Frontiers in Behavioral Neuroscience

Research Article
12 June 2014

Article Title

First Author [1,*], Co-Author [2] and Co-Author [2]

[1] *Laboratory X, Institute X, Department X, Organization X, City X , State XX (only USA, Canada and Australia), Country X*

[2] *Laboratory X, Institute X, Department X, Organization X, City X , State XX (only USA, Canada and Australia), Country X*

Correspondence*:
corresponding Author
Laboratory X, Institute X, Department X, Organization X, Street X, City X , State XX (only USA, Canada and Australia), Zip Code, X Country X, email@uni.edu

ABSTRACT

Refer to `http://www.frontiersin.org/Behavioral Neuroscience/authorguidelines` or **Table??** for abstract requirement and length according to article type.

Keywords: Text Text Text Text Text Text Text Text

1 INTRODUCTION

Spatial navigation: We are able to navigate and orient ourselves effortlessly through the world. Yet, when we put ourselves in a virtual world navigation becomes cognitively demanding. Why the discrepancy? Normally we rely on vision, audition, vestibular, and proprioceptive input to automatically guide us and update our position. The sensory information from all senses is automatically combined into a spatial representation in the brain involving a wide network of brain regions (for a review see **?**). Noteworthy is the existence of different reference frames for spatial orientations that seem to be processed in distinct neural correlates [**??**]. There are two distinct reference frames we use: egocentric, self-to-object representation, and allocentric, object-to-object representation. While the egocentric reference frame uses the individual itself as reference point, an allocentric representation is independent of the observer and is aligned to objects in the outside world. For example, the allocentric coordinate system of cardinal directions is aligned to the magnetic north pole [**?**]. While we are almost forced to exclusively use our egocentric reference to navigate in the real world, when we imagine the same path we tend to use a mixed strategy to determine where we are. Forming and maintaining spatial representations most of the time takes no conscious effort - it is in fact automatic and often even obligatory, meaning that it is hard to suppress and ignoring it takes conscious cognitive effort [**?**]. During navigation, spatial representations are not only constantly updated and maintained in parallel but also interact [**?**]. When exactly we use a specific reference frame for a certain task remains a difficult question because individual proclivities come into play [**?**]. Spatial navigation is a deep rooted and modularized cognitive skill based on spatial representations that are automatically formed and maintained (updated) in specialized brain areas based on multimodal sensory information. However, there are times when spatial updating fails, especially when we receive incomplete or contradicting sensory information. In such cases, we revert to so called offline strategies where we try to cognitively restore our spatial representations. As inconvenient as those cases may be for the individual, they enable researchers to study the mechanism of spatial updating in more detail: when is spatial updating automatic and obligatory, when does it brake down? What factors decide which reference frame we use for our spatial representation?

Spatial strategies: Researchers have discovered a phenomenon involving spatial updating and spatial representations in different reference frames (Klatzky et al., 1998) where, in a point-to-origin paradigm, participants experienced a virtual visual flow environment. Here, the turner group used an egocentric reference frame that was updated during the trajectory and the non-turner group responded as if they were still facing the original direction they started. However, those using the non-turner strategy solved the task correctly based on their strategy, applying an allocentric reference frame that stays constant throughout the trajectory.

VR/VE: The advent of virtual reality (VR) technology gave researchers the opportunity to create experiments where the availability and fidelity of visual, vestibular, proprioceptive and auditiv information channels can be controlled separately in a highly controlled way. This enables to systematically dissociate the different influences of the modalities on complex tasks like spatial navigation.

Individual factors and related works: Previous studies have looked at the individual factors that may influence the strategy used for spatial updating in a virtual point-to-origin task, such as gender, video gaming experience, ethnicity, response mode, navigation skills, cardinal direction proficiency, and decision certainty (Goeke et al., 2013; Sproll, 2013; Avraamides et al., 2004; Riecke, 2008).

Gramann hypothesised that non-turners responded as if they had not turned and were still facing the original direction because they solved the task in a more abstract, disembodied way applying an allocentric reference frame that stays constant during the passage. Thus, what was thought of as an error solving the task turned out to be a different strategy of solving the task where the answer is expressed in a different reference frame.
Avraamides and colleagues showed in [?] that an increased error (corresponding to non-turner behaviour) did not arise when participants performed an imagined triangle completion task and answered using spatial language instead of pointing. Thus they concluded that the non-turner answers in the pointing condition is due to the strong attachment of the pointing gesture to the current perceived body position (that is aligned with the hypothetical allocentric reference frame).
This hypothesis is notably different from the one used by Gramann. While they agree that participants giving turner answers update their egocentric reference frame according to the given stimulus (imaginary walking, visual flow, etc) they have different explanations for the non-turner answers. Whereas Gramann explains non-turner behaviour as a different strategy of solving the task using an allocentric reference frame, Avraamides sees non-turner answers as an artefact of the task, namely the conflict between a virtual body orientation and a physical body orientation. Here non-turner answers are not valid answers in an allocentric reference frame but errors due to an overriding of the virtual egocentric reference frame with a physical egocentric reference frame. He found that this conflict is not present when spatial language is used to give the answers. Avraamides explains this with more abstract and less embodied nature of spatial language compared to bodily pointing. In this it might be closer to a more cognitive representation of heading.
To enable a neutral discussion of the phenomenon we will use the terms turner and non-turner in the this study, referring only to behavioural observation whether participants incorporated the virtual turn in their response or not without making an implicit assumption which reference frame they use.
Several further studies (see table **??**) have investigated what factors determine the strategy selection in the individual but still no coherent picture has emerged. While individual proclivities seem to have a significant influence on strategy selection [?], we can again observe similar influences as for automatic spatial updating, e.g. a more prominent use of a turner strategy in studies with naturalistic scenes and vestibular input [?]. The first big cross-sectional study investigating the turner non-turner phenomenon was an online study conducted by Göcke and colleagues [?]. Their sample contained (after preprocessing) 260 participants from 15 countries, although the majority were from Spain and Germany. The task did not only contain left right (yaw) turns but also up and down turns (pitch). Answers were given via selecting one of four 3D arrows. In their analysis they found the factors gender, cardinal direction proficiency and decision certainty to be significant factors determining turner / non-turner behaviour, while this was not

the case for self-estimated general navigation skills or video gaming experience. However, it seems that a multitude of known and unknown factors influence the strategy use, leading to partially widely varying ratios of turners to non-turners in different studies . A large body of literature has well established the link between culture and cognitive style (Kitayama, Duffy and Uchida, 2007; Kitayama et al., 2009; Norenzayan, Choi and Peng, 2007; Varnum et al., 2008). Western cultures, such as the United States, tend to exhibit a more independent and analytic social orientation: emphasizing uniqueness, having relatively low sensitivity to social cues, and encouraging behaviours that affirm autonomy. On the other hand, other cultures such as China tend to exhibit a more interdependent and holistic social orientation: emphasizing harmonious relations with others, promoting sensitivity to social cues, and encouraging behaviours that affirm relatedness to others (Kitayama et al., 2007; Varnum et al., 2010). On the basis of such evidence, the link between social orientation and cognitive style has been widely accepted. Goeke, Koenig, and Gramman (2013) suggest looking at cultural background on reference frame proclivity to finally unravel the underlying factors determining human navigation strategies.

1.1 GOALS OF THE PRESENT STUDY

We conducted a simple point to origin task in lecture halls, thereby getting a classification of a very large number of subjects together together with demographic information. The goals were:

- replicate the gender bias found by Göcke et al. in [**?**]. We hypothesise based on the literature that females are more likely to be non-turners compared to males.
- extend the findings of **?**, predicting a higher amount of turners when spatial language instead of pointing is used, to the use written spatial language vs. pictograms
- investigate a possible influence of ethnicity on strategy selection

To answer these questions we designed our study with the main goal of having a very large sample size to cope with intrinsically noisy strategy classification data and the high individual differences. We settled on a design that could be executed with many participants simultaneously showing the stimulus on a projector and recording the answers via a paper questionnaire. This way, we were able to perform the experiment in lecture halls at the beginning of regular courses. We chose a small number of trials since earlier studies have shown that strategies are relatively stable over time [**?**].
As a consequence of the study design, we could not directly employ the same answering modes as in **?**. We instead used pictograms as the more embodied version while using answering in written spatial language as the equivalent of description on spatial language (see Fig. 3 B). We are aware that those answering modes are somewhat more abstract that the ones used by Avraamides and thus expect weaker effects.

2 MATERIAL & METHODS

2.1 PARTICIPANTS

A total of 507 participants took part in the study, 228 female, 273 male and 6 NA. Participants with missing gender and/or ethnicity data were cut out (n = 6).The average age was 20.5 years (SD = 3.2). We recruited a quite diverse spectrum of participants from 3 universities: the Simon-Fraser University (244 participants) and the University of British Columbia (183 participants), both in Vancouver, Canada and the University of Osnabrück in Germany (104 participants). An effort was made to recruit a sample with high ethnic diversity as can be seen in Fig. 1. Participants were not reimbursed.

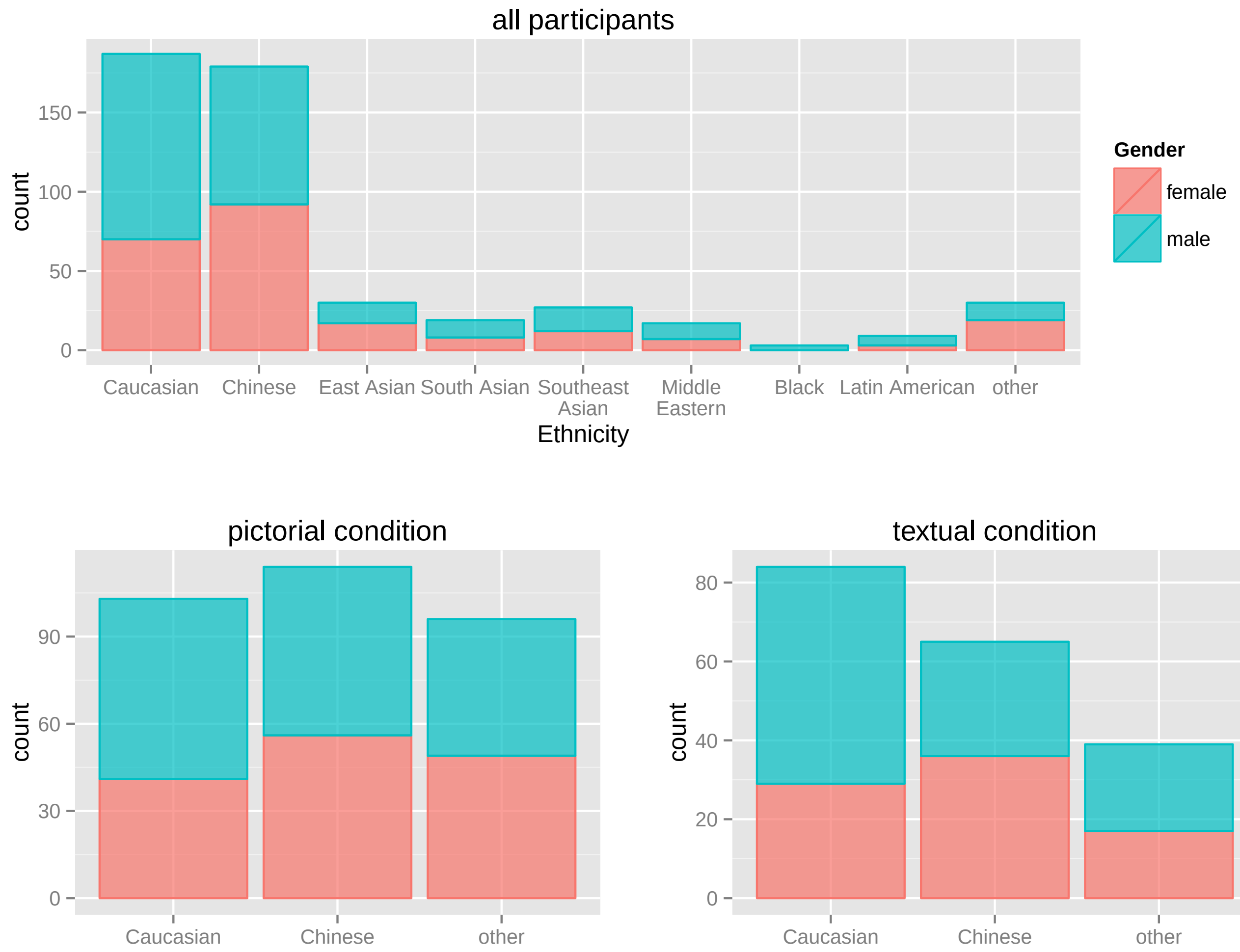

Figure 1: Demographics of the participants. The two main groups are Caucasian and Chinese, all other Ethnicities were pooled into a third group. While two thirds of the Caucasian participants were male, for all other groups the male female ratio was one to one. This distribution is also reflected in the allocation to the two conditions (lower two plots)

2.2 STIMULUS & APPARATUS

The stimulus shown to the participants was a passage through a virtual starfield providing optical flow without any landmarks. The trajectories consisted of an initial straight part, followed by a curve and a second straight part at the end. The curve angles used for the four trials were $60^\circ\ left$, $90^\circ\ right$, $90^\circ\ right$ and $60^\circ\ left$ in this order (paths are illustrated in Fig. 2). The velocity profile was smoothed to make the stimulus less artificial and prevent nausea. The first linear part included a $1s$ linear acceleration phase with $10\frac{m}{s^2}$, followed by a constant movement with $10\frac{m}{s}$ for $2s$. The turn was divided into an accelerating half and a decelerating half, the constant acceleration being $15\frac{\circ}{s^2}$, resulting in an overall turn time of $4s$ for 60° and $5s$ for 90°. The second linear part consisted of a $3s$ constant linear movement and $1s$ deceleration, thus being slightly longer than the first part.

However, it should be noted that velocities and distances are quite abstract in a starfield environment and the subjective perception highly depends on the starfield parameters chosen (star size, area and visibility

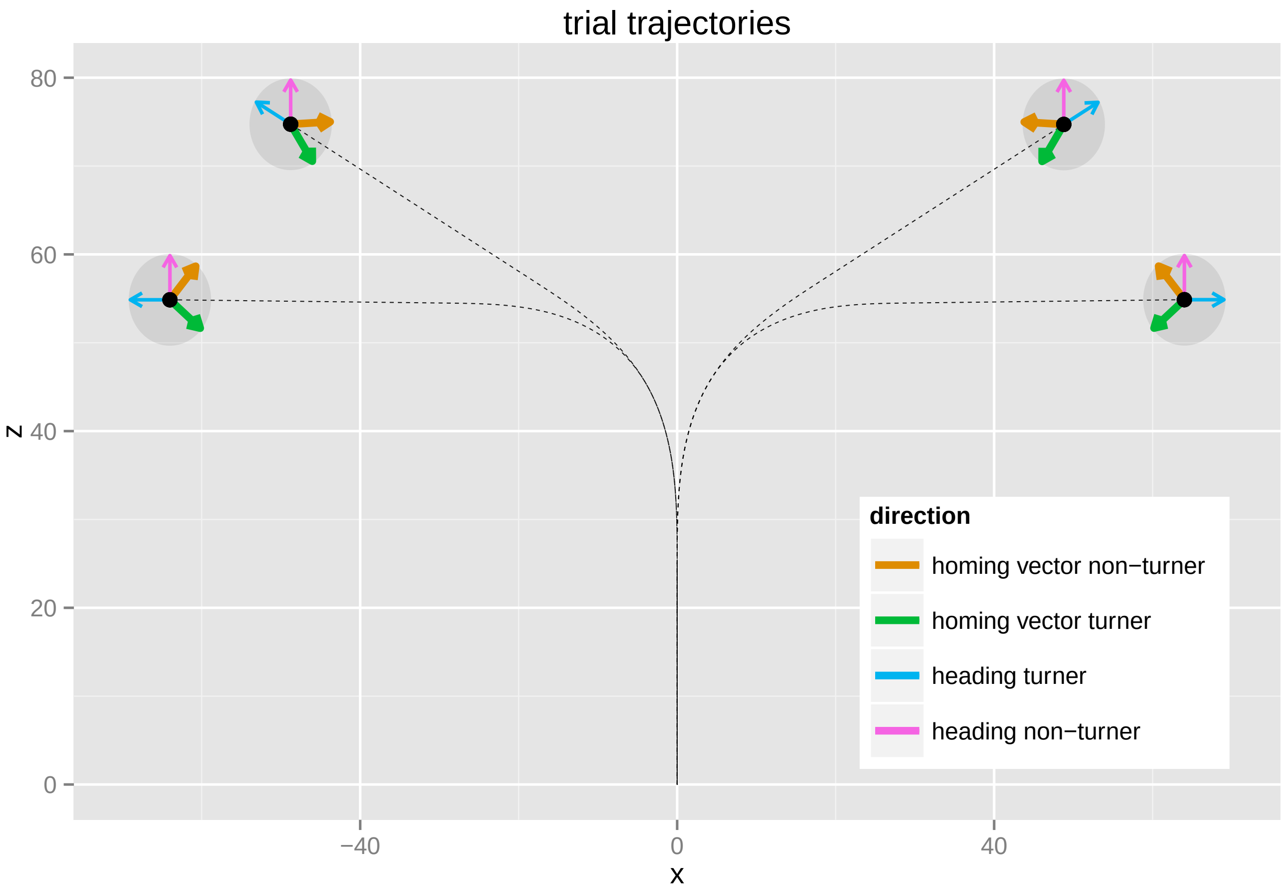

Figure 2: The trajectories of the four trials from birds-eye perspective. Thin arrows are the heading at the end of the trajectory while the thick arrows are the egocentric and allocentric homing vectors. X and Z axis are the displacement in the plane in meters.

range). The passages were programmed using *Vizard 4.0*. The code for the pre study can be found online (`http://github.com/leftbigtoe/starfield`) and can be executed with the free trial version of *Vizard 4.0*.

The answers were given via a multiple choice questionnaire (see appendix **??**). For each trial of the point to origin task the same four possible answers could be selected: front left, front right, back left, back right for the textual condition and the corresponding pictograms for the pictorial condition. The sequence of the items for each trial was randomized to avoid answering tendencies. The questionnaire was folded and sealed with tape, the part for assessing the demographic information being inside to prevent possible bias of the task performance by the demographic questions.
The stimulus was shown on the projectors available in the classrooms. Lights were dimmed where possible. Students were asked to group as closely as possible around the projector to minimize extreme viewing angles.

2.3 PROCEDURE

The experiment took place at the beginning or at the end of the classes. The experimenter was introduced by the lecturer, then the informed consent form was distributed and read by the students. All students

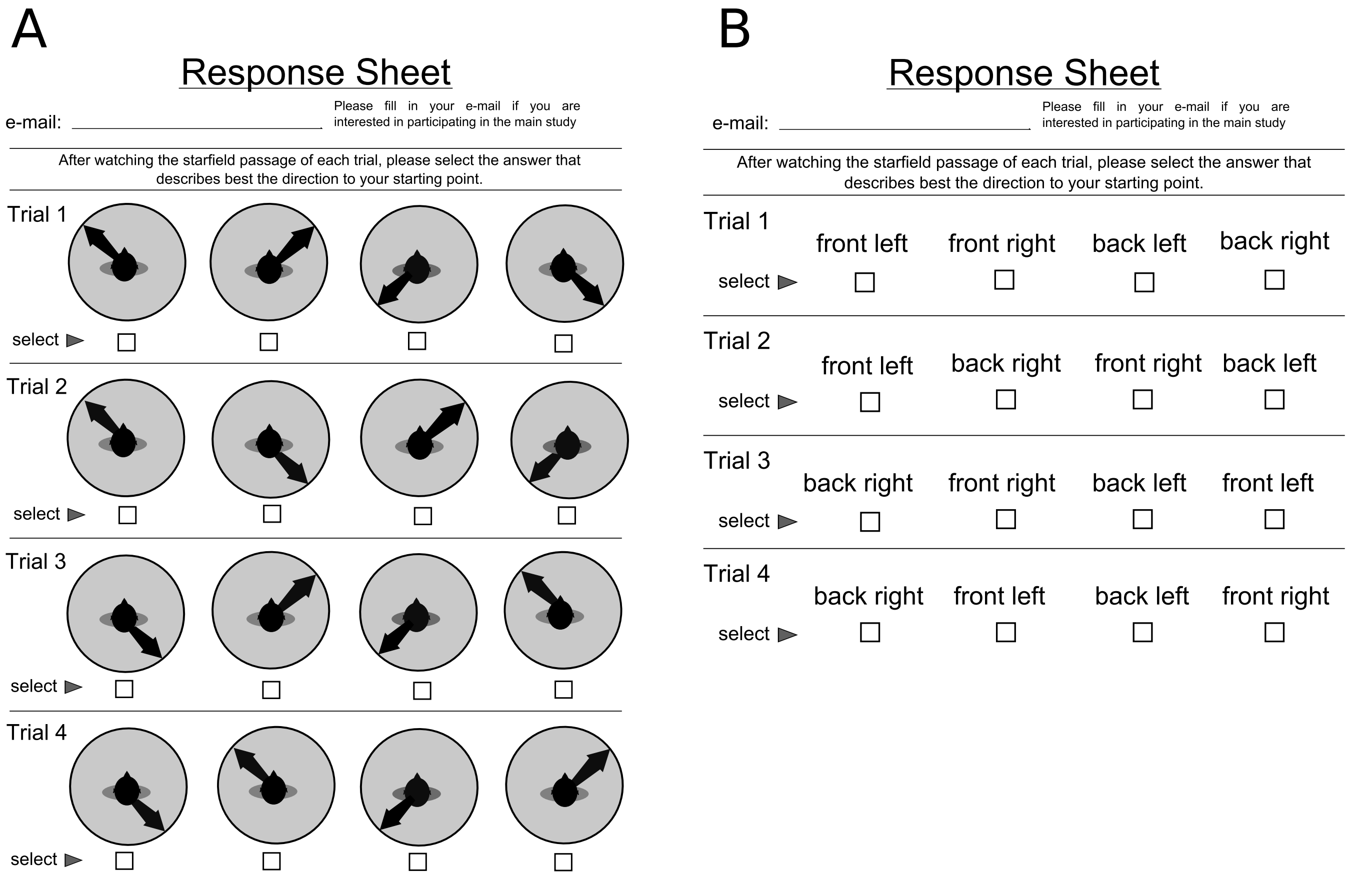

Figure 3: **A** Questionnaire for the pictorial condition **B** Questionnaire for the text condition

volunteering to participate in the study signed the consent form and were randomly handed a pictorial of text condition questionnaire. The experimenter then explained the task until no subject had further questions. The participants were asked to select the answers as quickly and intuitively as possible and not to perform mental arithmetic or similar strategies. They were also asked not to copy from their neighbours or discuss their answers until after the experiment. The trials were shown to the class, pausing after each trial until everybody was finished. No questions that could provide feedback were answered. After completing the task, the room was illuminated again and the participants asked to open their sheets and fill out the demographics questionnaire. The experiment took approximately 10 minutes.

2.4 PREPROCESSING

Before the analysis, the following preprocessing was performed on the collected data. Only participants that provided data for ethnicity and gender and had no missing answers for the navigation task were used ($n = 6$ participants excluded). For each trial the strategy used was classified (turner, non-turner, frontal pointing 1 or frontal pointing 2). In accordance with previous studies (e.g. **?**) we classified participants with consistent strategy use in 75% of the trials as users of the respective strategy. All others were classified with no preference. Only three participants were classified as frontal pointing 2 users and as no explanation could be given to this answering pattern (as discussed above), those answering patterns were considered to be due to inattentiveness. We thus excluded participants classified as frontal pointers 2 from further analysis due to sparseness of data ($n = 3$ participants excluded). Statistical analysis was performed with the remaining $n = 498$ participants.

2.5 DATA ANALYSIS

R 2.15.2 was used for data analysis. The multinomial regression model used for statistical analysis was the `multinom` implementation of the `nnet` package. The likelihood ratio test of the parameters was done using the `Anova` function of the `car` package.

3 RESULTS & DISCUSSION

3.1 GENERAL RESPONSE BEHAVIOUR

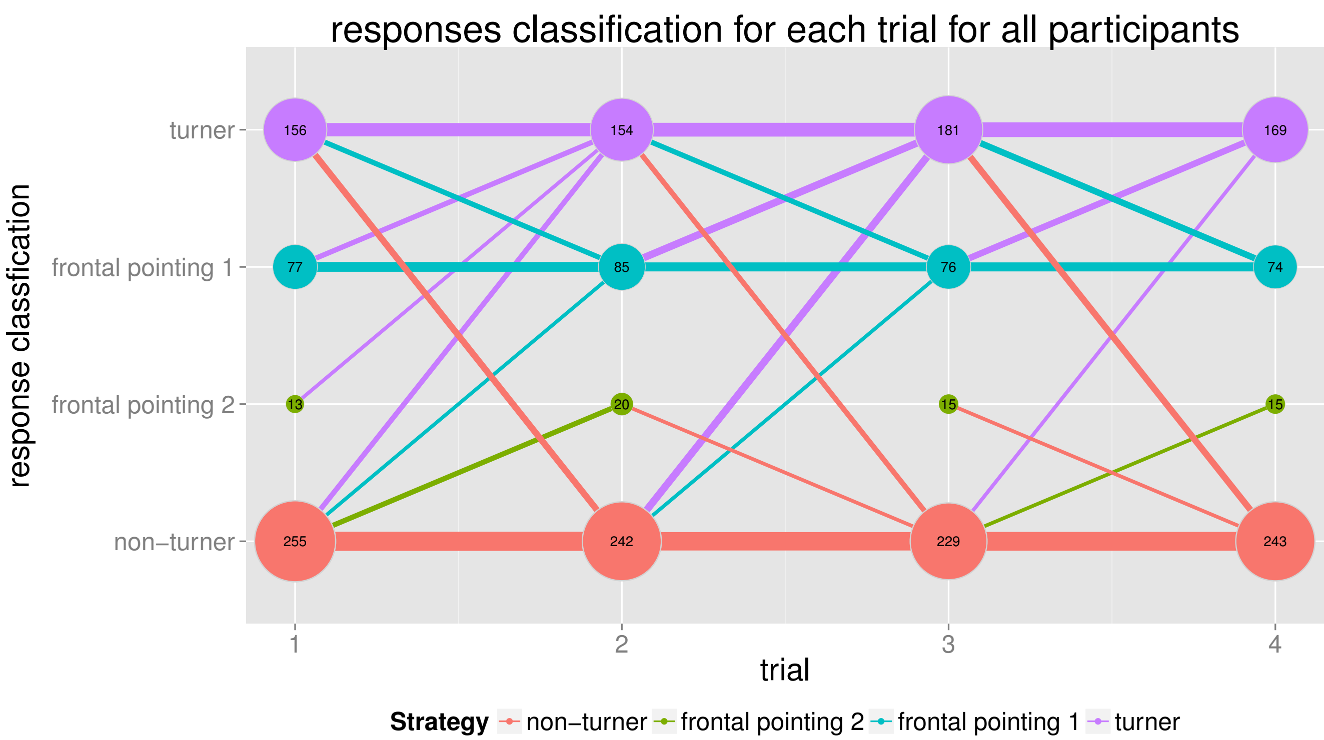

Figure 4: Total counts of answering types per trial. Y position and colour of the dots indicate the type of the answer, x position the trial and area of the dot corresponds to the count, also given by the number within the dot. The bars indicate how many changed from giving one answer type in a previous trial to which answer type in the next trial, e.g. a bar from frontal pointing 1 in trial 1 to turner indicates the amount of participants that changed from giving a frontal pointing 1 response in the first trial to a turner answer in trial 2. Thickness again stands for amount of people changing in this way. A cutoff of $n > 5$ for the bars was chosen to only show stable trends. Strategies are relatively stable. The turner strategy draws the most participants over time from all other strategies and is the only strategy that is growing overall while frontal pointing 2 is the most isolated. The interaction between frontal pointing 1 is highest with the turner answers, giving more evidence that frontal pointing one might be turners overestimating the turn. Non-turner interacts moderatly, mainly with the turner answers and the frontal pointing 2 answers

Looking at the total counts of responses over the trials (see Fig 4) shows relatively stable strategies, the two most prominent being non-turner answers (48.35%) and turner answers (32.93%). A smaller amount of participants gave frontal pointing responses, mainly frontal pointings 1 in the direction of the turn (15.57%). Only very few frontal pointings 2 in the opposing direction of the turn were given (3.14%. While non-turner and turner answers were correct and expected, both types of frontal pointings were only thought to be distractors. They were not correct in either reference frame, however a frontal pointing in the direction of the turn (frontal pointing 1) could be explained in two possible ways. First, by a turner that overestimated the turn (somewhere over $135°$). In this case the starting point would be in the frontal hemisphere. The second explanation could be a misunderstood task in which participants pointed from the starting to the end point. This was reported by a few participants after the experiment. For a frontal pointing in opposite direction of the turn (frontal pointing 2) in contrast no possible explanation could be found. We therefore assumed them to be simply a wrong answer due to inattentiveness or distraction. This is supported by the fact that it does not seem to be a very stable strategy: while 36 people (7.19%) gave a frontal pointing 2 once, only 8 (1.6%)gave it more than once and 3 more than twice (0.6%).

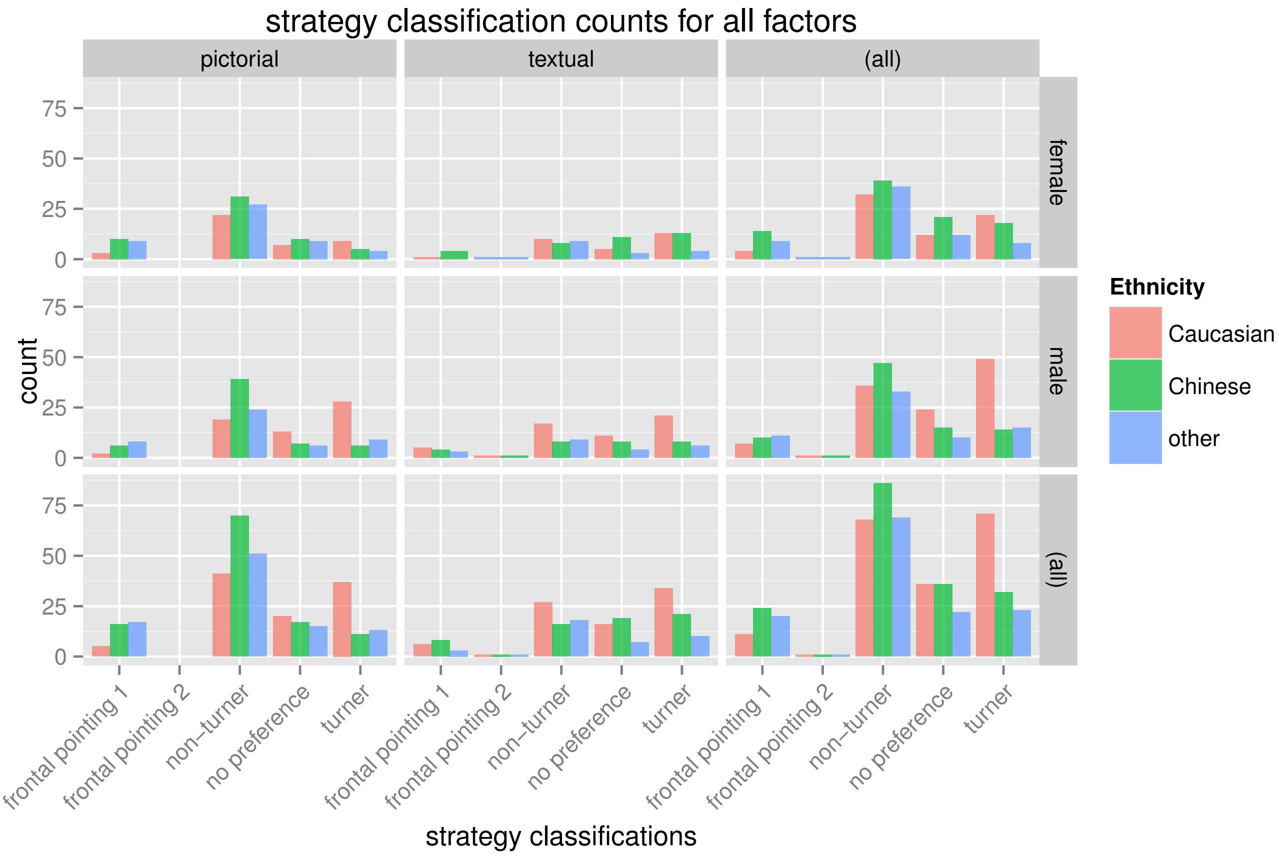

Figure 5: Total counts of preferred strategy classifications factored out into each of the model factors condition, ethnicity and gender and respective marginal sums. It can be seen that the two most dominant classifications were turner and non-turner followed by no preference, while the frontal pointing classifications, especially frontal pointing 2, were quite rare.

The overall counts of classification according to the 75% criterion (participants that used the same strategy in 75% of the trials) can be seen in Fig. 5. As expected, the two most prominent classifications were non-turner (44.78%) and turner (25.3%). 11.04% were classified as frontal pointing 1 users and only 0.6% had frontal pointing 2 as their preferred strategy. 18.88% of the participants did not show a clear preferred strategy and thus were classified as with no preference. Obvious already in this overview is the high amount of non-turners in the pictorial condition compared to the text condition and the high amount of male Caucasian turners, especially in the pictorial condition.

3.2 MULTINOMIAL REGRESSION MODEL

For statistical analysis a multinomial regression model was fitted We included the factors condition, ethnicity, gender and all interaction terms to model the preferred strategy. Accuracy of the model on the training data was 49.0% compared to 25% chance level. The precise parameter values can be found in Table **??**

	non-turner		turner		frontal pointing 1	
Parameter	Estimate	SE	Estimate	SE	Estimate	SE
(Intercept)	1.15	0.434	0.251	0.504	-0.847	0.69
EthnicityChinese	-0.0136	0.566	-0.944	0.744	0.847	0.822
EthnicityOther	-0.0469	0.58	-1.06	0.784	0.847	0.836
ConditionText	-0.452	0.699	0.704	0.729	-0.763	1.29
GenderMale	-0.766	0.564	0.516	0.605	-1.02	1.03
EthnicityChinese:ConditionText	-0.998	0.915	0.156	0.999	-0.249	1.49
EthnicityOther:ConditionText	0.453	1.04	0.396	1.21	-12	0.66
EthnicityChinese:GenderMale	1.35	0.787	0.0226	0.988	0.87	1.25
EthnicityOther:GenderMale	1.05	0.821	0.701	1	1.31	1.25
ConditionText:GenderMale	0.508	0.876	-0.824	0.884	1.85	1.6
EthnicityChinese:ConditionText:GenderMale	-0.775	1.24	0.119	1.35	-1.37	1.94
EthnicityOther:ConditionText:GenderMale	-1.08	1.39	-0.276	1.56	10.4	0.66

Likelihood ratio tests on the regression parameters revealed that the parameters Ethnicity ($p_{chi^2} < 0.001$)and Condition ($p_{chi^2} < 0.001$) were highly significant. Further, the interaction terms ethnicity & condition and condition & gender were found to be mildly significant ($p_{chi2} < 0.05$). In contrast to earlier studies [**?**], gender was not found to be significant at all. For an overview see Table **??**.

Parameter	LR chi^2	df	p_{chi^2}	
Ethnicity	26.8880	6	0.0001520	***
Condition	17.9785	3	0.0004444	***
Gender	2.1589	3	0.5400950	
Ethnicity:Condition	14.3335	6	0.0261252	*
Ethnicity:Gender	5.9970	6	0.4235304	
Condition:Gender	7.9853	3	0.0463172	*
Ethnicity:Condition:Gender	2.8220	6	0.8308366	

3.3 BOOTSTRAP CONFIDENCE INTERVALS FOR MODEL PERFORMANCE

To be able to further judge the accuracy a bootstrap analysis was conducted. For a review on bootstrap methods see [**?**]. Two kinds of bootstrap models were created: a naive one creating random classifications for every participant with uniform probability and one were the probability of the classifications were weighted based on the observed strategy counts. 10000 random classifications were created for each models and the confidence intervals calculated. The accuracy of our model lay outside of both bootstrap confidence intervals (naive: 23.5% - 28.7%, weighted: 29.7% - 35%) indicating a decent fit. A further interesting observation was, that the model only made two classifications, namely non-turner or turner but never frontal pointing 1 or no preference. This inability of the model to discriminate between all four strategies and the emergence of turner and non-turner as main strategies indicates some correlation between some of the strategies. No preference and frontal pointing 1 seem both the be correlated to one of those main strategies instead of being independent strategies. However, also the fact that there is more training data for the turner and non-turner classifications has to be taken into account, possibly also accounting for at least some of the bias of the model.

3.4 ODD RATIOS

From the regression parameters of the multinomial regression model, we directly calculated the odd ratios (ORs) for more detailed interpretation of the results. Odd ratios quantify the correlation of two variable appearing together. They are calculated by dividing the number of occurrences that a participant has a given b (the odds of a given b) divided by the number of occurrences of a given not b. An OR greater 1 shows a positive correlation of a with b while an OR smaller one indicates a negative correlation. ORs equal 1 mean no correlation.
In a In multinomial regression models parameters with more than two factors are dummy coded as dichotomous variables and comparisons are always performed by using one of two possible values for a factor as baseline and comparing it against the other value. To capture all effects, a script was written that created a model for every possible combination of base cases and extracted all significant odd ratios (Wald confidence intervals that did not contain 1). Note that changing the baseline values does not change the overall performance of the model, it rather "phrases the result in a different way". Due to the dichotomous dummy coding there is also a mirror symmetry among the reported effects (e.g. OR text makes turner instead of non-turner more likely = OR pictorial makes non-turner instead of turner more likely). This symmetry is also nicely visible in the plots. We decided to still report both ways to avoid introducing a bias by leaving to much implicit. In the next step, all odd ratios with values under 0.001 and over 100 were excluded. Those ORs were highly likely to be artefacts of sparse data, having huge confidence intervals indicating their unreliability. In the following, only ORs greater than one will be shown. Due to the dichotomous dummy coding of parameters, every effect indicating x to be less likely for a certain parameter having value b also means x is more likely if that parameter has its other possible value a. To avoid redundancy we will only present ORs greater than one (more likely). The ORs are plotted in Fig. 6 and the exact values and confidence intervals can be found in the appendix **??**

Ethnicity: (see Fig. 6 A) All Ethnicity effects were only with the pictorial condition as baseline. Chinese and Other were more likely to be frontal pointers 1 instead of turners compared to male Caucasians (Chin. OR: 14, Other OR12.45) and female Caucasians (Chin. OR: 6, Other OR 6.75). Further, compared to male Caucasians, Other were more likely to be non-turners instead of turners (OR:3.93). Chinese males were non-turners instead of no preference (OR:3.81) or turners (OR: 9.58). Vice versa, Caucasians were more likely to be turners instead of front pointers 1 compared to Chinese (male OR: 13.99, female: 6) or in the Other category (male OR: 12.45, female OR:6.75). Male Caucasians were also more likely to have no preference (OR: 3.81) or to be turners (OR: 9.58) instead of non-turners compared to Chinese. Last male Caucasians were more likely to be turners instead of non-turners (OR: 3.93) or to have no preference instead of being frontal pointers 1 (OR: 8.67) compared to males in the Other group.
The effects of ethnicity again seem to be more pronounced when the male baseline is used, possibly explained by the extreme amount of male Caucasian turners. Another noteworthy observation is that there no significant difference between the Chinese and Other group and their comparisons against the Caucasian group are quite similar. This can be interpreted in two ways: either a high similarity between the Chinese and Other groups or that Caucasians are quite unusual in their navigation behaviour compared to other ethnicities. It seems unlikely that the differences might be mediated by a difference in video gaming or navigation skills, since both were not significantly different in both groups as revealed by a Kruskal Wallis Test (self rated navigation skills $H = 0.17, df = 1, p = 0.68$ and gaming $H = 0.82, df = 1, p = 0.37$).

Condition: (see Fig. 6 B) While a significant effect of the condition for Caucasians can only be observed among females (OR: 3.18), a significant effect is present for both sexes among Chinese participants (male: 6.5, female: 10.07). In both cases, the pictorial condition makes a non-turner strategy more likely

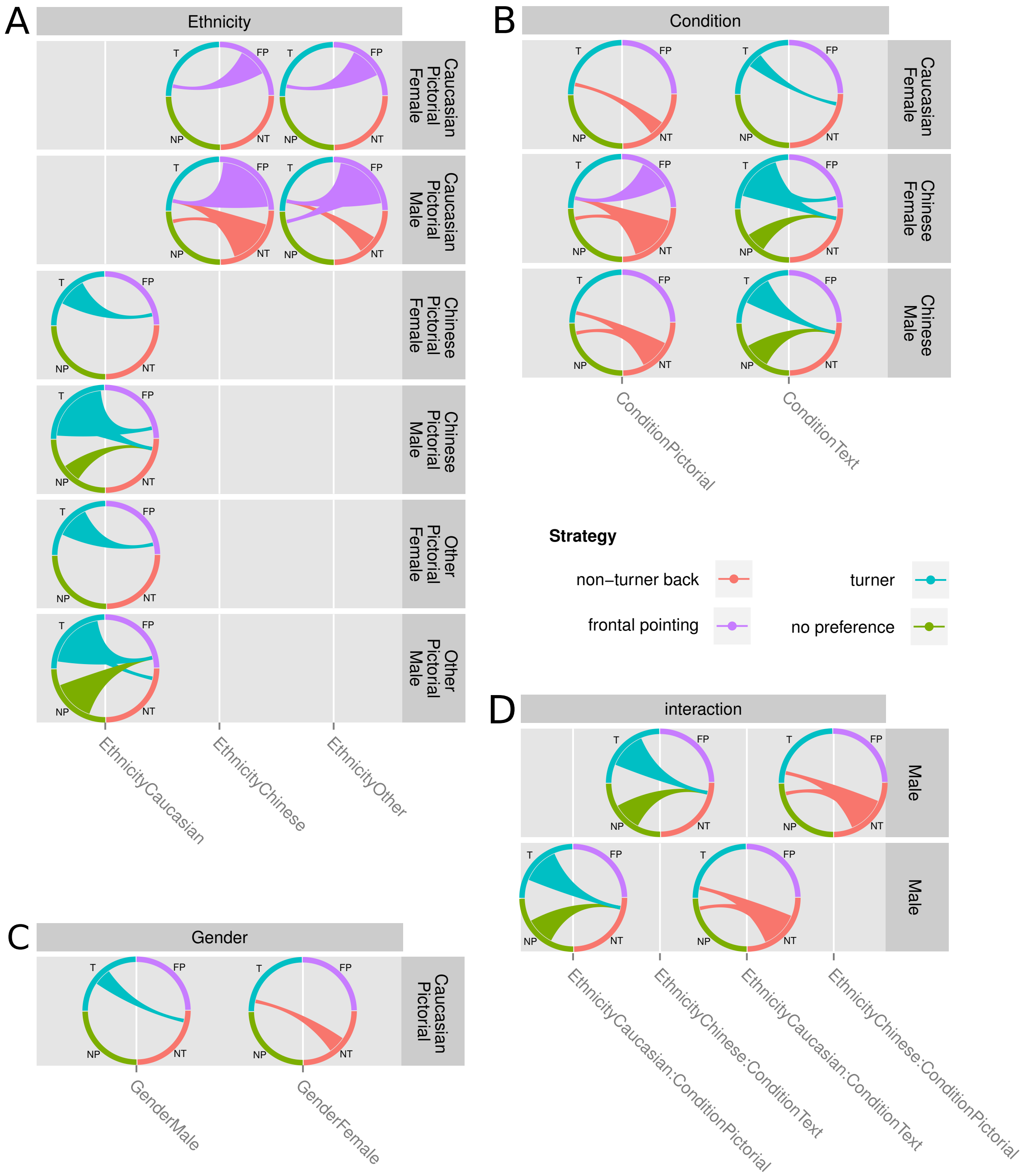

Figure 6: Significant and reasonable odd ratios. Each chord marks a significant comparison. The thin end is the baseline strategy, the thick end the strategy that is more likely instead of the baseline. Example left circle of **C**: for Caucasians in the pictorial condition being male means a classification as turner is significantly more likely than being a non-turner compared to being female (3.6 times more likely, see last row of **??** for exact value).
A: The effect of condition was significant for female Caucasians and both genders among Chinese participants. They were more likely to be non-turners or frontal pointers in the pictorial condition and turners or have no preference in the text condition.
B: Gender related ORs were only significant for Caucasians in the pictorial condition. Males were more likely to be turners while females were more likely to be non-turners.
C: All effects for Ethnicity only emerged in comparison to a pictorial baseline. Here Chinese and Other were more likely to be frontal pointers (men and women) or non-turners (only males). Vice versa, Caucasians were more likely to be turners compared to Chinese and Other, while having no preference was also more likely but only for males.
D: The interaction terms go into a similar direction than before, showing an opposing trend: while Caucasians are turners or have no preference in the pictorial condition where Chinese are more likely to be non-turners, this reverses for both ethnicities in the text condition. Here the effects only appear compared to a male baseline.

compared to a turner strategy. For Chinese participants a non-turner strategy is also more likely compared to a no preference strategy (male OR: 5.57, female OR:4.26). Among female Chinese subjects a front pointing 1 strategy becomes also more likely (OR: 6.5). On the other hand, the text condition has the opposite effect, rendering a turner strategy more likely in the same groups: Chinese males and females are now turners instead of non-turners (male OR: 6.5, female OR: 10.8 and have no preference instead of non-turner (male OR: 5.57, female OR: 4.26). Chinese females were also more likely to be turners instead of frontal pointers 1 (OR: 6.5). Only effect for Caucasians was again among females, an OR of 3.18 for being turner instead of non-turner. Among the other group, no significant effects for condition emerged. Effects are stronger compared to a non-turner strategy as baseline.
We did replicate the results of Avraamides and colleagues [**?**], showing that the use of spatial language indeed makes turner responses more likely. Moreover we could extend the findings, showing the effect also remains present for simple multiple choice response sheets using more the abstract pictograms and written spatial language for indicating the direction of origin. Interestingly, this effect is not significant in male Caucasians which could be due to already quite high amount of turners in this group in the pictorial condition. That there was no effect within the Other group might be due to the heterogeneous composition of different ethnicities within this group, averaging out any effects.

Gender: (see Fig. 6 C) Gender effects only emerged among the Caucasian group with the pictorial condition as baseline. Here males were more likely to use a turner strategy (OR: 3.6) while females tended more towards a non-turner strategy (OR: 3.6). In addition, a few implicit gender effects emerged, like the stronger difference between male Caucasians and male Chinese participants compared to their female counterparts.
Against our expectations, females were not in general more likely to be non-turners than males, contradicting the results of [**?**]. Gender was not found to be a significant model parameter, it only turned out to be significant within the interaction term of the model. Examining further, we found that the only significant OR for gender was found in comparison to the Caucasian / Pictorial baseline. All in all, our results suggest that the gender effect as found in **?**, where most participants were from Germany and Spain, could be an artefact of a very specific task and sample instead of a general bias in reference frame use.

Interactions (see Fig. 6 D) Only the interaction between Ethnicity and Condition yielded some significant ORs. The interaction again emphasized effects already seen before: in the pictorial condition Caucasians are more likely to be turners (OR: 7.75) or have no preference (OR: 5.89) both compared to a being a non-turner. The same holds for Chinese in the text condition where they are also more likely to be turners (OR: 7.75) or have no preference (OR: 5.89). Consequently, male Caucasians are more likely to be non-turners in the text condition (OR against no pref.: 5.89, OR against turner: 7.75) while the higher likelihood of a non-turner classification for Chinese males was found for the pictorial condition (same ORs). The interaction effects show common directions instead of influences of single parameters for given baselines. Thus, Chinese & text push in the same direction as Caucasian & pictorial, namely towards a turner or no preference strategy, while Chinese & Pictorial as well as Caucasian & text push in the other direction towards a non-turner strategy.

Another interesting observation is that the effects seem to group in a way that two strategies are likely to appear together, namely turner and no preference on the one side and non-turner and frontal pointing 1 on the other. This connects to the emergence of turner and non-turner as main classifications of the model and its inability to make frontal pointing 1 or no preference classifications observed above. Although the two correlating classifications don't always appear together, they never appear in different combinations. This fact was also reflected by the classification behaviour of the model, that classified the data only into turner or non-turner but never in no preference or frontal pointing 1. While 93% in the no preference group gave at least one turner answer, this was only the case for 31% in the frontal pointing 1 group. A possible explanation for the link between the turner and no preference strategies could be that no preference acts as a kind of pre-stage to a complete turner strategy. Participants with strong proclivities for the use of a non-turner strategy might start to partially apply a turner strategy for some of the trials. The data even suggests a temporal development in which turner responses become more frequent among participants in the no preference group as can be seen in Figure 7. The number of turner answers is the only one constantly growing and ends up being the most frequent question in the fourth trial. However, since the experiment only included four trials, conclusions about temporal development have to be taken with a grain of salt.

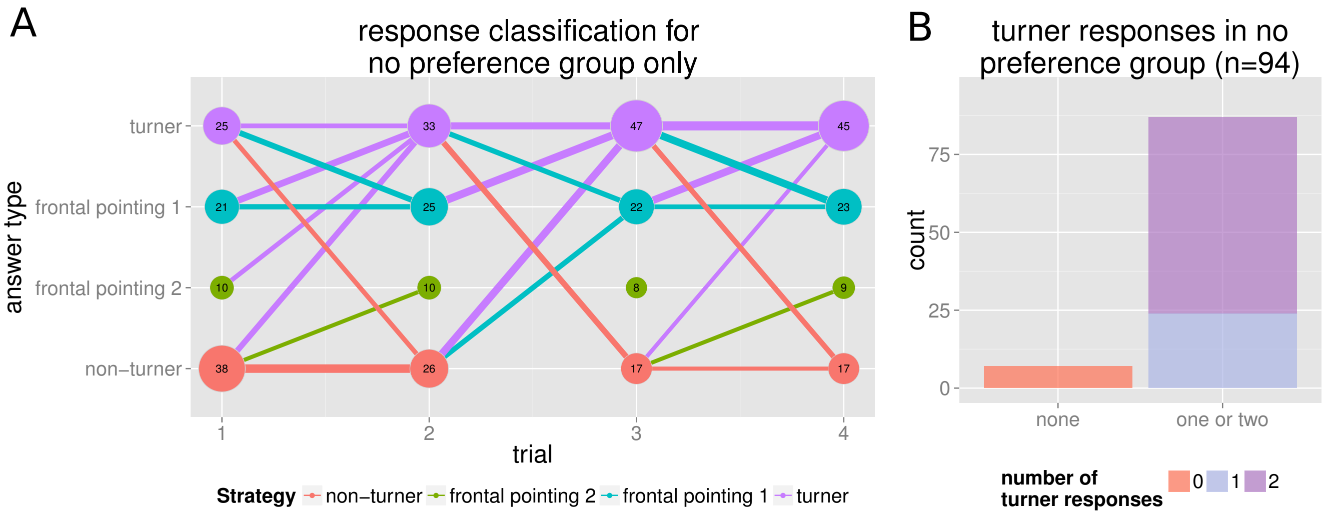

Figure 7: **A:** Strategy graph for the no preference group. While the number of frontal answers stay almost constant, the number of turner answers constantly grows and the number of non-turner answers shrinks. Also participants giving all sorts of answers before change to a turner answer in subsequent trials, whole exchange among other answering types is more limited.
B: 87 participants (93%) within the no preference group gave at least one turner answer.

4 CONCLUSION

4.1 LIMITATIONS

There were several limitations of the study. First, the small number of trials. Especially findings about a trend towards a turner strategy within the group classified with no preference have to be taken with care. Second, due to the nature of the study being conducted as classroom experiments, several limitations are present: a biased perception of the stimulus due to extreme viewing angle, interaction and copying between participants and simple issues like lack of motivation or inattentiveness. Third, although the experimenter took care to explain the task thoroughly, not all participants perfectly understood the task as indicated by the frontal pointings 1. Nevertheless, we think we minimised those issues wherever possible and were able to overcome the remaining noise via the large sample size.

4.2 REVISITING THE HYPOTHESIS

Concerning the initial hypothesis of the study we can conclude as follows:
Gender effects are quite limited. Our results contribute to the controversy around sex and gender differences in spatial navigation. We could not replicate a general influence of gender as in [**?**]. A gender influence appeared only in the pictorial condition and, even more interesting, only among Caucasians.

This may may be due to the extremely high amount of turners among male Caucasians.
Sex difference in human spatial abilities are well established in the literature [**??**], the most stable difference being found for mental rotation tasks. Here, women scored significantly worse compared to men, which was assumed to be correlated with the female bias towards the use of landmark based strategies compared to orientation based navigation strategies [**???**]. However, this view was somewhat challenged by Parsons and colleagues [**?**], who found, that the sex difference observed in mental rotation tasks vanished when a 3D virtual environment instead of a paper and pencil test was used for the task. They offered the possible explanation that it was the creation of a 3D representation from 2D drawings that actually caused or inflated the bias, not necessarily the task itself. If female participants in our study had higher difficulties in relating the 2D pictogram to the solution of the task, this could have been a reason for the higher amount of non-turners among females and why this bias vanished in the text condition.
Moreover, our findings might offer a possible explanation for the high controversy of gender differences in earlier studies. Our results can be read in the way that those differences are not universally present sex differences but gender differences tied to cultural background, explaining why their presence or absence is highly dependant on the sample demographics.

It is important how the question is posed. We were able to replicate the findings of [**?**] and extend them insofar as they also hold for a more abstract level were written spatial language and pictograms are used for answering instead of pointing and responding with spatial language. Our results add more evidence to the hypothesis that non-turner answers might indeed be due to a conflict of mental orientation and current body orientation that is more severe the more embodied the way of answering is.

Male Caucasians are a very specific subpopulation. Caucasians, especially males, seem to be a very specific subpopulation when it comes to virtual point to origin tasks. The ratio of male Caucasians using a turner strategy in the pictorial condition was extremely high while in all other groups the trend was exactly the other way around, strongly in favour for a non-turner strategy. This effect might have carried over to several other effects: the gender effect that was only observed among Caucasians, the condition effect that was not present for male Caucasians and the interactions effects that were only present against a male baseline and in comparing Chinese and Caucasians. We currently have no conclusive possible explanation for this effect and further research on this topic is needed.

4.3 FURTHER EFFECTS

An effect not hypothesised beforehand was the co-occurrence of the front pointing with the non-turner strategy on the one hand and the turner and no preference strategy on the other. We concluded that the border between the main strategies non-turner and turner might be harder to draw than previously assumed, especially during the first trials of a point-to-origin task. Interestingly the trend in the no preference group went clearly towards a turner strategy. Along the lines of Avraamides hypothesis this could mean that some participants, after an initial confusion due to the conflict of actual and virtual body orientation, get to a point were they resolve the conflict and adapt the virtual orientation as the one relevant for solving the task. The fact that we observed a trend in this direction and not towards a stable non-turner strategy might be due to our more abstract answering modes of which none involved physical pointing, the most embodied form of answering. We considered our answering modes more in between the continuum spanned by physical pointing and verbal description with spatial language.

4.4 OUTLOOK

The search for gender differences might be a complicated quest since our results suggest an interaction with task and possibly also with ethnicity. Instead of directly searching for sex differences, future studies should focus on investigating these interactions and aim for demographically more diverse samples. Our work gives more evidence to the embodied reference frame conflict hypothesis of **?**, however further investigations are needed to determine if non-turner answers are reflecting the use of an allocentric

reference frame or the use of an egocentric reference frame that is still aligned with the physical body orientation. An focused investigation of the turner / non-turner behaviour over more trials without feedback, looking for a resolution of a the hypothetical reference frame conflict might be fruitful.
The newly found influence of ethnicity on the strategy selection for triangle completion tasks adds a new facet to the influence of individual proclivities, motivating more studies with demographically diverse samples to get a more complete picture.

DISCLOSURE/CONFLICT-OF-INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

The statement about the authors and contributors can be up to several sentences long, describing the tasks of individual authors referred to by their initials and should be included at the end of the manuscript before the References section.

ACKNOWLEDGEMENT

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text.

Funding: Text Text Text Text Text Text Text Text.

SUPPLEMENTAL DATA

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text.

FIGURES

Figure 1. Enter the caption for your figure here. Repeat as necessary for each of your figures.