Mental rotation does not account for sex differences in left–right confusion

Sebastian Ocklenburg, a,⇑, Marco Hirnstein b, Hanno Andreas Ohmann a, Markus Hausmann b

a Department of Biopsychology, Faculty of Psychology, Ruhr-University Bochum, D-44780 Bochum, Germany
b Department of Psychology, Durham University, South Road, Durham DH1 3LE, United Kingdom

1. Introduction

Left–right confusion (LRC) is a common phenomenon experienced by many humans in various situations. Apart from anecdotal evidence, a number of scientific studies reported that healthy adults sometimes experience difficulty when telling left from right (Harris & Gitterman, 1978; McMonnies, 1990; Wolf, 1973). Particularly women believe themselves to be more prone to LRC than men when they are asked to self-assess their ability to make fast and accurate left–right judgements (Hannay, Ciaccia, Kerr, & Barett, 1990; Hirnstein, Ocklenburg, Schneider, & Hausmann, 2009; Jaspers-Feyer & Peters, 2005; Jordan, Wüstenberg, Jaspers-Feyer, Fellbrich, & Peters, 2006). However, whether sex differences in self-ratings have behavioural consequences remains unclear as some studies report lower accuracy (Bakan & Putnam, 1974; Ofte, 2002; Ofte & Hugdahl, 2002b) and slower reaction times in LRC tasks (Snyder, 1991) in women compared with men, while others report no sex differences (Teng & Lee, 1982; Williams, Standen, & Ricciardelli, 1993). Jordan et al. (2006) suggest that the conflicting results in LRC can be explained by sex differences in specific spatial abilities. For example, in the Bergen Left–Right Discrimination Test (Ofte, 2002; Ofte & Hugdahl, 2002a, 2002b), participants have to mark either the left or right hand of stickman figures that were drawn either from the front or from the back. Performance in this task may be influenced by the fact that participants have to mentally rotate figures that are shown from the front to make a left–right decision. Mental rotation refers to the ability to mentally rotate two- or three-dimensional objects. Typically, men outperform women in tasks of mental rotation by effect sizes of more than one standard deviation (Linn & Petersen, 1985; Masters & Sanders, 1993; Peters, Lehmann, Takahira, Takeuchi, & Jordan, 2006; Voyer, Voyer, & Bryden, 1995). Jordan et al. (2006) argue that several studies that reported sex differences in LRC (e.g., Bakan & Putnam, 1974; Ofte, 2002; Ofte & Hugdahl, 2002b; Snyder, 1991) are confounded by mental rotation.

In their own study Jordan et al. (2006) tested participants on LRC tasks that did or did not involve mental rotation. In the LCR task that involves mental rotation, participants had to navigate through a three-dimensional virtual maze on the basis of a map they had seen before. Here, a sex difference emerged with men being significantly faster than women. Participants had to mentally rotate their memorized overview image of the maze depending on their position in order to make a correct left–right decision at junctions in the maze. The assumption that the sex difference in this task may be confounded by sex differences in mental rotation is partly supported by a significant positive correlation between the time needed to navigate through the maze and the performance in the Mental Rotation Test (Peters et al., 1995). However, only men, and not women, showed this correlation suggesting that mental rotation may not entirely account for the observed sex dif-
ference. In the task that did not involve mental rotation, participants had to decide as quickly as possible whether an object (a bunch of pencils) is to the left or right of another object (an iced-tea can) viewed solely from the participants' perspective. No sex differences were observed in this LRC task.

However, there are findings that argue against Jordan et al.'s (2006) notion and suggest that sex differences in LRC exist independently of mental rotation. For example, in the Left–Right Commands Task, Hirnstein et al. (2009) asked their participants to follow verbal commands concerning their own left and right body parts. This task showed a significant sex difference with a large effect size ($d > 0.8$), although no mental rotation was required. Similarly, in the pointing-hands task, in which participants viewed pictures of pointing hands in different orientations and had to indicate either the pointing directions or whether they saw a left or right hand, women committed more LRC than men.

Taken together, it remains still unclear from these studies whether sex differences found for some LRC tasks are a mere artefact of sex differences in mental rotation. To investigate whether sex differences in LRC exist independently of mental rotation, we compared the susceptibility to LRC in men and women that were matched for their mental rotation ability as measured with the Mental Rotation Test (Peters et al., 1995; Vandenberg & Kuse, 1978). Two LRC tasks were used: the Left–Right Commands Task and the Bergen Left–Right Discrimination Test. Both tasks showed robust sex differences previously (Hirnstein et al., 2009; Ofte, 2002; Ofte & Hugdahl, 2002b). If sex differences in LRC are independent of mental rotation, then men and women matched for their mental rotation abilities should still display sex differences in both tasks.

2. Methods

2.1. Participants

Overall, 91 neurologically healthy women ($N = 50$) and men ($N = 41$) participated in the present study. The mean age was 23.5 years ($SD = 3.45$) for women and 25.29 years ($SD = 3.68$) for men. All participants were right-handed, as determined by the Edinburgh Handedness Inventory (EHI; Oldfield, 1971). The laterality quotient (LQ) provided by this test is calculated as $LQ = ([R - L]/(R + L))/100$, resulting in values between $-100$ and $+100$. Positive values indicate right-handedness, while negative values indicate left-handedness. Women had a mean LQ of 91.36 ($SD = 15.12$), while men had a mean LQ of 87.37 ($SD = 17.43$). There was no sex difference in LQ ($t(89) = 1.17, p = 0.25$).

2.2. Procedure

Each test session began with two behavioural LRC tasks, the Left–Right Commands Task (Hirnstein et al., 2009) and the Bergen Left–Right Discrimination Test (Ofte, 2002), in counterbalanced order. Subsequently, they completed the EHI, the left–right self-rating questionnaire and the Mental Rotation Test (Peters et al., 1995). Performing the LRC tasks at the beginning of each test session prevented possible stereotype activation effects of the self-rating questionnaire.

2.3. Mental Rotation Test

To select groups of men and women that are matched for mental rotation abilities, all participants performed the Revised Vandenberg and Kuse Mental Rotation Test by Peters et al. (1995). This paper-and-pencil test consists of two subtests with 12 items each. Each item consists of five cube figures, one of them being the target figure. Of the other four figures, two are rotated versions of the target figure whereas the other two cannot be matched with the target figure via rotation. Participants have 3 min to finish each subtest. A score of ‘one’ per item was given, if both rotated versions of the target had been identified correctly. In all other cases, a score of ‘zero’ was given, resulting in an overall score between zero and 24 for each participant.

2.4. The Bergen Left–Right Discrimination Test

The original paper–pencil version of the Bergen Left–Right Discrimination Test (Ofte, 2002) was adapted for use on a computer. Stimuli were presented on a standard PC monitor using Presenta® (Neurobehavioral Systems, Inc., Albany, USA). The stimulus set consists of 96 line drawings of a figure with a height of 11 cm. When the head of the figure is highlighted in black, the figure is viewed from the back, so that the left hand of the figure is presented on the left side of the participant. When the head of the figure is highlighted in white, the figure is viewed from the front, so that the left hand of the figure is presented on the right side of the participant. In half of the trials the figure is viewed from the back, in the other half it is viewed from the front. The shoulders are represented by a black triangle. The arms of the figures are located at different positions in relation to the body, with no arm, one arm or both arms crossing the vertical midline of the figure. Circles at the end of the arms represent the hands of the figures. For half of the figures the left hand is coloured red, for the other half the right hand is highlighted in red (see Fig. 1).

Participants had to decide by button press whether the labels ‘R’ or ‘L’ below the figure matched the left or right hand highlighted in red. For example, if the right hand was highlighted in red and the label below the figure showed ‘R’, the trial was correct. Labels were correct in 50% of the trials (matching trials) and incorrect in the other 50% of trials (mismatching trials). The response keys were arranged vertically to prevent possible spatial stimulus–response compatibility effects. Each stimulus was presented until participants pressed one of the two response keys. After a response was made a blank screen was presented for 1000 ms before the next trial began. The original set of 48 stimuli was used twice, once with left and once with right hands highlighted in red. Stimuli were presented in pseudo-randomised order, and LRC rates as well as reaction times for correct left–right decisions were calculated as dependent variables.

Fig. 1. Stimuli of the Bergen Left–Right Discrimination Test. The stickman figures are viewed from the front (white head) or from the back (black head) and exemplify the three possible arm positions (no arm, one arm or both arms crossing the vertical midline of the stickman figure). Participants have to decide whether the label below the stickman figure corresponds to the hand highlighted in red (here in grey). The first two stickman figures show correct items, whereas the third stickman figure shows an incorrect item.
2.5. Left–Right Commands Task

In the Left–Right Commands Task, participants were asked to follow verbal commands concerning their own left and right body parts, while being recorded with a video camera (Hirnstein et al., 2009). At the beginning of the experiment, participants sat upright on a chair with their hands on their knees (starting position). Sixty verbal commands were presented via loudspeakers in a pseudo-randomised order. The task consisted of three conditions with 20 commands each. In the control condition, neutral commands which did not involve any left or right body parts were given (e.g., ‘Open your mouth!’). In the simple condition, participants were asked to move one left or right limb (e.g., ‘Lift your left hand!’), and in the complex condition participants were asked to move two left or right body parts at the same time (e.g., ‘Touch your left foot with your right hand!’). Participants were asked to follow each command and return to the starting position afterwards. After 750 ms the next command was presented. Only if participants follow the command correctly, e.g., ‘Lift your left hand!’ but mix up left and right, i.e., they lift the right instead of the correct left hand, is this considered as LRC. LRC error percentage scores are calculated for both simple and complex commands.

2.6. Self-rating questionnaire

The LRC self-rating questionnaire was used previously (Hirnstein et al., 2009) and originally adopted from Jordan et al. (2006). The questionnaire consists of eight items. The first four items focus on left–right judgements (LRC-items, e.g., “As an adult, I have noted difficulty when I quickly have to identify right versus left.”). These items were derived from Hannay et al. (1990). The last four items are more generalised directional questions (DIR-items, e.g., “Are you (quickly) able to locate north without the aid of a compass?”) derived from Jaspers-Feyer and Peters (2005). For each item participants had to indicate on a five-point scale whether they had ‘no problems at all’ (one point) or ‘severe’ problems (five points), resulting in individual scores between 4 and 20 for LRC- and DIR-items.

3. Results

Effect sizes are shown as Cohen’s d or the proportion of variance accounted for (partial $\eta^2$). P-levels for post hoc t-tests were adjusted using Bonferroni correction.

3.1. Mental Rotation Test

As expected, the sex difference in mental rotation was significant ($t(89) = 5.54; p < 0.001, d = 1.17$), with men solving more items correctly (mean ± SD) (14.10 ± 3.78) than women (9.56 ± 3.97). To investigate whether sex differences in LRC persist when men and women were matched for mental rotation performance, 23 pairs of men and women with identical mental rotation scores were formed. In cases in which more than one participant of the same sex scored identical, pairs were additionally matched for age. There was no significant age difference between women (22.8 years, SD = 3.66) and men (24.09 years SD = 3.54) in the matched group ($t(44) = 1.19, p = 0.24$). The remaining 45 participants who could not be matched were excluded.

3.2. Bergen Left–Right Discrimination Test

A $3 \times 2 \times 2$ mixed ANOVA with Arm position (no arm, one arm, both arms crossing the midline) and View (back view, front view) as within-participants factor and sex as between-participants factor was calculated for LRC rate and reaction time. For LRC rates, only the main effect of sex was significant ($F_{(1,44)} = 8.77, p < 0.01, \eta^2 = 0.17$), indicating that overall men made fewer errors than women (see Fig. 2). The arm position x view interaction only approached significance ($F_{(2,42)} = 3.09, p = 0.051, \eta^2 = 0.07$). All other main effects and interactions were not significant (all $F_{(1,44)} < 0.82, ns$). For response times only the effect of arm position was significant ($F_{(2,42)} = 5.55, p < 0.01, \eta^2 = 0.11$). Post hoc t-tests revealed longer reaction times when both arms of the figure crossed the vertical midline than when only one arm ($p < 0.05$) or no arm ($p < 0.01$) crossed the vertical midline. In contrast, no difference in reaction times was observed between trials in which one arm or no arm crossed the vertical midline of the figure ($p = 1.00$). All other main effects and interactions were not significant (all $F_{(1,44)} < 2.43, ns$).

3.3. Left–Right Commands Task

The mean LRC rates are shown in Fig. 3. A $2 \times 2$ mixed ANOVA with condition (simple, complex) as within-participants factor and sex as between-participants factor was calculated. The main effect of sex was significant ($F_{(1,44)} = 5.49, p < 0.05, \eta^2 = 0.11$), indicating less LRC in men than women. The main effect of Condition was also significant ($F_{(1,44)} = 18.31, p < 0.001, \eta^2 = 0.29$). Participants showed more LRC in the complex than in the simple condition. The interaction between sex and condition was not significant ($F_{(1,44)} = 1.23, ns$).

3.4. Self-rating questionnaire

The mean self-rating scores were subjected to a $2 \times 2$ mixed ANOVA with self-rating type (LRC, DIR) as within-participants factor and sex as between-participants factor. Women ($2.78 \pm 0.64$) rated themselves as being more prone to LRC and DIR than men ($1.96 \pm 0.45$) as indicated by a significant main effect of sex ($F_{(1,44)} = 25.29, p < 0.001; \eta^2 = 0.37$). Moreover, participants reported more difficulty with DIR ($2.57 \pm 0.79$) than LRC ($2.17 \pm 0.87$) as indicated by a significant main effect of self-rating type ($F_{(1,44)} = 8.57, p < 0.01; \eta^2 = 0.16$). The interaction between sex and question type did not approach significance ($F_{(1,44)} = 0.30, ns$).

3.5. Relationship between sex, mental rotation performance and LRC

To investigate the relationship between LRC and mental rotation performance in men and women, moderated multiple linear regression analyses were performed for both LRC tasks. The overall

![Fig. 2. Mean LRC rates (%) and standard error means in the matched sample for three different arm positions. The number (0, 1, 2) corresponds to the number of arms crossing the vertical midline of the stickman figure in the back and front view condition of the Bergen Left–Right Discrimination Test for women (white bars) and men (black bars).]
LRC rate was used as dependent variable and the mental rotation score, sex and the interaction (sex \times MRT score) were included in the regression model as predictors.

For the entire sample (91 participants), the moderated multiple regression model was significant for both the Left–Right Commands Task \((F_{(3,87)} = 9.12, R^2 = 0.17, p < 0.001)\) and the Bergen Left–Right Discrimination Test \((F_{(3,87)} = 3.50, R^2 = 0.07, p < 0.05)\). However, only the predictor sex was significant or at least approached significance \((β = −0.28, t = −2.44, p < 0.05; β = −0.23, t = −1.93, p = 0.057, respectively)\), indicating that men had higher LRC rates than women. Neither mental rotation performance \((β = −0.20, β = −0.07, respectively)\) nor the interaction significantly predicted LRC rates \((all p ≥ 0.75)\). For the matched sample (46 participants), the moderated multiple regression models were significant for both the Left–Right Commands Task \((F_{(3,42)} = 3.39, R^2 = 0.14, p < 0.05)\) and the Bergen Left–Right Discrimination Test \((F_{(3,42)} = 4.63, R^2 = 0.18, p < 0.05)\). Again, only sex was a significant predictor \((β = −0.33, t = −2.35, p < 0.05; β = −0.41, t = −2.95, p = 0.01, respectively)\), and neither mental rotation performance \((β = −0.16, β = −0.11, respectively)\) nor the interaction significantly predicted LRC rates \((all p ≥ 0.95)\).

4. Discussion

Previous research suggests that sex differences in LRC tasks are often confounded by the participants’ ability to mentally rotate objects \((Gardner & Potts, 2010; Jordan et al., 2006)\). Given that on average men are known to outperform women on mental rotation \((Linn & Petersen, 1985; Masters & Sanders, 1993; Peters et al., 2006; Voyer et al., 1995)\), it might be assumed that sex differences in LRC are a mere artefact of sex differences in mental rotation \((Jordan et al., 2006)\). The aim of the present study was to investigate directly whether sex differences in LRC still exist when male and female participants were matched for their mental rotation skills as measured by the Revised Vandenberg & Kuse Mental Rotation Test \((Peters et al., 1995)\).

In line with these studies, a sex difference in favour of men was also observed in the Mental Rotation Test in the overall sample of the present study. When men and women were matched for their mental rotation performance, women still believed they are more prone to LRC than men, a finding that is in line with previous research \((Hannay et al., 1990; Hirnstein et al., 2009; Jaspers-Feyer & Peters, 2005; Jordan et al., 2006)\). About 37% of the variance in the results of the self-rating questionnaire was accounted for by the factor sex. The sex difference in self-rated LRC susceptibility corresponds to sex differences in LRC tasks in the present study.

Specifically, the results revealed robust sex differences in the Bergen Left–Right Discrimination Task \((Ofte, 2002)\) and the Left–Right Commands Task \((Hirnstein et al., 2009)\), suggesting that the pronounced sex difference in LRC represents a genuine phenomenon that exists independently of the well known sex difference in mental rotation. This interpretation is further supported by the observation that MRT scores did not significantly predict LRC rates, neither in men nor in women as indicated by the multiple regression analyses. Moreover, the results of the regression analyses suggest that individuals with relatively low mental rotation abilities with respect to their gender do not perform better in LRD, for example, as a compensatory mechanism.

In the Left–Right Commands Task that does not involve mental rotation, women were more susceptible to LRC than men, replicating previous findings by Hirnstein et al. \((2009)\). This finding is also in line with Snyder \((1991)\) who assessed LRC with the Right–Left Orientation Test that requires manual localization of own lateral body parts. However, other studies that used LRC tasks not involving mental rotation did not always find sex differences. These conflicting findings may be explained by differences in task difficulty. For example, in the test used by Jordan et al. \((2006)\), participants had to decide as fast as possible whether an object was to the left or right of another object. The authors state that this rather simple task did not reveal sex differences in LRC, probably as a result of ceiling performance.

In the Bergen Left–Right Discrimination Task women made more left–right errors than men. This is in line with previous findings of the paper-and-pencil version of this task \((Ofte, 2002)\). Moreover, participants needed longer to solve the more complex items in which both arms of the figure crossed the midline than those in which none or just one arm crossed the midline. Initially, we assumed that performance in the Bergen Left–Right Discrimination Task is influenced by mental rotation, since participants may have to mentally rotate the front view figures to make a left–right decision. However, the suggestion that the additional cognitive operations in this test consist of mental rotation \((Ofte & Hugdahl, 2002b)\) has not been tested directly before. Surprisingly, the regression analyses revealed that the mental rotation score did not significantly predict the LRC rate in the Bergen Left–Right Discrimination Task in both the entire and the matched sample. These findings suggest that participants are using other strategies than mentally rotating the front view figures to make a left–right decision. For example, participants could have simply used their knowledge that the left and right sides of the front view figures were opposite to the left and right sides of their own bodies \((Ofte & Hugdahl, 2002b)\).

To disentangle sex differences in LRC from those in mental rotation, Auer et al. \((2008)\) proposed to categorise left–right discriminations into egocentric discriminations regarding the own body and extra-egocentric discriminations regarding objects outside the own body. Egocentric left–right discrimination is conducted from a participant’s own viewpoint and therefore never involves mental rotation. Sex differences in egocentric LRC tasks, like in the Left–Right Commands Task of the present study, therefore exist independently of sex differences in mental rotation. In contrast, sex differences in extra-egocentric left–right discrimination tasks, in which left–right decisions regarding objects outside the own body have to be performed, may be confounded by sex differences in mental rotation. Here, left–right errors can be the consequence of either a failure in egocentric LRC or mental rotation. However, the present study also suggests that extra-egocentric tasks, which appear to involve mentally rotating objects outside one’s own body like the Bergen Left–Right Discrimination Test, do not necessarily require mental rotation. Thus, when claiming that a sex difference in an extra-egocentric task is dependent upon sex differences in mental rotation it is advisable to test whether mental rotation is actually involved.
There is also evidence that left–right discrimination and mental rotation probably involve sex-specific but distinct neuronal mechanisms, further supporting the view that sex differences in LRC exist independently of sex differences in mental rotation. In general, mental rotation is reflected by core activations in the parietal lobe, centred in the intraparietal sulcus (Jordan, Heineze, Peters, & Jäncke, 2001). However, it has been observed that men and women exhibit different patterns of cortical activation during mental rotation. Jordan, Wüstenberg, Heinze, Peters, and Jäncke (2002) found that both men and women showed activation in premotor areas, but only women exhibited significant bilateral activation in the intraparietal sulcus, the superior and inferior parietal lobule and the inferior temporal gyrus. In contrast, men revealed significant activation in the right parieto-occipital sulcus, the left intraparietal sulcus, the left superior parietal lobule and the left motor cortex. The authors assume that these differences in brain activity reflect different cognitive strategies in solving mental rotation tasks, an assumption that has also been corroborated by others e.g., Hugdahl, Thomsen, & Erslund, 2006).

Much less is known about the neuronal mechanisms of sex differences in LRC. An early study that addressed this topic measured cortex. The authors assume that these differences in brain activity cannot be unequivocally attributed to either left–right discrimination or mental rotation. A less confounded fMRI study (Auer et al., 2008) used an egocentric LRC task in which male participants had to show numbers with the fingers of either the left or the right hand. Brain activations during this task specific to left–right discrimination (and not involved in mental rotation) were found in the right hemisphere in the medial and middle frontal gyrus, the precuneus, the postcentral gyrus, the angular gyrus, the lingual gyrus, and the superior temporal gyrus. In the left hemisphere, specific activations were found in the superior and middle temporal gyrus and in the precuneus. Female participants were not included in this study. In a recent transcranial magnetic stimulation (TMS) study (Hirnstein, Bayer, Ellison, & Hausmann, submitted for publication), participants had to decide whether the left or right hand of a stick-man figure was highlighted in red (Bergen Lateral Discrimination Test, Ofte, 2002). A significant impairment in this LRC task was observed only after stimulation of the left angular gyrus, not the right angular gyrus. This finding was present for men and women. Previous studies have also shown that particularly the left hemisphere is involved in LRC (Gold, Adair, Jacobs, & Heilman, 1995; Hannay et al., 1983; Hirnstein et al., 2009). The left angular gyrus is involved in semantic processing of language and assessing the meaning of words (Price, 1998). Sholl and Egget (1981) concluded that LRC arises from erroneous verbal labelling and not erroneous spatial/perceptual encoding. This might suggest that LRC is a verbal labelling problem rather than a deficit in mental rotation ability and spatial orientation in general. However, whether this also accounts for sex differences in LRC remains unclear. It seems likely that individual differences in verbal labelling play a crucial role in determining individual differences in the ability to discriminate left from right.

References