their previous location. This process of reorientation is referred to as spatial updating (Reiser, 1989), and it is possible to maintain orientation with cues at and beyond your local environment. For example, if an individual is sitting facing a computer and turns to the right to answer the door, the computer is now on the individual's left. Equally when walking through a building, it is possible to track your position with reference to the street from which you entered the building. Evidence suggests that spatial updating is an automatic cognitive process (Reiser, 1989; Farrell & Robertson, 1998). In terms of spatial updating within a computer generated environment such as that used in the current experiment, Riecke, Cunningham and Bulthoff (2007) found that participants provided with the realistic tour of the environment were able to engage in automatic spatial updating, regardless of whether they had access to physical motion cues, suggesting that visual cues were sufficient for spatial updating.

Wang and Brockmole (2003) investigated spatial updating within nested environments, specifically, a room within a university campus. After learning the locations of key targets within both the room and the wider campus, blindfolded participants were required to track key objects within either the room or the campus. Participants were then required to point to non-visible objects outside of the room within the larger environment. They found that although participants could maintain orientation within the local environment of the room, they struggled to accurately track and update locations within the external setting of the wider campus.

In the current experiment, different groups of participants were told that males are better than females at spatial tasks, that females are better than males, or that there is no difference between males and females. They were allowed to explore a digital model of the

University of Southampton's Psychology building and then were asked to point to an occluded object external to the building from each of two rooms, an internal room or an external room with a partial view of the surrounding campus.

The results of Wang and Brockmole (2003) would suggest that maintaining orientation to the occluded external target cue via spatial updating within an internal room would be a difficult task, as the spatial relationship between the external target cue and the participants' position is constantly changing. In a room with a partial external view, however, participants do not have to rely solely on spatial updating. They have a view of an alternative external cue that has a constant spatial relationship with the external target cue. If the participants have learned the spatial relationship between the external cues via an allocentric array, then they should find it easier to point in the direction of the occluded target cue from the external room than from the internal room.

From previous research, we make several predictions regarding Experiment 1. Firstly, given the reliable gender difference found between males and females in pointing at occluded positions in both real world situations (Lawton et al., 1996) and digitally created mazes (Lawton & Morrin, 1999), it is expected that males will outperform females at the orientation task. Secondly, given the difficulty of spatial updating in respect to objects within an external array (Wang & Brockmole, 2003) it is predicted that participants will perform poorly in the internal room. Thirdly, if we replicate the findings of Moè, we should find a depression of performance in the stereotype threat condition and a facilitation of performance in the lift condition. Finally, if task difficulty and stereotype threat interact in the same manner as shown in maths problems (Keller, 2007), we should find the effect only in the more difficult spatial updating task in the internal room.

Experiment 1

Method

Design.

This study used a 2 (Room type) x 2 (Gender) x 3 (Condition) mixed design. Participants were divided by gender (Male and Female) and systematically placed in one of the three conditions, Stereotype Threat, Stereotype Lift or Control. The dependent variable was orientation error as measured by pointing accuracy in the Internal and External rooms.

Participants

Participants were 72 Psychology undergraduate students (37 males and 35 females) attending the University of Southampton who were allocated course credits for participation in the 20-minute session. They were between the ages of 18 and 25 (*M* = 21.62*, SD* = 1.51). All participants were presumed to be of a similar cognitive ability given their similar education history, and all participants' first language was English. Participants were randomly assigned to the three conditions, Stereotype Threat, Stereotype Lift or Control (Stereotype Threat: Males = 12, Females = 12; Stereotype Lift: Males = 12, Females = 11; Control: Males = 13, Females = 12) .

All participants had normal or corrected normal vision. As part of their studies, all participants had experience of the real campus and Psychology building, but not of the computer model.

Materials and Apparatus

The study took place within a windowless research cubicle, measuring 2.4 m in length by 1.3 me wide, with a height of 2 m. The cubical contained a single desktop computer. The computer used a standard Windows 7 operating system, with keyboard and mouse, placed on a 1.3-m wide desk in the center of the rear wall. The computer was connected to three identical 15 inch LCD monitors. The monitors were placed horizontally so that the displayed image was shown continuously across all three screens.

A virtual model of the Psychology Building, University of Southampton, and its surroundings was created for this study. This model was developed by Dr Matt Jones, University of Southampton, using 3DSMax 2012. The programme placed the participants within the environment and offered a first person perspective (Figures 1a and 1b). The model enabled participants to freely explore both the surrounding area and the third floor of the Psychology building. The remaining areas of the campus, including other buildings and additional floors within the Psychology building, were not accessible. Participants controlled their movement using the “FORWARD” “LEFT” and “RIGHT” arrow keys, but could not look up or down, or interact with items within the environment. Participants were unable to use the “BACK” arrow key in order to better simulate real life movement.

Procedure

Before starting the spatial task, participants were given one of three stereotype priming statements, (Table 1). The spatial task was comprised of two set of trials, acquisition and orientation test trials. Participants completed two acquisition trials to familiarize themselves with the environmental layout. For the first acquisition trial, participants were required to explore the outside campus environment, allowing the participant to see, but not enter, surrounding buildings, including the physics building and the student union. Participants were provided written instructions regarding a route to take within the environment. At each junction, participants were provided with instructions as to which direction to take. This route ensured that participants saw and identified all visible landmarks in the surrounding area and did not become trapped or lost. As movement was controlled by the participants, they were free to divert from the suggested route if they wished to explore further. Once participants had explored the surrounding environment, they were instructed to make their way back to the starting location and to indicate to the experimenter that they were happy to proceed to the second acquisition trial. For the second acquisition trial, participants were required to enter the virtual building and to explore the third floor of the model. Participants were again provided with a route to follow in order to ensure that they had visited key locations within the building and looked out of key windows; participants were free to explore the environment for as long as they wished and no time limit was placed on these trials.

Participants were given two Orientation test trials. Participants were digitally placed in an internal or an external facing room (order of room presentation was counterbalanced) within the Psychology building and asked to turn to face a non-visible, but previously seen, external target cue. The target cue was one of four areas, the Car Park (south of the Psychology building) the Physics Building (west of the building), the Main Campus (North of the building), or the Union Building (east of the building) (See Figure 1c for layout of University Environment). Participants were not able to leave the room they had been placed in but were free to move within the room. The internal room overlooked the inner courtyard of the building. The external room overlooked one of the external cues except the target cue. The cues provided were sufficient to complete the orientation task.

Results and Discussion

Training Trials

Time spent exploring during the initial two exploratory trials revealed little difference between the stereotype conditions (External exploratory trial Threat: M=82.68 s, SE = 15.98; Lift M=83.35 s SE=15.23; Control M=61.54 s, SE=17.92; Internal exploratory trial Threat: M=182.63 s, SE = 22.98; Lift: M= 194.85 s, SE=21.89; Control: M=151.54 s, SE=25.77).

Overall males spent less time on these trials than females (Males M=89.83 s, SE=14.60; Females M=162.37 s, SE=14.72). These conclusions were confirmed by a 2 (trial type) x 2 (gender) x 3 (condition) mixed design Analysis of Variance (ANOVA) with trial duration the dependent variable. The main effect of trial type was significant $F(1, 52) = 83.39$, $p<0.000$, $\eta_p^2 = .61$, with the internal exploration trial lasting longer than the external exploration trial. The main effect of gender was significant, $F(1, 52) = 12.23$, $p=0.001$, $\eta_p^2 = .19$, with males taking less time to complete the exploratory trials than females. The main effect of stereotype condition was not significant, $F<1$, nor were any of the interactions between the factors (trial x gender: $F(1, 52) = 2.87$, $p=0.096$, $\eta_p^2 = .05$; trial x condition: $F<1$; trial x gender x condition: $F<1$; gender x condition: $F(1, 52) = 1.38$, $p=0.261$, $\eta_p^2 = .05$)

Orientation trials

Figure 2 illustrates the orientation error among the groups during the orientation trials. Overall males performed better than females except in the threat condition. For both males and females, performance in the external room was better than in the internal room. The effect of stereotype threat was more pronounced in the male groups; the males in the threat condition appeared to perform worse than those in the lift and control condition. This finding was true for both the internal and external room. These observations were confirmed via a 2 (room type) x 2 (gender) x 3 (condition) ANOVA with orientation error the dependent variable. The main effect of room type was significant, $F\ (1, 66) = 17.29$, $p < 0.000$, $\eta_p^2 = .21$, with participants recording less orientation errors in the external room. There were no significant interactions between room type and gender or room type and condition, nor was there a significant three way interaction, $Fs< 1$.

Both the main effect of gender, $F(1, 66) = 7.14$, $p = 0.009$, $\eta_p^2 = .098$, and the main effect of condition were significant, $F(2, 66)= 5.26$, $p = 0.008$, $\eta_p^2 = .138$, and the interaction was also significant, $F(2, 66) = 5.05$, $p = 0.009$, $\eta_p^2 = .133$. Further analysis of the simple main effects (Kepple, 1973) revealed that the effect of condition was only significant in the male participants, $F(2, 66) =10.45$, $p < 0.000$, $\eta_p^2 = .28$. A post hoc independent t-test using the Bonferroni correction revealed males in the threat group recorded more orientation errors than those in the lift and control groups. Two further post hoc independent t-tests using the Bonferroni correction further revealed that performance by females in the Lift condition did not significantly differ from the performance of males in the control condition in either the internal or external rooms.

In Experiment 1, overall males performed significantly better than females on the spatial task, which agrees with the reliable gender difference seen in other spatial tasks. Females in the control and lift conditions performed poorly compared to the males in the equivalent conditions. Particularly the results of the control group suggest females are worse at this task than males, which is typical of many spatial tasks which show gender differences (e.g., Lawton & Kallai, 2002). There was also a significant effect of room type, with all participants performing better in the External rather than the Internal orientation test trial. This result supports the findings of Wang and Brockmole (2003) and suggests that spatial updating with respect to an external target cue was more difficult than learning to orient using the spatial relationship between the external cues and the allocentric array.

In Experiment 1, there was also a significant effect of stereotype priming for males but not for females. This result replicates the findings of Rosenthal et al. (2012) and Haussmann et al. (2009). The failure to find a priming effect in females could be due to two factors. Firstly, the priming had no effect on the female participants because the statement that “males are better than females at spatial tasks” is a commonly held belief. The result would be that neither the threat nor lift priming statement would have been sufficient to have an impact on female participants’ beliefs and thus their performance. Secondly, females may be unable to maintain orientation, particularly in the internal room, leading to a floor effect and making the stereotype manipulation irrelevant (Spencer, 1999). Females' performance in the internal room was at chance and so could not be influenced by the stereotype threat. By chance, it would be expected that participants would get a score of 90 degrees orientation error (average in a possible orientation error range between 0-180). A further single sample t-test showed that females' performance across the three groups in the internal room was no better than a chance score of 90 degrees $t(34) = 0.04$, $p=.96$. A similar t-test for male participants showed that performance across the three groups was better than chance, $t(37) = 2.20$, $p=.0.037$.

Given Doyle and Voyer's (2016) meta-analysis showing that stereotype threat and lift is not significant in the majority of spatial tasks, it is important to explore these two possible explanations for the lack of an effect of stereotype manipulation in female participants in Experiment 1. In order to assess whether an easier task would be more vulnerable to the stereotype manipulation Experiment 2a provided additional orienting cues and assessed whether this improved female performance in the internal room. Experiment 2b was then performed to assess whether the easier task is vulnerable to the stereotype manipulation used in Experiment 1.

Experiment 2a

From previous research (Wang & Brockmole, 2003), it is apparent that individuals struggle to maintain orientation within nested environments due to limited visual cues linking their environments. By providing participants with clear visual links between the internal and external environments they do not have to rely solely on spatial updating and it was anticipated that orientation error would be reduced.

Experiment 2a investigates whether the inclusion of additional visual cues within a nested virtual environment will assist participants in their ability to successfully orient. Using the same computer generated model as in Experiment 1, one group of participants was provided with additional orienting cues within the environment. Participants were divided into two groups, Group No Cue and Group Cue. Although both groups explored the same environment, color bands, were added to the top of each of all internal and external walls for Group Cue (see Figures 3a and b) to aid orientation. North, South, East and West facing internal walls had uniquely colored bands running along their lengths. The color matched the color of the band on the external walls adjacent to the four external target cues. For example, all south facing internal walls had green bands that matched the green band on the external wall adjacent to the Car Park Target Cue that was south of the building. If the internal orienting cues do aid orientation, then the female participants who are exposed to these cues should perform as well as the males in the internal rooms.

Method

Design

This study used a 2 (Room type) x 2 (Gender) x 2 (Condition) between participant factorial design. The condition was whether participants received Orientation cues or No Orientation cues. The dependent variable was orientation error, as measured by pointing accuracy. Participants were randomly allocated between the two conditions.

Participants

Participants were 90 psychology undergraduate students (31 males, 59 females), from the University of Southampton, who completed the study in partial fulfilment of a research participation scheme. They were between the ages of 18 and 32 years (*M* = 21.91 years, *SD* =

1.32). All participants were presumed to be of a similar cognitive ability given their similar education history, and all participants' first language was English. Participants were recruited via the use of a departmental recruiting system and randomly allocated to the two conditions Orientation cues or No Orientation cues (Orientation cues: Males = 15, Females = 30; No Orientation cues: Males = 16, Females = 29).

All participants had normal or corrected normal vision. Participants had experience of the real campus and Psychology building, but not of the computer model. Participants who took part in Experiment 1 were not eligible to participate within this study.

Materials and Apparatus

The apparatus used in this study was identical to those used in Experiment 1. The study took place in the same windowless research cubicle.

Procedure

Conditions varied with the number of additional orientation cues provided within the environment. Within the No Cue condition, participants were given no additional orientation aids to assist navigation and orientation. Group No Cue explored the same virtual model as participants in Experiment 1 (Figures 1a and b). The external orienting cues present for Group Cue were colored bands 3 m in height (relative to the building) and running the length of the external walls. The cue on the wall adjacent to the southerly Car Park Target Cue was Green, the cue for the Main Campus was Blue, for the Physics Building it was Red and for the Union Building it was Yellow. The Internal Orienting cues were similarly colored bands running the length of each internal wall, 0.5 m in height (relative to the wall) and were one of four colors based on the wall's cardinal direction (Figures 3a and b); cues on southerly facing walls were green, cues on northerly facing walls were blue, cues on westerly facing walls were red, and cues on easterly facing walls were yellow.

As in Experiment 1, participants were allowed to explore both the exterior and interior of the building. Orientation trials were also the same as in Experiment 1. On separate test trials, participants were placed within an internal facing room and an external facing room. On both occasions, they were asked to turn to face a target which was not visible from the room but that they would have seen during exploration. Orientation error was recorded.

Results and Discussion

Figure 4 illustrates the orientation error among the four groups. Whereas there was little difference between the two male groups in either the internal or external rooms, the females that received the additional orienting cues in the internal room performed better than those without the cues and similar to the male participants. These observations were confirmed via a 2 (room type) x 2 (gender) x 2 (condition) mixed design ANOVA with orientation error the dependent variable. The main effect of room type was significant, $F(1, 86) = 8.93$, $p = 0.004$, $\eta_p^2 = .09$, because participants had lower orientation errors in the external room than the internal room. There was no significant interaction between room type and gender, $F < 1$, or between room type and condition, $F(1, 86) = 3.66$, $p = 0.059$, $\eta_p^2 = .04$. There was not a significant three way interaction, $F(1, 86) = 2.09$, $p = 0.152$, $\eta_p^2 = .02$.

The main effect of gender was significant, $F(1, 86) = 5.31$, $p = 0.024$, $\eta_p^2 = .058$, with males performing better overall than females. The main effect of condition was significant, $F(1, 86) = 6.40$, $p = 0.013$, $\eta_p^2 = .07$, with participants given the additional orienting cues recording lower orientation error than those without. The interaction was not significant, $F(1, 86) = 2.46$, $p = 0.120$, $\eta_p^2 = .03$. A post hoc independent t-test using the Bonferroni correction on the female orientation errors revealed that the effect of condition was significant, with females who were given the additional orienting cues performing better than females without the cues.

Experiment 2a once again demonstrated that participants performed better in the External orientation test trials than the Internal orientation trials. This difference was particularly the case for females without the additional orientation cues. However, with additional orientation cues, females could orientate in the internal room to the same level as they did in the external room and to the same level as males.

Experiment 2b tested the possible reasons for the lack of a stereotype effect in females seen in Experiment 1. The absence might have been due to the task in Experiment 1 being too difficult, as the results of Experiment 1 demonstrated, in the absence of internal orientation cues, females performed at chance. Alternatively, the lack of effect of the priming statements on female participants may have been due to the common assumption that males outperform females in spatial tasks and so the control and lift statements were not believed and so had no effect.

Experiment 2b

Experiment 2b provided similar stereotypic control, threat and lift statements to those used in Experiment 1. However, Experiment 2b varied from Experiment 1 in two ways. Firstly, the environment in Experiment 2b included the extra orienting cues used in Experiment 2a in order to assess whether the reason no stereotype threat was seen in females was due to the difficulty of the task or the general acceptance of the notion that males outperform females in spatial task. Secondly, the environment in Experiment 2b was presented in a passive manner similar to that used by Hegarty et al. (2006). Participants were shown a video of the route taken around the building rather than being allowed to explore the building. Hegarty et al. (2006) found that learning environments via video or desk top Virtual Environments leads to a high level of performance in pointing accuracy tests. Presentation of the environment via video will allow us to assess the generality of the benefit of an external viewpoint and of the stereotype effect.

Method

Design

A 2 (room type) x 2 (gender) x 3 (condition) design was applied in which stereotype conditions varied on 3 levels (threat, lift, and control) between the participants. The dependent variable was orientation error.

Participants

Participants were 160 undergraduate students (37 males and 123 females) attending the University of Southampton who were allocated course credits for participation in the 20-min session. They were between the ages of 18 and 26 (*M* = 20.42, *SD* = 1.13). The experiment was run over six sessions, with no more than 30 students attending each session. In sessions 1 and 4, participants were told that females perform better than males in certain spatial tasks, meaning that males attending these sessions would be in the threat condition and females in the lift condition. In sessions 2 and 5, participants were told that males perform better than females in certain spatial tasks so that females attending these sessions would be in the threat condition and males in the lift condition. In sessions 3 and 6, participants were told that there is no difference between males and females in certain spatial tasks so that all students attending these sessions were in the control condition. The numbers of participants in the three conditions were as follows: Threat: Male = 12, Female = 41; Lift: Male = 12, Female = 42; Control: Male = 13, Female 40. All participants had experience of the real campus, but not of the computer model. Participants who took part in Experiments 1 or 2a were not eligible to participate within Experiment 2b.

Materials and Apparatus

A video following the same routes given to the participants in Experiments 1 and 2a, exploring the inside and outside of the computer generated building was presented on a large screen at the front of a 50-seat lecture theater. Participant responses to questions regarding demographics and orientation within the building were recorded via Turningpoint software and clickers.

Procedure

The study was run over six different sessions with different sets of participants in each session. Participants were shown a 2-min video exploring the outside of the computer generated building used in Experiments 1 and 2 from a first person perspective. The route followed the directions given to the participants in the first stage of those experiments. The participants were then shown a 2-min video exploring the inside of the building following the route taken by the participants in Experiments 1 and 2. Prior to the orientation trials, the participants were given one of three priming speeches. In sessions 1 and 4, participants were told that females perform better than males in certain spatial tasks. In sessions 2 and 5, participants were told that males perform better than females in certain spatial tasks. In session 3 and 6, participants were told that there is no difference between males and females in certain spatial tasks (See Table 1).

During the orientation test trials, participants were shown videos of the rooms used in Experiments 1 and 2a. On each wall of the room there was a letter (Figure 5). Participants were asked to respond on the clickers to indicate whether they would be facing wall A, B C or D, if they were to face the external target cue. A response referring to the correct wall would be given a score of 0 degrees orientation error, a response referring to either wall adjacent to the correct wall would be given a score of 90 degrees and a response referring to the opposite wall would receive a score of 180 degrees.

Results

Figure 6 illustrates the orientation error among the groups. As with Experiments 1 and 2a, performance in the External room was better than in the Internal room. Both males and females that were told their gender normally performs worse than the opposite sex (Group Threat) performed worse than those told that they would normally either perform better than the opposite sex (Group Lift) or the same as the opposite sex (Group Control). Performance in the Internal room for all males in the Lift condition was without error. These observations were confirmed via a 2 (room type) x 2 (gender) x 3 (condition) mixed design ANOVA with orientation error as the dependent variable. The main effect of room type was significant, $F(1, 154) = 15.04$, $p < 0.01$, $\eta_p^2 = .09$, indicating that orientation error was smaller in the external room than in the interior room. There was a significant interaction between room type and condition, $F(2, 154) = 5.65$, $p = 0.004$, $\eta_p^2 = .07$. Further analysis of the simple main effects revealed that the effect of room type was only significant in the threat condition, $F(1, 154) = 20.74$, $p < 0.01$, $\eta_p^2 = .34$. There was also only a significant effect of condition in the internal room, $F(2, 308) = 14.60$, $p < 0.01$, $\eta_p^2 = .165$. Post hoc independent t-tests using the Bonferroni correction demonstrated participants in Group Threat made significantly more orientation errors than participants in either Group Lift or Group Control. The interaction between room type and gender was not significant, $F(1, 154) = 2.16$, $p = 0.144$, $\eta_p^2 = .014$, nor was the three way interaction between room type, gender and condition, $F<1$.

The main effect of gender was not significant, $F(1, 154) = 1.36$, $p = 0.246$, $\eta_p^2 = .009$. The main effect of condition was significant, $F(2, 154) = 8.88$, $p < 0.000$, $\eta_p^2 = .103$. However, as stated there was a significant condition x room type interaction showing that the effect of condition was only significant in the internal room. The gender x condition interaction was not significant, $F<1$.

The results of Experiment 2b showed that both males and females are susceptible to the stereotype threat. For males this finding confirms the results of Experiment 1 and extends the generality of the effect to a passive presentation of the environment. The results from the females in Experiments 1 and 2b suggest that including the internal cues made the task easier and allowed the females to maintain orientation in the internal room. In Experiment 1, maintaining orientation in the internal room was too difficult for the female participants in any of the stereotype conditions and thus the priming statements were irrelevant. Making the task easier in Experiment 2b meant performance was vulnerable to the stereotype manipulation. There was not a significant effect of stereotype manipulation in the external room for either males or females. This suggests that the external plus internal cues made the task too easy to be influenced by the stereotype manipulation and supports the Keller (2007) findings that difficulty interacts with stereotype manipulation.

General Discussion

The aim of this paper was firstly to examine whether performance on a novel spatial orientation task could be influenced by a stereotype manipulation, given the weakness of the effect in spatial tasks noted by Doyle and Voyer (2016) and then to examine whether the influence was moderated by task difficulty as demonstrated in maths ability (Keller,2007). Results from Experiments 1 and 2b demonstrate that the participants' ability to orient within a nested environment was influenced by the presence of a stereotype manipulation. Data suggest that those exposed to a stereotype lift recorded the lowest mean orientation errors, with participants exposed to a stereotype threat manipulation recording the highest mean orientation errors. However, although numerically different, there was no significant difference between the lift and control groups in either experiment. There was a significant difference between the threat and control groups for males in Experiment 1 and for both males and females in Experiment 2b, suggesting that threat had the greater effect on spatial performance.

Task difficulty was manipulated by requiring the participants to maintain orientation with reference to an external target cue. In the Internal rooms, this could only be done by spatial updating, involving continually updating their spatial relationship as they moved through the building. In External rooms, orientation could be established by learning the constant spatial relationship between the visible external cue and the occluded external target cue. Research from Wang and Brockmole (2003) suggested that the former would be more difficult. Results from Experiment 1 and 2 of the current paper support and extend these previous findings, as participants in the novel computer generated environment performed better in the External room.

Experiment 1 demonstrated an effect of stereotype threat in both the difficult internal and the easier external room, suggesting task difficulty did not interact with the stereotype threat. This result will be discussed later. The effect of a stereotype threat was, however, only visible in male performance. Female performance in all three groups was not significantly different from chance, suggesting the task, at least in the internal room, was too difficult and so stereotypes had become irrelevant (Spencer, 1997).

Having made the task easier in Experiment 2a by the inclusion of internal orientation cues which maintained a constant spatial relationship with the external cues, Experiment 2b showed an effect of stereotype threat not only in males and but also in females. In Experiment 2b, the effect was only seen in the more difficult internal room. This finding supports the results of Keller (2007) but extends it to a spatial task. Returning to why we found the effect in males in the easier external room in Experiment 1 but not in Experiment 2b, it may be that the addition of the internal cues in the latter study in addition to the available stable relationship of the external cues meant the task was too easy and once more the stereotype threat was ineffective. Overall the results suggest that for the stereotype threat to affect performance, task difficulty has to be at a specific level. If it is too difficult, as for the female participants in Experiment 1, stereotype threat is irrelevant. If the task is too easy, as in the External room trials in Experiment 2b, with both internal and external cues present, the stereotype threat will not have an impact on performance.

Findings from Experiment 2b, using our naturalistic orientation task, both extend and support those reported by McGlone and Aronson (2006) and Moè (2009) who found that the addition of a stereotype threat had a negative impact on the performance of female participants when they were asked to complete a mental rotation task. It also should be noted that, unlike Rosenthal et al. (2012), we did not find an effect of stereotype lift. Although numerically, in both Experiments 1 and 2b participants performed better in the lift groups than the control groups, the results were not significant. It may well be that task difficulty plays a crucial role when the effect is found and that the current set of tasks were too easy for the lift condition to boost performance above that of controls.

The difference between the results for females in Experiment 1 and 2b is presumed to have been due to reduction of the difficulty of the orientation task in the internal room. It has to be acknowledged that there were other changes between the studies that may have resulted in the change in results. In Experiment 1, participants explored the VE individually, while in Experiment 2b participants were shown a video within a group. Hegarty et al (2006) used exploration of a VE and passive exposure to a video and found similar levels of performance in both means of presentation. Inzlitch and Ben-Zeev (2000) noted that when females were presented maths problems in a group containing males and females, the females performed worse than when they were in a group of all women. Possibly the female participants in the group presentations of Experiment 2b, which all contained male and female participants, were more susceptible to the stereotype threat than female participants exploring the VE alone.

Another factor which might have impeded the stereotype lift in females compared to the threat might be that the lift priming statement for females that "on certain tasks females outperform males in their spatial performance" is not as forceful as the threat priming statement "males are often found to outperform females on spatial tasks". However, if this was a factor it might be expected that we would see a bigger effect in lifting males performance and see the threat having a less of an effect. The fact that for both males and females the lift effect is weak compared to the threat effect suggests it is the more to do with spatial performance being more vulnerable to a threat statement than the difference between the lift and threat statements.

The studies within the current paper have focussed on consistent gender differences. However, it must be noted that recent work has demonstrated differences between individuals in the absence of sex differences (e.g. Hegarty et al , 2006; Ishikawa, & Montello, 2014; Weisberg, Schinazi, Newcombe, Shipley, & Epstein, 2014; Weisberg & Newcombe, 2016). For example, Weisenberg and Newcombe showed that imprecise navigators based on their poor accuracy in a pointing task had lower verbal and spatial working memory. The results of Experiment 2b suggest that the stereotype threat manipulation is effective across both genders; however, it would further our understanding of why stereotype threat is effective if we were to investigate the differences within the gender groups to assess how much of this difference can be apportioned to differences in such things as spatial working memory.

The environment chosen for these experiment was a virtual model of the Psychology building, University of Southampton, and its surroundings, a location with which all of the participants would be expected to be familiar. Participants were all Psychology undergraduates at the University of Southampton. They all had experience of attending lectures in the building on which the computer model was based. It would not seem unreasonable to assume that they had taken the routes around and through the actual building depicted in the experiments. Research has documented that familiarity with an environment improves participants' ability to navigate and orient (Ruddle, Payne, & Jones, 1997). Presson, (1987) and Siegel and White (1975) have argued that substantial familiarity with an environment leads to the development of orientation free survey knowledge. It would again extend our understanding of how the stereotype manipulation affects a more difficult task if we were to test the participants in an environment with which they were unfamiliar.

Limitations to the conclusions of the current studies should be mentioned. The cognitive ability of the participants was not directly measured. The participants were all psychology students and their first language was English. Therefore the cognitive ability of participants was considered to be fairly uniform. However, it cannot be ruled out that differences such as that between Group Threat and Group Lift were not a result of stereotype threat but a result of differences in cognitive ability. It also should be noted that the number of males compared to females in Experiments 2b was quite small. However, the results of the male participants in Experiments 1 and 2b were very similar, showing an effect of stereotype threat manipulation. Taking the two experiments together along with those of the much larger sample of females in Experiment 2b, would seem to provide reasonable evidence for the conclusion that the stereotype threat manipulation was effective in this novel spatial test.

Finally, the studies in the current paper suggest two ways that orientation within a nested environment such as a building can be improved. This has important implications for helping people in an emergency to find an exit point quickly. Firstly, allowing a viewpoint of some part of the external array allowed the participants to remain oriented in terms of any other point on the external allocentric array. For the design of a building, it would be beneficial wherever possible to provide windows on the main thoroughfare through which people can view the outside and thus remain orientated. A second way the studies allowed participants to remain oriented was the inclusions of internal orientation cues. The orientation cues maintained a constant spatial relationship to the cues in the external environment. This meant the participants did not have to rely on constantly updating their changing position, a task which, as the current studies and those of Wang and Brockmole (2003) have shown, is a difficult thing to do.

Conclusion

The studies of this paper have shown that maintaining orientation within a nested environment is difficult but can be improved by exposure to part of the external array or the inclusion of internal cues that have a constant spatial relationship to the external array. The accuracy of orientation can be influenced by the presence of a stereotype threat manipulation. Participants exposed to a stereotype threat manipulation recorded higher orientation errors than control participants or participants exposed to stereotype lift. The effectiveness of the stereotype threat is influenced by task difficulty. Spatial skills have been suggested to be strong predictors of achievement in science, technology, engineering and maths (STEM) subjects (Casey, Pezaris, & Nuttall, 1992). Given the disparity in the number of females to males going into STEM subjects (Chipman, Krantz, & Silver, 1992), finding ways to enhance spatial ability would seem to be an important step in addressing the disparity. Understanding the reasons for the occurrence of stereotype threat in both spatial and mathematical abilities, particularly in females, would go some way to enhancing ability in these areas.

References

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*(1), 29-46.

Campbell,S. M., & Collaer, M. L. (2009). Stereotype threat and gender differences in performance on a novel visuospatial task, *Psychology of Women Quarterly, 33,* 437-444.

Carlson, L. A., Holscher, C., Shipley, T. F., & Dalton, R. C. (2010). Getting Lost in Buildings. *Current Directions in Psychological Science, 19*, 284-289.

Casey, M. B., Pezaris, E. & Nuttall, R. L. (1992).Spatial ability as a predictor of math achievement - the importance of sex and handedness patterns. *Neuropsychologia*, *30*, 35-45.

Chipman, S. F., Krantz, D. H., & Silver, R. (1992). Mathematics anxiety and science careers among able college women. *Psychological Science, 3*, 292–295.

Dabbs, J. M., Chang, E.L., Strong, R. A., & Milun, R. (1998). Spatial ability, navigation strategy and geographic knowledge among men and women. *Evolution and Human Behaviour, 19,* 89-98.

Doyle, R. A. & Voyer, D. (2016) Stereotype manipulation effects on math and spatial test performance: A meta-analysis. *Learning and Individual Differences, 47,* 103–116.

Duff, S., & Hampson, E. (2001). A Sex Difference on a Novel Spatial Working Memory Task in Humans. *Brain and Cognition, 47,* 470–493.

Farrell, M. J., & Robertson, I. H. (1998) Mental rotation and the automatic updating of body-centred spatial relationships. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 2,* 227 -233.

Flore, P., C., & Wicherts, J. M. (2015), Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology. 53*, 25-44.

Hausmann, M., Schoofs, D., Rosenthal, H. E. S. & Jordan, K. (2009). Interactive effects of sex hormones and gender stereotypes on cognitive sex differences – a psychobiosocial approach. *Psychoneuroendocrinology 34*(3): 389-401.

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, *34*(2), 151-176.

Heil, M., Jansen, P., Quaiser-Pohl, C., & Neuberger, S. (2012) Gender-specific effects of artificially induced gender beliefs in mental rotation. Learning and Individual Differences, 22, 350-353.

Hund, A. M., & Minarik, J. L. (2006). Getting from here to there: Spatial anxiety, wayfinding strategies, direction type, and wayfinding efficiency. *Spatial Cognition and Computation, 6*, 179–201.

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment:Why femalesare susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11*, 365–371.

Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive psychology*, *52*(2), 93-129.

Keller, J. (2007) Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance. *British Journal of Educational Psychology, 77*, 323–338

Keppel, G. (1973). *Design and analysis: A researchers' handbook*. Englewood Cliffs, NJ: Prentice-Hall.

Lawton, C., Charleston, S. I. & Zieles, A.S. (1996). Individual- and gender-related differences in indoor wayfinding. *Environment and Behavior, 28,* 204-219.

Lawton, C., & Kallai, J. (2002). Gender differences in wayfinding strategies and anxiety about wayfinding: A cross-cultural comparison. *Sex Roles, 47*, 389–401.

Lawton, C., & Morrin, K. A. (1999) Gender Differences in Pointing Accuracy in Computer-Simulated 3D Mazes. *Sex Roles, 40*, 73-92.

McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience and spatial reasoning. *Journal of Applied Developmental Psychology, 27*, 486-493.

McGlone, M. S., Aronson, J., & Kobrynowicz, D. (2006). Stereotype threat and the gender gap in political knowledge. *Psychology of Women Quarterly, 30*, 392-398.

Meneghetti, C., Muffato, V., Suitner, C., De Beni, R., & Borella, E. (2015) Map learning in young and older adults: The influence of perceived stereotype threat. *Learning and Individual Differences, 42,* 77-82.

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

Moè, A. (2009) Are males always better than females in mental rotation? Exploring a gender belief explanation, *Learning and Individual Differences, 19,* 21-27.

Nguyen, H. H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93,* 1314–1334.

Nichols, A. L., & Maner, J. K. (2008). The good-subject effect: Investigating participant demand characteristics. *The Journal of General Psychology, 135,* 151–166.

Prestopnik, J. L., & Roskos-Ewoldsen, B. (2000). The relations among wayfinding strategy use, sense of direction, sex, familiarity, and wayfinding ability. *Journal of Environmental Psychology, 20,* 177–191.

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*(1), 55-71.

Rosenthal, H. E. S., Norman, L., Smith, S. P. & McGregor, A. (2012). Gender-based navigation stereotype improves men's search for a hidden goal. *Sex Roles, 67*. 682-695.

Riecke, B.E., Cunningham, D.W., & Bulthoff, H.H. (2007) Spatial updating in virtual reality: The sufficiency of visual information Psychological Research – *Psychologische Forschung, 71*, 298-313

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental, Learning, Memory, and Cognition*, *15*(6), 1157 - 1165.

Sandstrom, N. J., Kaufman, J., & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task. *Cognitive Brain Research, 6,* 351- 360.

Schmader, T., Johns, M., & Barquissau, M. (2004) The Costs of Accepting Gender differences: The Role of Stereotype Endorsement in Women's Experience in the Math Domain, *Sex Roles, 50*, 835-851.

Schmitz, S. (1999) Gender differences in acquisition of environmental knowledge related to wayfinding behavior, spatial anxiety and self-estimated environmental competencies. *Sex Roles, 41,* 71-93.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*(1), 80-83.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613-629.

Steele, C. M. (1999). Stereotype threat and black college students. *American Association for Higher Education Bulletin, 52*(6), 3-6.

Thorndyke, P. W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology, 14,* 560-589.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental Rotations: A group test of three dimensional spatial visualisation. *Perceptual and Motor Skills, 47*(2), 599-604.

Walton, G. M., & Cohen, G. L. (2003). Stereotype Lift. *Journal of Experimental Social Psychology, 39,* 456-467.

Wang, R. F., & Brockmole, J. R. (2003). Human navigation in nested environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 398-404.

Weisberg, S. M., Schinazi, V. R., Newcombe, N. S., Shipley, T. F., & Epstein, R. A. (2014). Variations in cognitive maps: Understanding individual differences in navigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 669.

Weisberg, S. M., & Newcombe, N. S. (2016). How do (some) people make a cognitive map? Routes, places, and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(5), 768.

Figure Captions

Figure 1a. View of external environment during exploration seen by participants.

Figure 1b. View from Internal Orientation test room. It can be seen that no external cues are visible.

Figure 1c. Map of University Environment

Figure 2. Group mean orientation error in Experiment 1

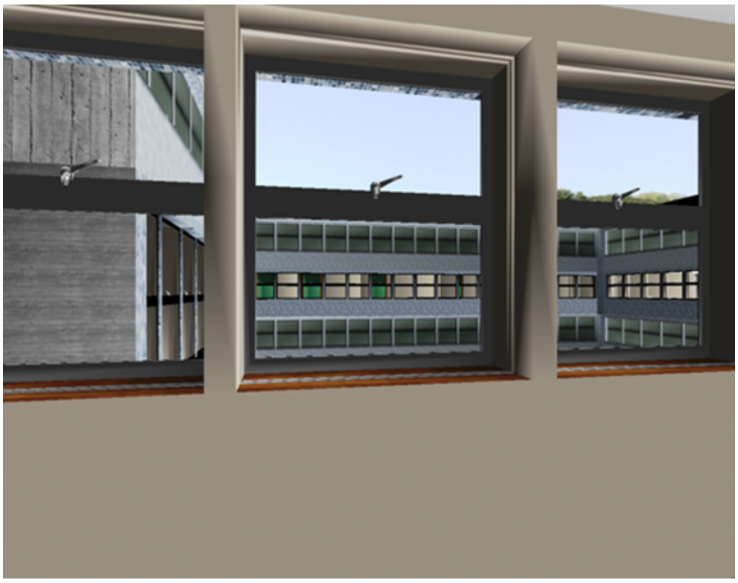
Figure 3a. View of External Target (Car Park) seen by participants in Group Cue with orienting cue running along centre of building wall.

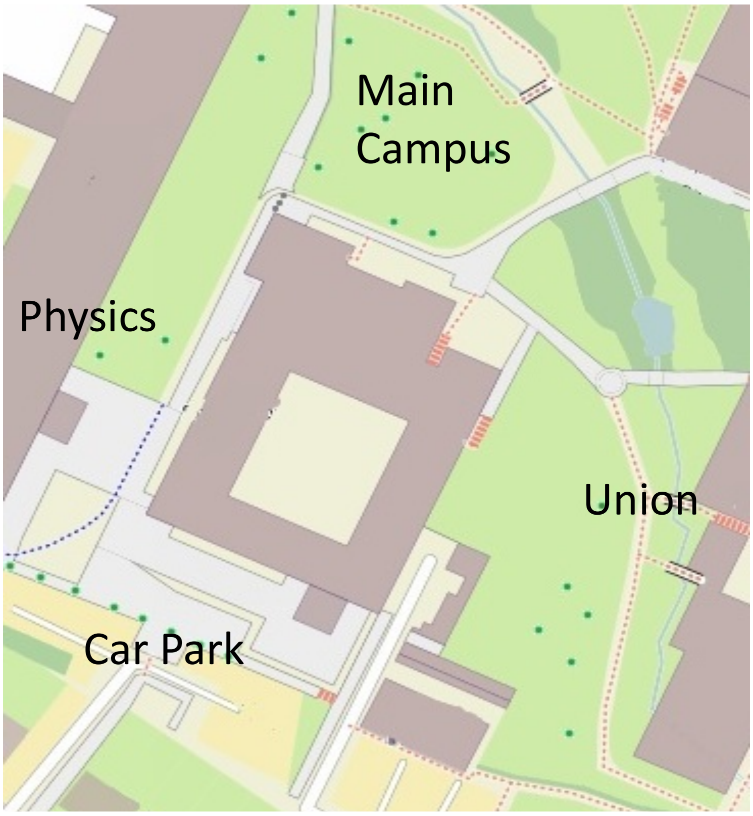
Figure 3b. View from External Orientation test room for Group Cue with orienting cue running along top of window.

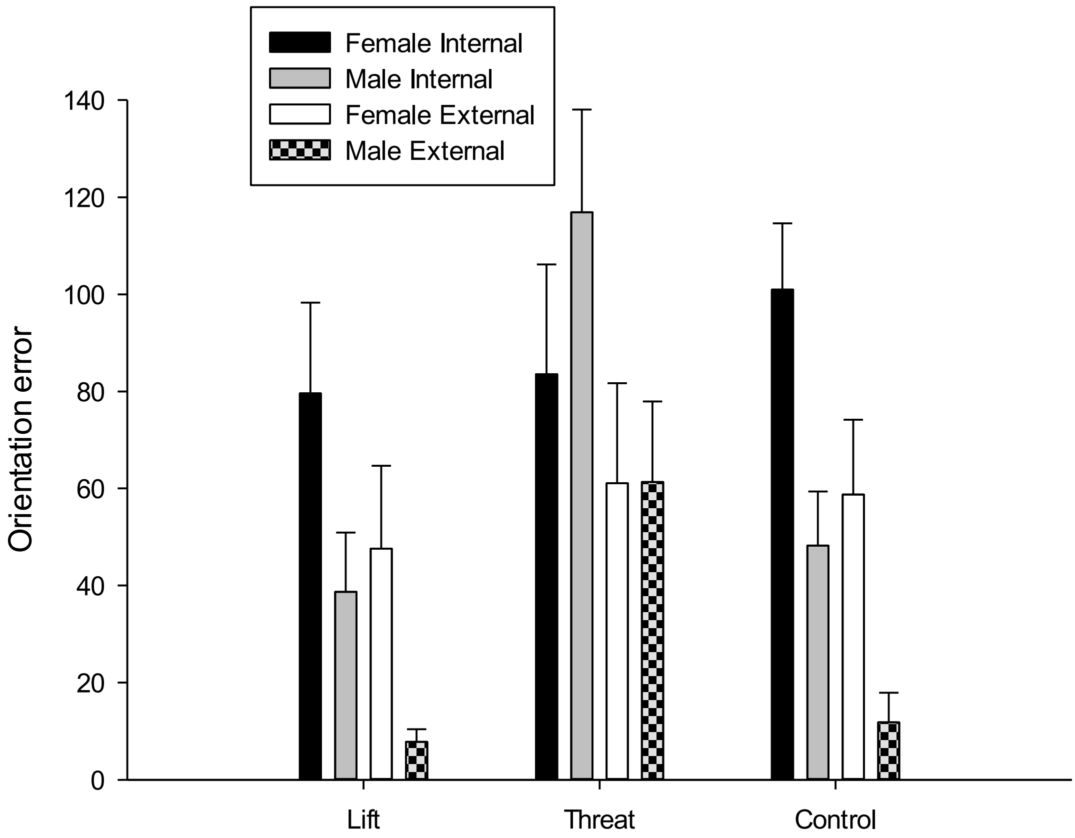
Figure 4. Group mean orientation error in Experiment 2a.

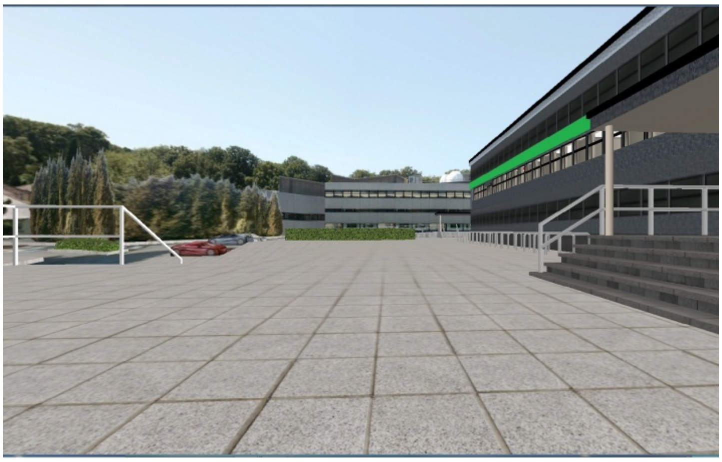
Figure 5. View from Internal Orientation test room with orienting cues and labelled walls.

Figure 6. Group mean orientation error in Experiment 2b.

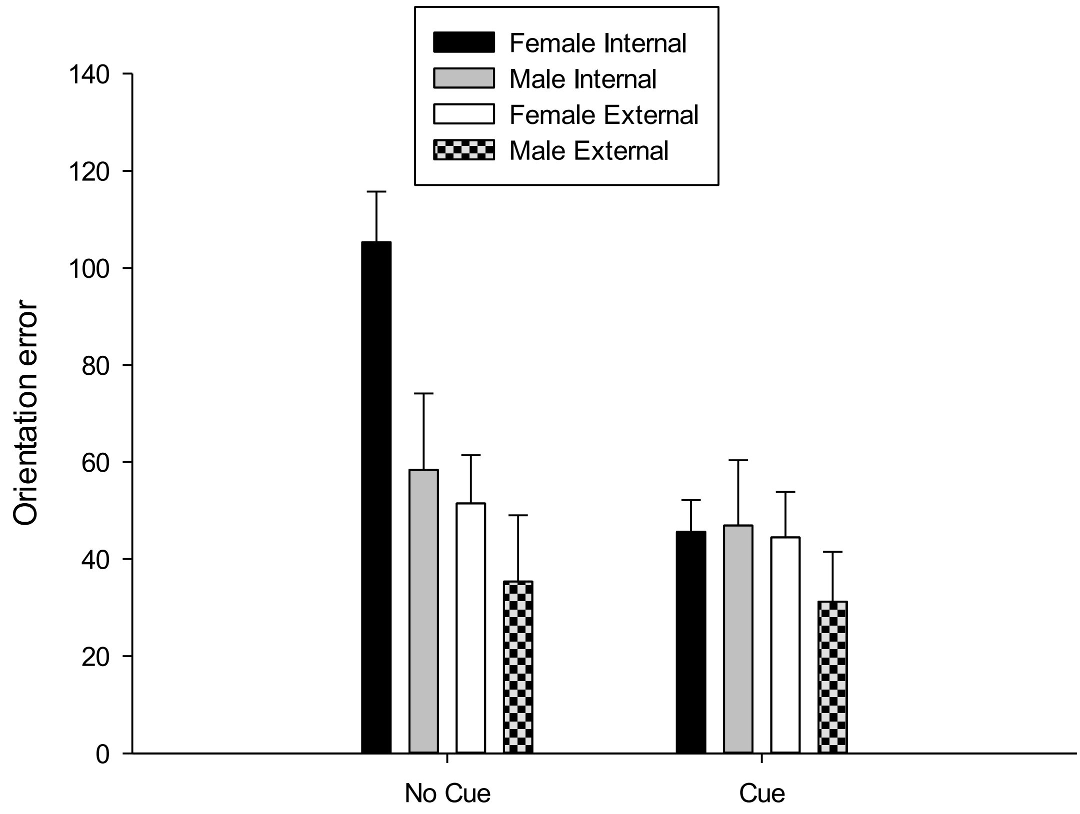



Main
Campus
Physics
Union
Car Park


Female Internal
Male Internal
Female External
Male External
Orientation error
0
20
40
60
80
100
120
140
Lift
Threat
Control



Experiment 2a



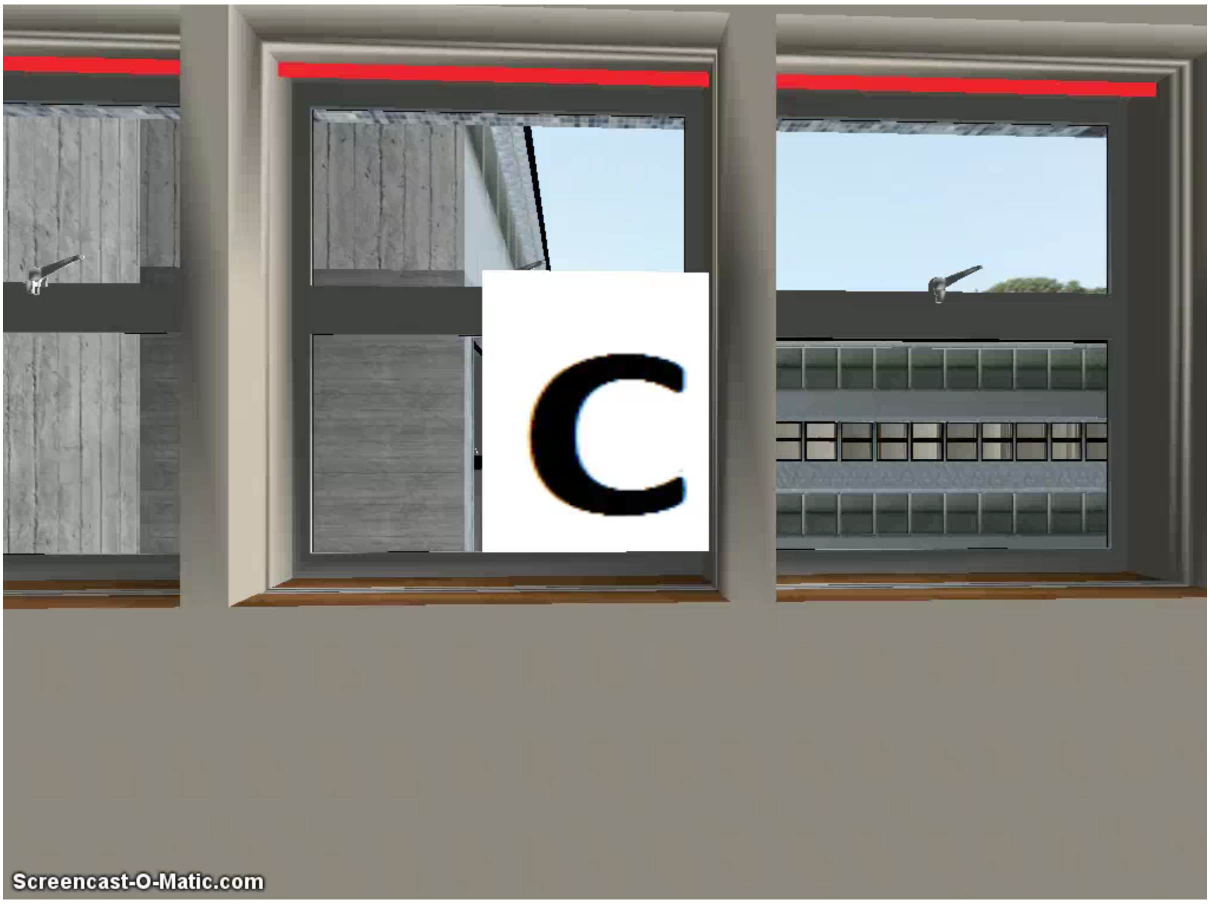


Figure 5 View from internal rooms in orienting cues and labelled walls.

Experiment 2b

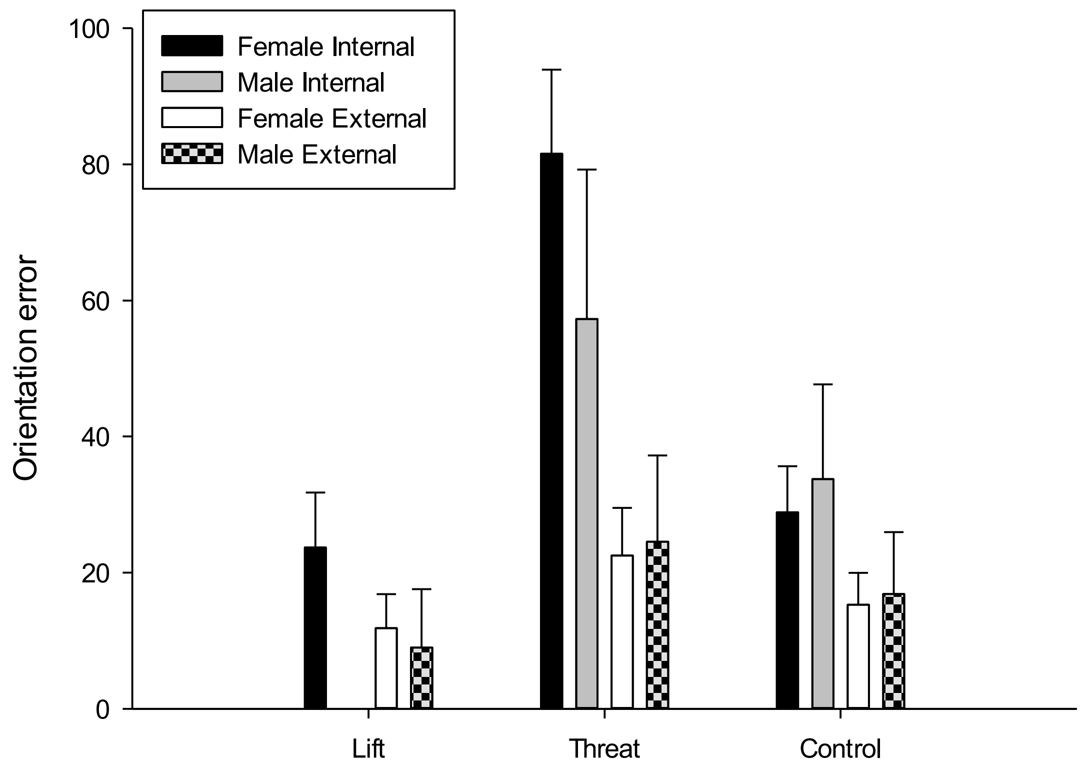


Table 1: Stereotype Priming Scripts

Males outperform females

Today you will be taking part in a spatial task. There has been much research into this area, some of which has looked specifically at gender differences. It has been found that males and females use different strategies when navigating, although males are often found to outperform females on spatial tasks (Dabbs, Chang, Strong, & Milun, 1998).

Dabbs, J. M., Chang, E.L., Strong, R. A., & Milun, R. (1998). Spatial ability, navigation strategy and geographic knowledge among men and women. *Evolution and Human Behaviour, 19,* 89-98.

Females outperform males

Today you will be taking part in a spatial task. There has been much research into this area, some of which has looked specifically at gender differences. It has been found that males and females use different strategies when navigating, and on certain tasks females outperform males in their spatial performance (Duff, & Hampson, 2001).

Duff, S., & Hampson, E. (2001). A Sex Difference on a Novel Spatial Working Memory Task in Humans. *Brain and Cognition, 47,* 470–493.

No difference between males and females

Today you will be taking part in a spatial task. There has been much research into this area, some of which has looked specifically at gender differences. It has been found that males and females use different strategies when navigating, although no significant differences have been found in their spatial performance (Sandstrom, Kaufman, & Huettel, 1998).

Sandstrom, N. J., Kaufman, J., & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task. *Cognitive Brain Research, 6,* 351- 360.