Spontaneous movements during a spatial perspective-taking task

Sergiu T. Popescu

Ecole des Hautes Etudes en Sciences Sociales

Centre National de la Recherche Scientifique

Paris, France

Mark Wexler

Université Paris Descartes

Centre National de la Recherche Scientifique

Paris, France

Word count (text excluding title, references, author affiliations, acknowledgments, figures and figure legends, but including the abstract): **8898**

Abstract

While cognitive processes can affect body movements in a variety of ways, the reverse effect of movements on cognition is less well studied. Gestures are thought to support concurrent reasoning and communication in subjects performing them, but the complexity of the tasks used to study these interactions has only allowed qualitative evaluations of the role of movements. Here we investigate a spatial perspective-taking task that is demanding even for experts, but is simple enough to allow a quantitative analysis of concurrent spontaneous movements. The task, which involves reporting object locations from points of view differing from that of the participant, was studied in apprentice and expert air traffic controllers, who perform similar tasks frequently. We found spontaneous low-amplitude head rotations that were related to concurrent task parameters: participants turn the head as if to align themselves with the task-relevant viewpoint, but with amplitude that is 10-100 times too small for visual alignment. Significantly, we found evidence that these movements are not epiphenomenal empty gestures, but actually play a causal role in the cognitive task: while movements before response onset are tightly coupled to task parameters, no coupling is found in movements performed later. Furthermore, the causal role of spontaneous movements is modulated by expertise: movement parameters partly predict response times in the perspective-taking task, but only in apprentice air traffic controllers. Taken together, these results suggest that motor activity, via incipient movements, triggers sensorimotor predictive mechanisms that drive cognitive perspective change, with subsequent inhibition of most of the overt movement.

Keywords: perspective taking, spontaneous movement, sensorimotor prediction, spatial cognition, gestures

Introduction

People have been found to produce spontaneous and sometimes systematic movements in a variety of spatial reasoning tasks with no explicit motor requirements (Chandrasekharan, Athreya, & Srinivasan, 2010; Chisholm, Risko, & Kingstone, 2014; Chu & Kita, 2008, 2011; De Maeght & Prinz, 2004; Knuf, Aschersleben, & Prinz, 2001; Popescu & Wexler, 2012). Because the tasks in questions do not require verbal communication, the observed movements are not speech-accompanying gestures (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001). The frequency of the spontaneous movements increases with the difficulty of the spatial task (Chandrasekharan, Athreya, & Srinivasan, 2010; Chisholm et al., 2014; Chu & Kita, 2011). Preventing the execution of spontaneous movements seems to lead to lower task performance (Chu & Kita, 2011; but no effect in Chandrasekharan et al., 2010). Such movements are even produced when merely observing a spatial task (De Maeght & Prinz, 2004).

One role of these movements may be to reduce cognitive load by offloading some cognitive operations onto the physical environment via the action-perception loop (Goldin-Meadow et al., 2001; Kirsh, 1995a). For instance, instead of performing a mental rotation of a physical object, one may be able to put the object into the required orientation, and to 'read off' its appearance (Kirsh & Maglio, 1994). In other situations, one may be able to offload visuo-spatial memory representations onto the environment by rearranging objects, or by pointing (Kirsh, 1995b). This sort of offloading results in movements that fall outside explicit task requirements.

Another possible role of spontaneous movements may be to activate sensorimotor prediction mechanisms (Davidson & Wolpert, 2005; Wolpert & Flanagan, 2001). The output of these mechanisms, namely the anticipated perceptual effects of the incipient actions, could contribute to performing cognitive tasks such as mental image transformations (Wexler, Kosslyn, & Berthoz, 1998; Wohlschläger & Wohlschläger, 1998) and spatial perspective taking (Popescu & Wexler, 2012). Once the prediction mechanisms are activated, most of the simulated movement can be overtly inhibited, explaining the generally observed small amplitude of spontaneous movements. Brain imaging studies that have found motor activation during tasks such as mental rotation (Vingerhoets, De Lange, Vandemaele, Deblaere, & Achten, 2002; Zacks, 2008) support the sensorimotor prediction mechanism.

How does expertise in a domain involving spatial tasks affect concomitant motor activity? One might expect that, with prolonged practice, experts may better internalize the underlying spatial reasoning processes and therefore display less overt motor activity during frequently practiced spatial tasks. For example, experts may rely more on memory representations acquired during their training (Frank & Macnamara, 2017; Logan, 1988). However, some experimental findings suggest that, on the contrary, experts may produce more motor activity than non-experts. Kirsh & Maglio (1994) found that novice Tetris players (a video game involving rapid rotation of geometric figures) rely more on mental rotations, while more advanced players offload these rotations to the action-perception loop, performing more manual rotations. Trafton et al. (2006) found that expert scientists are more likely to produce gestures related to spatial transformations than journeymen.

In addition, several recent findings using interference paradigms support the claim that motor activity plays a causal role in visuo-spatial tasks in experts. In a dual-task paradigm, in athletes compared to control participants, the recall of a sequence of movements presented visually was more affected by an interfering sensory-motor task (tapping the tips of hands and feet) than by a verbal task (repeating a verbal pattern; Moreau, 2013). Similarly, disturbing the motor planning via a dual motor interference task lowered the performance of mental abacus calculations in experts while disturbing the visual or proprioceptive feedback did not; in non-experts, the three types of disturbances had the same effect (Brooks, Barner, Frank, & Goldin-Meadow, 2017). This later study suggests that experts rely more on motor planning than other inputs as compared to non-experts (see also Beilock, Lyons, Mattarella-Micke, Nusbaum, & Small, 2008).

Motor activity has also been shown to contribute to increased performance in spatial perspective taking (SPT) tasks. In these tasks participants are required to adopt an imaginary viewpoint (a novel perspective) that is different from the physical viewpoint or from the one used during learning (the initial perspective). The novel perspective may differ from the initial one by a rotation, a translation, or a combination of the two; however, rotations of imagined perspective were found to be the most cognitively demanding (Presson & Montello, 1994; Rieser, 1989). The larger the angle of rotation between the initial and the novel perspective, the lower the task performance (as measured by response time and error rate)—the *angular disparity effect* (Easton & Sholl, 1995; Rieser, 1989; Roskos-Ewoldsen, McNamara, Shelton, & Carr, 1998). When accompanied by active movements, even in the absence of vision (e.g., when blindfolded), performance on the SPT tasks involving perspective rotations is better compared to conditions in which only the visual information is provided (Klatzky, Loomis, Beall, Chance, & Golledge, 1998; Rieser, 1989). Even when the perspective rotation task is only imagined, performance is better when imaging a rotation of the self than when imagining an equivalent rotation of the environment around oneself (Wraga, Creem, & Proffitt, 2000). The mere simulation of movements has been shown to facilitate performance in other domains, involving dynamics of the body and objects (Hegarty, 2004; Schwartz & Black, 1999) and thus seems to be a cognitive strategy beneficial in a variety of tasks. The simulation of movements amounts to a motor contribution, because motor imagery involved in simulating movement has been shown to share behavioral features and neural mechanisms with overt movements (Jeannerod, 2001; Zacks, 2008).

In a previous study we measured spontaneous movements during a spatial perspective-taking task requiring a imaginary perspective rotation (Popescu & Wexler, 2012). Participants were shown simple maps composed of two streets: the street they were initially located on, and the street containing their destination. The task was to report whether they would turn left or right at the intersection to reach the destination. When the participant's initial street was oriented upwards on the map, the task was trivial: destinations on the right required right turns, and ones on the left required left turns. The more the initial orientation deviated from upwards—we call the difference the *deviation angle*—the more difficult the task became, leading to higher response times and error rates. Despite having no explicit motor task, we found that participants spontaneously produced miniature head rotations whose direction and amplitude were reliably related to parameters of the cognitive task—the direction and the magnitude of the angle of imaginary perspective rotation.

When we call our task “spatial perspective taking,” this is simply how we described the task to our participants. They could have performed it using a variety of different strategies, which we will come back to in the Discussion.

Present study

In the present study we created a more challenging version of our previous SPT task and tested it on two groups with different levels of expertise. One goal of the present study was to test whether any spontaneous movements play a causal role in spatial reasoning, or if they are mere contaminations of the motor system by the cognitive task, epiphenomena playing no causal role in the task at hand. Another goal of the present study was to test whether a population of individuals with high spatial ability and who are extensively trained on perspective taking tasks would still produce spontaneous head rotations with similar characteristics as those of untrained individuals.

A group of experienced air traffic controllers, and another group of apprentice controllers were tested. One of the selection criteria for this profession is high spatial ability, which is tested on admission examinations for air traffic controller training. Later, during their work, the main duty of air traffic controllers is to monitor aircrafts, usually several simultaneously (Shorrock & Isaac, 2010). The position of the planes is displayed on a monitor on which north points up, while giving indications to pilots from the *pilot’s* point of view (for example, warning the pilot of another aircraft at the pilot’s 3-o’clock direction), most often different from controller’s point of view, depending on the orientation of the plane Hence, air traffic controllers must switch reference frame for every plane. They are therefore ideally suited for the purpose of our study as having both a high spatial ability and an extensive practice of spatial perspective taking under high cognitive load conditions.

We used a spatial imaginary perspective-taking task that specifies a “you-are-here” position on a simple map along with a “you-are-looking-there” facing direction; together they define an *imaginary orientation* (see Figure 1). The angle between the physical “north-up” and the imaginary orientation is the main predictor variable in SPT tasks; we will refer to it as the *deviation angle*. Participants were required to verbally report the position of a target relative to that imaginary orientation or perspective. Verbal report allowed participants to choose from multiple responses (see next paragraph) in a way that interfered as little as possible with spontaneous skeletal movements.

In order to make the task challenging for spatially skilled individuals and to avoid a ceiling effect, we modified our previous stimuli and task (Popescu & Wexler, 2012). Instead of a binary left-right response, we required participants to report a more precise position of a target from an imaginary perspective; this was done using a clock-face standard, which allowed for ten possible responses (corresponding to ten discrete positions of the hour hand, excluding twelve and six o’clock). If this SPT task is non-trivial even for air traffic controllers or apprentices, we should observe the usual angular disparity effect: the average response onset time and error rate should increase with the increase of absolute deviation angle.

When a perspective-taking task is cognitively demanding, if spatial updating uses a prediction based on motor processing, we expect that the motor execution would not be fully inhibited and we should observe spontaneous head rotations whose spatial parameters on each trial are specifically related to the geometry of the cognitive spatial task being performed.

If spontaneous movements play a causal role in performance, they must occur before response onset. Spontaneous movements occurring after response onset would be expected to have different spatial parameters; more specifically, those parameters should not be related to the geometry of the cognitive task. If, on the contrary, spontaneous movements are only epiphenomena—not playing a causal role in performance but simply triggered by some sort of association—then spontaneous movements around the response onset, as well before as after it, should share common spatial parameters.

Methods

Participants

Participants were recruited among the air traffic control personnel of the French air force base in Cinq-Mars-la-Pile. Twenty-one unpaid participants took part in the experiment. Due to a technical problem, the data of one participant were not recorded. Out of the 20 remaining participants, 3 were removed due to low level of confidence in their responses by the automatic speech recognition system (mean confidence over all trials below 95%). The final sample of 17 participants consisted of four females and thirteen males (mean age 27.6, SD 8.7 years). Ten participants were qualified air traffic controllers (mean age 32.4, SD 8.5 years). Their professional experience following their apprenticeships ranged between 1.3 and 25.3 years, with a median of 4.5 years. Seven other participants were apprentice controllers (mean age 20.9, SD 1.3 years). All had normal or corrected-to-normal vision and were naïve as to the purpose of the experiment.

Task and stimuli

Participants were presented with a simple map on a computer display, representing three objects (see **Figure 1**). They were required to imagine that they were standing on the green circle on the map (i.e., *location point*, it defines the imagined position) looking towards the direction of a star (*facing point*, depicted by the yellow star with blue contour; it represents the imagined orientation). For the task that the subject performed, the line between the location point and the facing point defined the “12 o’clock direction.” The third point on the map was shown by a three-colored bull’s-eye that we refer to as the *target* *point*. The participants' task was to say out loud, as quickly and as accurately as possible, what the position of the target point was, relative to their imagined orientation which was defined as 12 o'clock. The response was given as an hour on a standard clock face. For example, the correct responses are five o'clock in **Figure 1A** and nine o'clock in **Figure 1B**. Participants were told to say, in French, the number corresponding to the hour (for example, "cinq" for 5 o'clock).

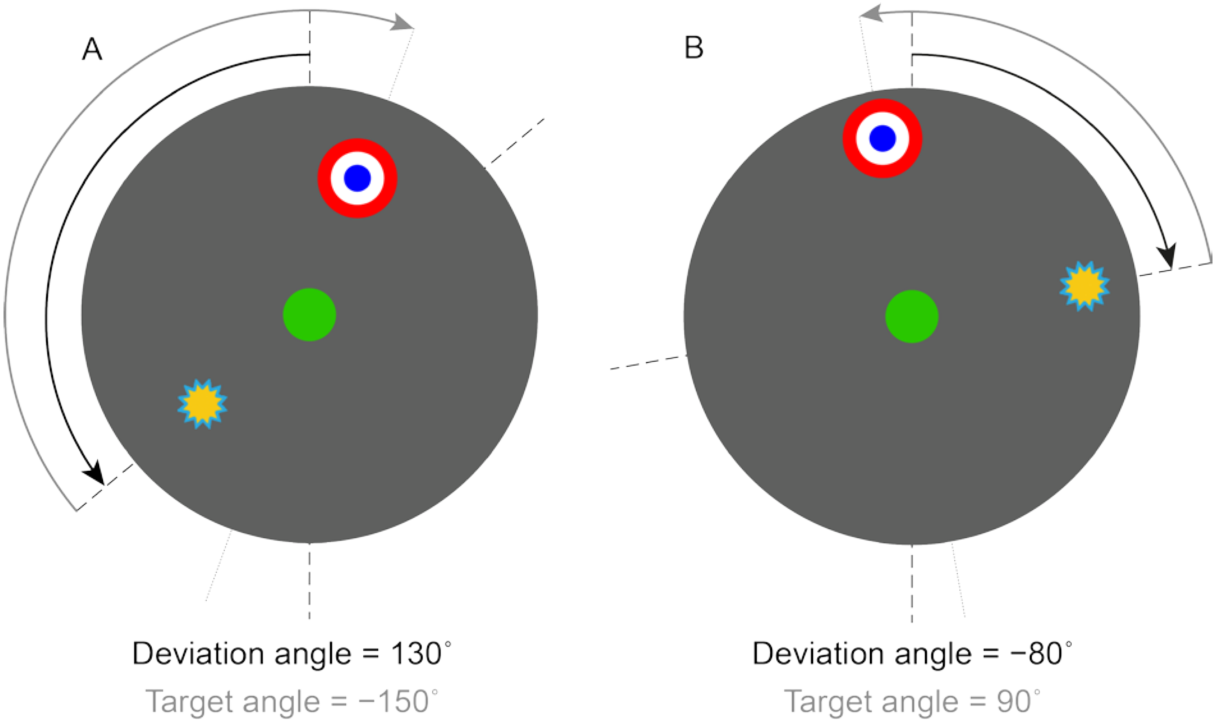


Figure 1. Stimuli and task. Participants were required to imagine being at the position of the green circle (*location point*) facing the yellow and blue star (*facing point*). We refer to this orientation as the *imagined orientation*. If the imagined orientation corresponded to 12 o'clock, the task was to verbally report the direction of the three-colored bull's-eye (*target point*) using the hours on a clock face. We call the angle between the upward-facing direction and the imagined orientation the *deviation angle*. **(A)** An example stimulus with a deviation angle of +130° (counterclockwise angles are positive in our convention) and a target angle of 150°. The correct response is 5 o'clock. **(B)** An example stimulus with a deviation angle of 80° and a target angle of +90°. The correct response is 9 o'clock.

An important parameter defining each stimulus was the *deviation angle* (see **Figure 1**), defined as the angle between the imagined orientation (the line between the location point and the facing point) and the upward, twelve-o'clock direction in the image. We take the upward direction as the "zero" of the deviation angle because previous studies (Shepard & Hurwitz, 1984) showed that it is easiest to perform the task when one's initial imagined orientation is upward. Deviation angles are taken as positive counterclockwise and negative clockwise. The other parameter defining a stimulus was the *target angle,* defined as the angle between the imagined orientation and the target point (see **Figure 1**). The correct response on each trial was the target angle. Varying the target angle allowed us to have multiple dissimilar trials with the same deviation angle.

Some geometrical details of the stimuli (but not the deviation and target angles) were randomized on each trial, in order to increase the visual variability of the stimuli. On each trial, the location point was displayed in a random position inside a square subtending about 2.9 degrees of visual angle (dva). The distances between the location point and the facing point, and between the location point and the target point, were randomly and independently chosen to lie between about 2.9 and 5.9 dva. The location point was displayed as a green circle with diameter about 0.7 dva, the facing point a yellow star with a blue outline and diameter 0.7 dva, and the target point a blue, white and red bull's-eye with diameter 0.9 dva. The background was black.

Design

We used a randomized factorial design. The deviation angle—our main predictor variable—had twelve levels: ±30°, ±55°, ±80°, ±105°, ±130°, ±155°. In our convention, counterclockwise angles are positive and clockwise ones are negative. The secondary independent variable, the target angle, had ten levels: −30° (“1 o’clock”), −60° (“2 o’clock”), −90° (“3 o’clock”), −120° (“4 o’clock”), −150° (“5 o’clock”), +150° (“7 o’clock”), +120° (“8 o’clock”), +90° (“9 o’clock”), +60° (“10 o’clock”), and +30° (“11 o’clock”). The crossing of the levels of the deviation angle and target angle created 120 conditions, each of which was repeated twice, resulting in 240 trials.

Procedure and apparatus

Each trial lasted 6 s. The visual stimulus was displayed throughout the length of the trial. The stimuli were displayed on a laptop monitor (Dell Latitude E5520, 15.6” display, 1920 x 1080 resolution). Participants performed the task while standing in front of the laptop display that was positioned at a confortable height and at approximately 70 cm from the participants’ eyes.

In order to obtain the participant’s response, throughout the length of the trial we recorded an audio stream from a lightweight headset microphone (Logitech, USB Headset A-00018). Audio data was recorded at 44.1 Hz, 16 bits/sample. The audio data was stored for subsequent off-line analysis.

Spontaneous head movements were recorded using a laserBIRD optical motion tracker (Ascension Technology, Burlington, VT). The tracker measures the 3D position and orientation of a moving sensor with respect to a stationary emitter, at a sampling frequency of 120 Hz. The sensor was rigidly attached to a light yet tight-fitting helmet that participants wore on their heads during the experiment. The emitter was held in place using a nearby tripod. In the right-handed coordinate system we use the positive X-axis points to the subject’s right, the Y-axis points forwards for the subject (into the monitor), and the Z-axis points upwards. Positive and negative rotations about the X-axis correspond to pitch rotations tipping the head backwards and forwards, respectively. Positive and negative rotations about the Y-axis correspond to roll rotations to the subject's right and left, respectively. Positive and negative rotations about the Z-axis correspond to yaw rotations to the subject's left and right, respectively.

The main experimental session consisted of 240 trials per participant. Each trial began when the participant pressed a mouse button. Halfway through the session participants were encouraged to take a break. Prior to the main experimental session, participants performed about 5 learning trials. The entire experimental session including the instructions, learning trials and breaks lasted about 40 min.

Speech recognition

The audio stream recorded on each trial was analyzed off-line in order to extract responses and response times. This analysis was performed by a specially written program using the French version of the Microsoft speech recognition engine. The program was designed to recognize a spoken response among a preset list of possible responses and to return, along with the recognized response, its onset time (which was how we calculated response time) and the confidence level of speech recognition. The possible responses were the French numbers from one to eleven. No individual calibration was performed. The program sometimes reported no response, either because the participant did not respond during the 6 s of the trial, or because the response was not identifiable.

After processing all trials using the speech recognition program, trials found to have correct and incorrect responses as concerns the task were analyzed separately.

On 3641 trials the speech recognition program detected responses that were correct as concerns the task, out of a total of 4080 trials for all 17 participants (89.2%). Out of these

3641 trials, there were 137 trials (3.8%) with a detection confidence level below 95%. Manual investigation of these 137 trials revealed that 2 trials (1.5%) were incorrectly recognized. These 2 trials were eliminated from the analysis. Thus, we kept 3639 trials for our analysis of correct responses.

As an indication of potential voice recognition errors, we manually checked all the trials for the participant who had the lowest average level of confidence, 94.5%. Out of 199 trials for which our program detected a correct response, there were 2 erroneously recognized responses (1.0%), which were the same 2 trials mentioned in the previous paragraph. We therefore expect a rate of recognition errors of no more than about 1% for trials where speech recognition finds correct responses.

We found a total of 422 out of 4080 trials (10.3%) where the speech recognition program found responses that were incorrect as concerns the task. We manually checked all of these 422 trials by hand, and found that 77 of them (18.2%) were incorrectly recognized. We eliminated those 77 trials from analysis. Thus, we analyzed the remaining 345 trials with incorrect responses.

Finally, on 17 trials (0.4%) the speech recognition program failed to detect a response. We removed these trials from the analysis.

Motion data

We found a small number of missing motion tracker samples. We eliminated from the analysis trials with 10 or more missing samples (0.3% of all trials).

We concentrated our analysis of motion data on head rotations (see Popescu & Wexler, 2012 for a discussion of this choice). We converted rotation matrices to an axis-angle representation of rotations using the inverse Rodrigues formula. To characterize the spontaneous rotation on each trial, we selected the sample on each trial that had the largest rotation angle (Popescu & Wexler, 2012).

Statistical analysis

All reported *p*-values are rounded to the nearest thousandth; unless otherwise specified, the statistical tests are considered significant at $p < 0.05$. All analyses using bootstrap included at least 10^4 resamples. All reported t-tests were two-tailed.

Outlier values are defined using the first and third quartiles (*Q1* and *Q3*) and interquartile range (*IQR = Q3 – Q1*): outliers are values lying outside the interval [*Q1 – 1.5 IQR, Q3 + 1.5 IQR]*; far outliers are values lying outside the interval of *[Q1 – 3 IQR, Q3 + 3 IQR]*.

We applied a linear mixed-effects (LME) model (Pinheiro & Bates, 2010) to fit the correct reaction time data. The computations were performed using lme4 library (Bates, Maechler, Bolker, & Walker, 2014; Bates, Maechler, Bolker, & Walker, 2015) running in open-source software R (R Core Team, 2013). The random effects of all models were maximum, as recommended by Barr, Levy, Scheepers, & Tily (2013). To define statistical significance of fixed effects, we used a backward stepwise simplification approach that allowed us to find a model with only statistically significant fixed effects: we started by an initial full model including all factors and meaningful interactions; at each simplification step we then fitted all models that differed from the current one by dropping a single term in fixed effects part while maintaining the same random effects and compared these reduced models to the original one with a likelihood-ratio test (Pinheiro & Bates, 2000). We further excluded from the model, the terms that could be dropped without a significant increase in unexplained deviance as indicated by the likelihood-ratio tests and fitted a new model without these terms, thus obtaining the initial model for the next step. We repeated the procedure until no more terms could be dropped from the model. Thus the last model was the best fit for the data. The fixed-effects terms in this final model could not be further excluded without a significant worsening of the model fit (as indexed by a significant increase of the residual deviance; the likelihood ratio tests of all terms were significant). The statistical significance of the terms in the final model was assessed via a similar procedure for each term: a full final model with the effect in question was tested against a reduced model without the effect in question; we reported the *p*-value of the likelihood ratio test of this comparison (Winter, 2013).

Results

Response time

We defined response time (RT) as the onset time returned by our speech recognition program. Overall, the mean RT for correct trials (the mean of individual median reaction times) was 2.02 ± 0.41 s (± between-participant standard deviation). On average, the ten qualified controllers responded faster (1.85 ± 0.40 s) than the seven apprentice controllers (2.25 ± 0.29 s) by about 400 ms.

To analyze RTs, we tested a linear mixed-effects model with the following fixed effects: level of expertise, absolute magnitude of the deviation angle, direction (sign) of the deviation angle, and all the interaction terms. The random effects terms were maximal and included, for each participant, an intercept and random slopes for the effects of all the terms used as fixed-effects predictors (including correlation parameters; for more details see Data Analysis in the Methods section). Visual inspection of residual plots did not reveal any obvious deviations from normality and a mild deviation from homoscedasticity.

The model showed a significant effect of absolute deviation angle on RT ($\chi^2(5) = 13.8$, $p = 0.017$) and a trend towards significance of the level of expertise ($\chi^2(1) = 3.65$, $p = 0.056$). All other fixed-effect terms did not reach significance. Note that a separate comparison of controllers' and apprentices' *median* RTs revealed a significant difference (2-tailed t-test, $t(15) = 2.23$, $p = 0.041$).

Since the sign of the deviation angle did not affect RT, we collapsed data for positive and negative deviation angles. The linear regression of absolute deviation angle on RT showed that the constant coefficients for all participants and the regression slopes for

13/17 (76%) participants were significantly different from zero (bootstrap with 10^4 resamples). At the population level, as estimated by the linear mixed effects model, the controllers are faster than the apprentices by 295 ms (SE 150 ms). Every additional degree of deviation angle increases the RT by 1.3 ms (SE 0.3 ms) for both groups (as the difference in slope between the controllers and apprentices was not significant). Given that the effect of the direction of deviation angle was not significant, **Figure 2** shows the mean RT as a function of the absolute deviation angle.

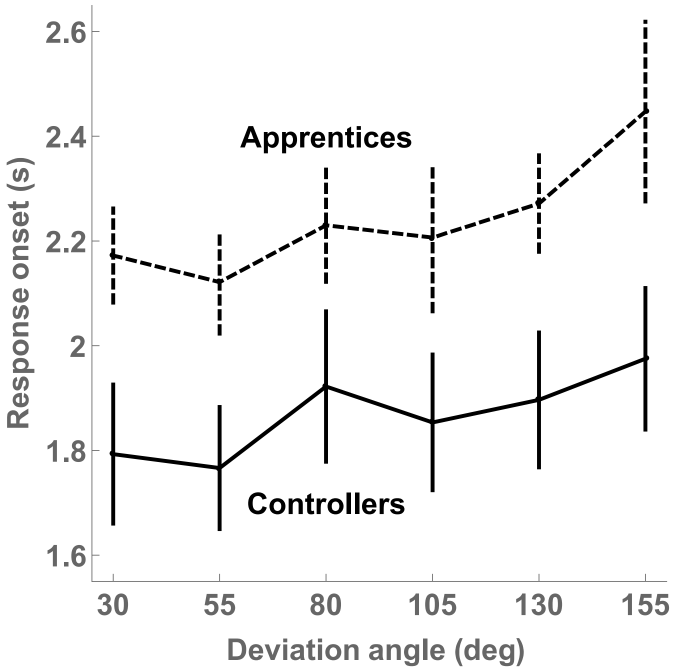


Figure 2. Response onset time as a function of absolute deviation angle. Each line represents the mean of participants' median RTs; the black line shows the data for qualified controllers and the dashed line—for apprentice controllers; the error bars are standard errors of the mean (of participants' median RTs).

In brief, there is a marginally significant difference suggesting that controllers are faster than apprentices, probably reflecting their higher level of expertise. Nevertheless, the sensitivity to the deviation angle is the same in both groups. The RT shows the usual angular disparity effect: it increases with absolute deviation angle. The fact that this effect is of similar magnitude for both groups leads to the conclusion that the faster RTs of controllers are not due to their faster processing of deviation angle but to a more general effect of expertise on the cognitive processing of the task. The curiously similar bump around 80-deg deviation angle in the two curves is much less pronounced when we do not collapse positive and negative deviation angles, and thus seems to be an accidental artifact of averaging.

The mean reaction time of incorrect trials across subjects (the mean of individual median reaction times) was 2.57 ± 0.71 s. On average participants took 556 ms longer to respond on trials where they responded incorrectly compared to trials on which they gave a correct response; the RT difference between correct and incorrect trials was statistically significant (paired samples t-test, $t(16) = 5.73$, $p < 0.001$). The fact that RTs on error trials were longer than on correct trials suggests that there was no speed-accuracy trade-off.

Error rate

The mean proportion of trials answered incorrectly was 0.087 ± 0.045 (± between-subject SD). The error rate for controllers was very similar to that of apprentices, respectively 0.087 ± 0.052 and 0.087 ± 0.036 and the difference between the two groups was not statistically significant (2-tailed t-test, $t(15) = 0.02$, $p = 0.98$). **Figure 3** shows mean error rates as a function of absolute deviation angle across participants: it appears that the error rate did not increase with the absolute deviation angle. Due to the limited number of incorrect trials and to the absence, at the descriptive level, of an effect of deviation angle on error rate, we have not used a more elaborate technique for error analysis.

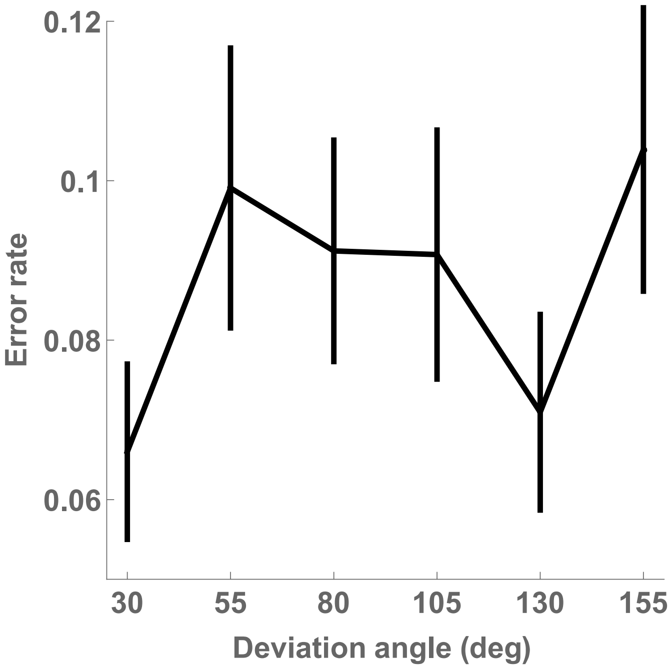


Figure 3. Mean error rate as a function of absolute deviation angle. The line represents the mean of participants' error rates at each absolute deviation angle. Error bars display standard errors of the mean.

It should be mentioned that there were no errors that would have corresponded to a total lack of spatial updating of imaginary perspective (i.e. keeping the perspective at zero degree deviation angle instead of adopting a new one required by the stimulus) and that only 8/337 (2.4%) errors were inversion errors, corresponding to an inversion of Location Point with Facing Point. Furthermore, over two thirds of the errors (69%) were by just "one hour" or 30°; this confirms that the participants were indeed performing the task as required.

In summary, participants took longer to respond on incorrect trials, suggesting the absence of a speed-accuracy trade-off. There was no difference in error rate between controllers and apprentices. Qualitatively, the errors corresponded to a partial failure during the processing of spatial updating (as denoted by the absence of errors that would result from either a total lack of updating or a systematic inversion of orientation). Due to the limited number of incorrect trials, only correct trials were further analyzed.

Spontaneous head rotations before response onset

For a spontaneous movement to affect the task response, it must occur before the response; therefore, we first analyzed spontaneous head rotations observed before response onset. If our hypothesis of motor contribution to spatial updating is correct and there is indeed a specific geometric relationship between deviation angles and the spontaneous head rotations, we expect that the amplitude of a head rotation, along with its direction as well as its axis should be related in a specific way to the deviation angle. We will first analyze the amplitude of head rotations, i.e., the absolute values of head rotation angles; we will then turn to the analysis of their directions and rotation axes. In this and all subsequent analyses of spontaneous movements, we only used the trials in which our voice recognition software identified a correct response.

As discussed above, we characterize spontaneous movements by the largest rotation (i.e., a rotation by the largest angle) before the onset of the response. This maximal rotation yields three values: the amplitude of the maximal rotation, its axis, and the time at which the maximum occurred (by definition less than or equal to the RT). Across participants, the mean amplitude of maximal head rotations executed before response onset was 2.46 ± 1.39 deg if we exclude an outlying value from one controller (and 3.23 ± 3.46 deg if we include it). The mean of the controllers' rotations was less than that of apprentices, respectively 2.21 ± 1.45 deg (3.54 ± 4.45 deg with the outlier) and 2.78 ± 1.36 deg, but this difference (with and without an outlier) was not statistically significant (t-test, without the outlier, $t(14) = 0.81$, $p = 0.43$). **Figure 4** shows head rotation amplitude as a function of the absolute deviation angles, across participants.

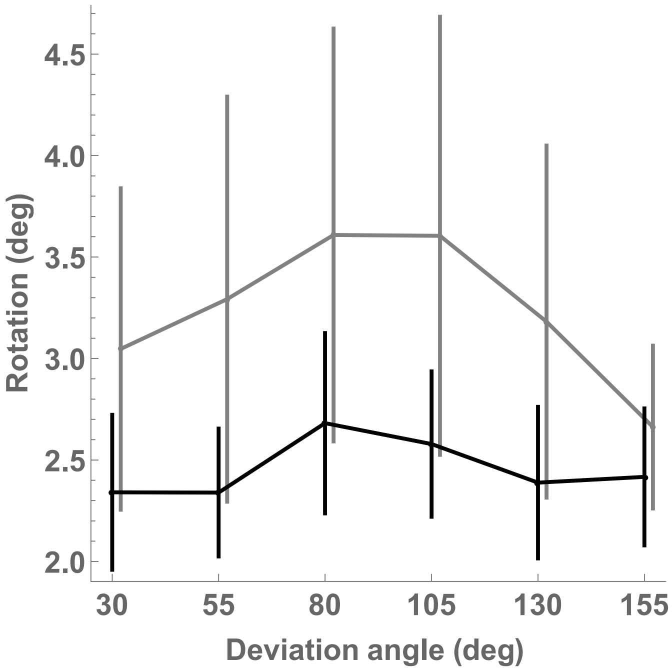
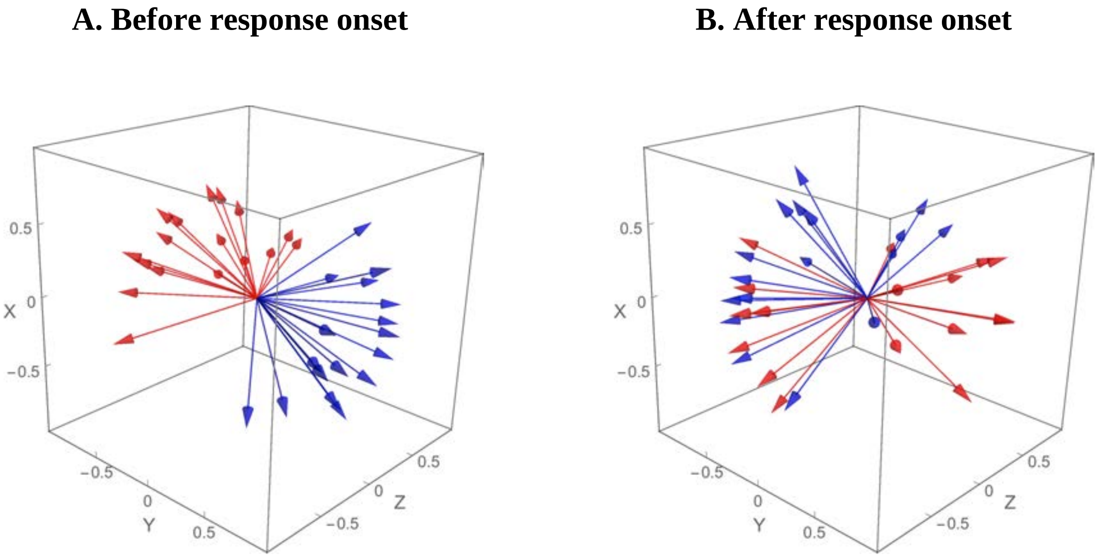


Figure 4. Absolute angle of head rotation as a function of absolute deviation angle, for rotations before response onset on trials answered correctly. Each line represents participants' mean head rotation angles; the black line shows the data for participants excluding the far outlying mean of one controller and the gray line—including the outlier. Error bars are standard errors of the mean. For better readability, the grey line is slightly offset to the right.

The regression of the head rotation amplitude on the absolute deviation angle shows that the mean constant term was 2.41 ± 1.54 deg (3.46 ± 4.6 deg with the outlier controller). The constant term for each of the seventeen participants was statistically significant while the regression slopes of 15/17 (88%) participants were not significantly different from zero (bootstrap over individual data with 10^4 resamples).

For controllers the mean constant term was 2.00 ± 1.28 deg (3.83 ± 5.92 deg with a far outlier) and for apprentices it was 2.94 ± 1.77 deg. This observed difference between controllers and apprentices was not statistically significant (t test, without the outlier, $t(14) = 1.24$, $p = 0.24$). This result suggests that participants produced head rotations of more or less constant amplitude, independent of deviation angle, before response onset.

We therefore assessed whether opposite directions of deviation angles (clockwise or counterclockwise) induced head rotations in opposite directions about some rotation axis: whether deviation angles to the right produced rotations in the opposite direction (but about the same axis) as deviation angles to the left. We computed the mean rotation axes separately for positive and negative deviation angles for each participant. **Figure 5A** shows these axes. It can be seen that axes for counterclockwise positive deviation angles (shown in red) and for clockwise negative ones (in blue) point in roughly opposite directions; we will refer to these axes as positive axis and negative axis, respectively.



A. Before response onset

B. After response onset

Figure 5. Individual rotation axes for positive and negative deviation angles. The individual axes for positive and negative deviation angles are shown in red and blue, respectively. **(A)** Data for rotations before response onset. **(B)** Data for rotations after response onset.

Each 3D rotation axis is a unit vector with three components, X (pitch), Y (roll), and Z (yaw). To understand more precisely how these components of the axes were related for deviation angles of opposite signs, we compared the positive and negative axes, component by component. **Figure 6** shows, separately for each component, its values for positive (CCW) and negative (CW) axes, or deviation angles. If, for example, the positive and negative axes oppose on the Y component, the *signs* of the Y component for positive and negative axes should be opposite; this is the case for Y components of most axes, as shown in **Figure 6**. The proportions of participants with X, Y and Z components of opposite signs are 0.65, 0.94, 0.12, respectively. Only the proportion of oppositions about Y-axis is significantly higher than 0.5 that would be expected by chance alone (sign test, $p < 0.001$).

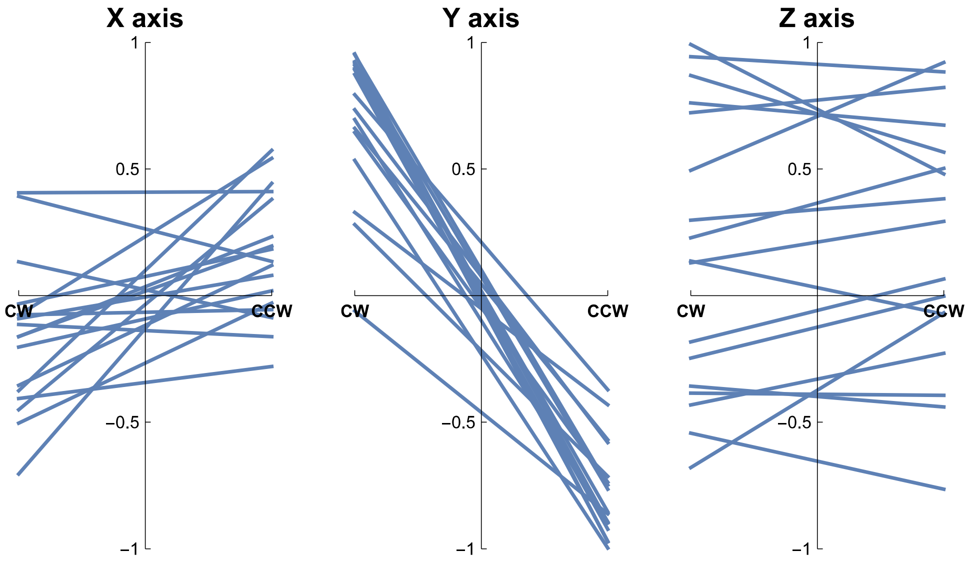


Figure 6. X, Y, and Z components of individual mean head rotation axes, for positive (CCW) and negative (CW) deviation angles. Each line connects two values of a component observed on the mean axes for CCW and CW deviation angles.

The proportion of oppositions about Z-axis is significantly lower than 0.5 (sign test, $p = 0.002$). The Z-axis corresponds to the upward direction. The yaw rotations about this axis probably reflect an idiosyncratic component of rotation execution. The X components of the axes appear to be smaller and random. We can summarize these results by saying that positive and negative deviation angles generally lead to opposite roll rotations (Y-axis), combined with yaw rotation independent of the sign of the deviation angle.

The oppositions between the positive and negative axes come from their Y components. Head rotations about this axis correspond to roll movements. Because the Y-axis points forwards, and we're using a right-handed coordinate system, rotations about the positive Y-axis correspond to rolls to the right, and negative to the left. Recall that in our convention, positive deviation angles correspond to counterclockwise deviations from the upward vertical in the image plane, and negative to clockwise. This means that the rotations about the negative Y-axis (left) for positive deviation angles (counterclockwise) are in the *same direction* as the deviation of the imagined orientation from the upward vertical, and similarly for negative deviation angles. We can summarize this by saying that, through their spontaneous head movements, participants attempted to align themselves with the facing direction *on the stimulus map itself*—rather than in the space represented on the map. If participants had imagined themselves in the space represented on the map, they would have performed rotations about the vertical (yaw) axis in order to align themselves with the facing direction—which is not what we have found. Because spontaneous rotation angles are much smaller than deviation angles, these attempted alignments with the map are very partial.

To summarize our results so far, we have found that for all participants RT increases with increasing deviation angle. Apprentices are somewhat slower than controllers. Before response onset, participants perform spontaneous head rotations during the trials. These rotations are of more or less constant amplitude with a mean value of roughly 2.5 deg. Importantly, the rotations are directionally related to the deviation angle: for opposite directions of deviation angle, the spontaneous rotations are in opposite directions, too, at least as concerns the roll component. The head rotations are roughly parallel to the direction of the deviation angle as it is displayed on the stimulus map: the head rotations are clockwise/counterclockwise about a front-back axis for clockwise/counterclockwise deviation angles.

Spontaneous head rotations on correct trials after response

Why analyze rotations occurring after response onset if they cannot affect the response any longer? If observed spontaneous rotations of the head are epiphenomenal, with no causal role on performance, rotations performed after the response should be functionally similar to those before the response. If, on the contrary, we find that rotations before the response are more closely related to the task than those after the response, this would argue that the rotations performed before response onset are not epiphenomenal but have a distinct causal role in performance. To compare rotations before and after response onset, we analyzed the maximum rotations after response onset in the same manner as those before response onset. We found the largest head rotation in the period after response onset with respect to head orientation at response onset. We searched for the maximum rotation during a period equal to the RT (or up to the end of the trial when the RT was greater than half the duration of the trial).

Figure 5B shows individual head rotation axes, separately for positive and negative deviation angles (depicted in red and blue, respectively). It can be seen that the axes for opposite deviation angles are much less clearly opposing than those of head rotations occurring before response onset that are shown in **Figure 5A**. It can also be seen that, compared to before response onset, a handful of axes changed direction after response onset. This probably corresponds to repositioning of head into its initial position, from the rotation before response onset. If the repositioning were systematic, a clear directional relationship to deviation angles would be found, but in a direction opposite to that observed before response onset.

To quantify the directional opposition of axes, we analyzed the positive and negative axes, component by component, in a similar way as for rotations before response onset. **Figure 7** shows the values of axis components, separately for positive and negative individual mean axes.

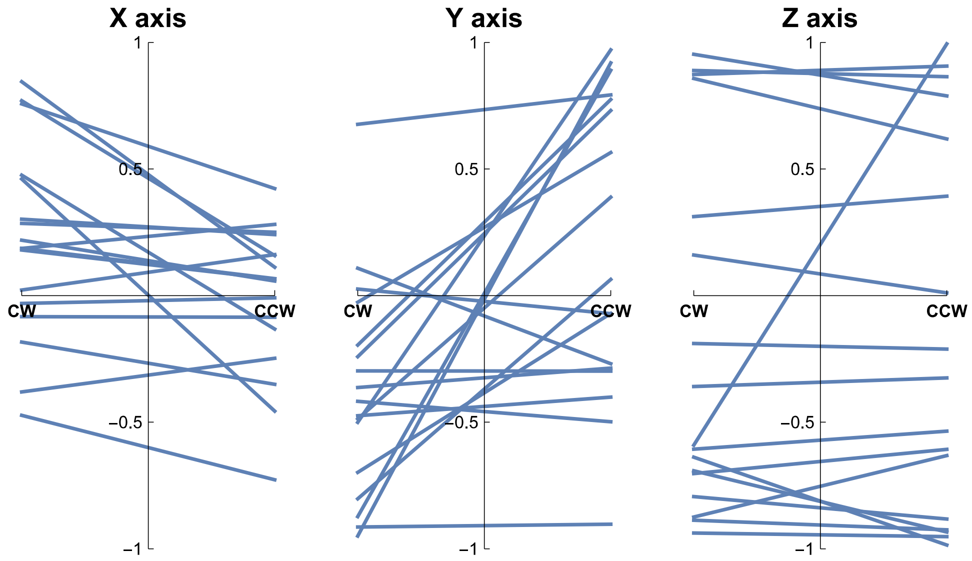


Figure 7. X, Y, and Z components of individual mean head rotation axes, for positive (CCW) and negative (CW) deviation angles. Each line connects the two values of a component, observed on the mean axes for CCW and CW deviation angles.

The proportions of participants with X, Y and Z components of opposite signs are 0.12, 0.58, 0.06, respectively. None of these proportions is significantly higher than 0.5 (as evidenced by non significant sign tests). The proportions of X and Z components are significantly lower than 0.5 (sign tests, $p = 0.002$ and $p < 0.001$, respectively), corresponding probably to rotations about a stereotyped axis. Compared to spontaneous head rotations before the response, those after response onset do not display a directional relationship to deviation angle.

Relation between spontaneous head rotations and RT

If spontaneously produced head rotations are related to the cognitive task, the time of the maximal rotation (TMR) may be correlated with the task response time. For each trial, we computed the time of the maximum head rotation observed during the entire 6 seconds of the trial. We then calculated, individually for each participant's correct trials, the correlation between the TMR and the RT. **Figure 8** shows the individual correlations, for controllers and apprentices.

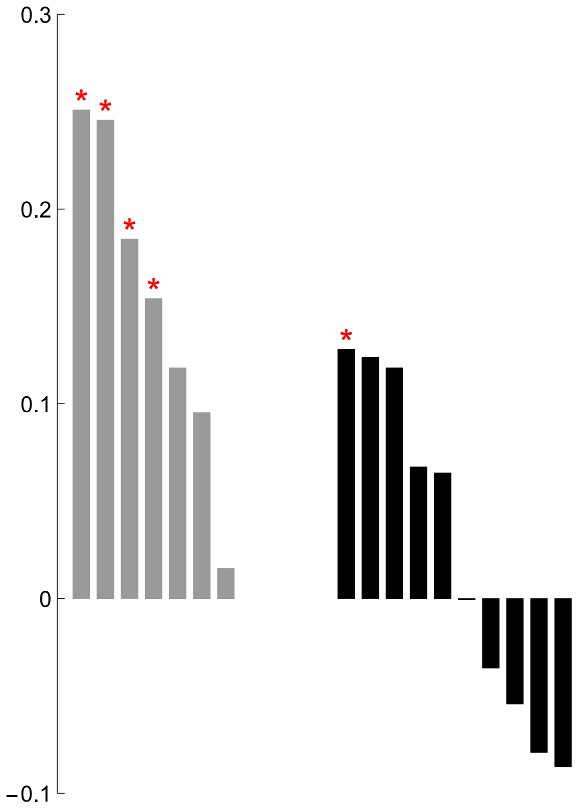


Figure 8. Individual linear correlations between the time of maximum head rotation across the entire trial and RT. Grey bars correspond to data for apprentices and black bars show data for controllers. Asterisks indicate correlations that are significantly positive.

Averaging the individual correlations over participants, we found that the mean correlation for the controllers was 0.02, and the population mean was not statistically different from zero ($t(9) = 0.9$, $p = 0.4$). In contrast, the apprentices' mean correlation was 0.15, with the population mean significantly greater from zero ($t(6) = 4.78$, $p = 0.003$, 2-sided). Moreover, bootstrap tests on individual data revealed that 4/7 apprentices had significantly positive correlations, while only 1/10 controllers did so. Finally, the population mean correlation was significantly greater in apprentices than in controllers ($t(15) = 3.04$, $p = 0.008$, 2-sided). It thus appears that in apprentices, or at least in some of them, spontaneous head rotations have a functional relation to the task performance because their rotations have a weak yet significant correlation with their RTs, and thus allow to (very partially) predict RTs. The controllers' rotations seem, on the contrary, to be less related to task performance.

Discussion

While performing an imaginary perspective-taking task, participants spontaneously produced miniature head rotations. These rotations were largely about the naso-occipital axis and were clearly directionally related to the deviation angle (the angle between the physical and the imaginary perspective)—but only before response onset. The rotations observed after response onset are not directionally related to the deviation angle. This indicates that rotations before and after response onset are functionally distinct and provides support to the idea that the incipient rotations prior to the response might be causally related to task performance, rather than being just epiphenomenal or causally ineffective movements driven by association with underlying cognitive processes.

Why would there be such a specific geometric relation between the spontaneous movements and the spatial cognitive task? These movements are difficult to explain as cognitive offloading (Risko & Gilbert, 2016), because only a small fraction of the mental transformation is actually offloaded: participants systematically turn in the direction to align themselves with the imaginary perspective, but their tiny rotations hardly help them to achieve the alignment. Instead, we propose that participants perform a simulation of self-motion corresponding to their imagined viewpoint rotation. Such a simulation could be activated through motor plans (that also lead to the incipient body rotations), activating sensorimotor prediction mechanisms. Once these mechanisms are activated, the rest of the movement can be overtly inhibited. The inhibition of most of the overt movement would account for the low amplitude of the observed spontaneous rotation, as compared to the actual rotation required to align oneself with the imaginary perspective.

Additionally, in apprentices but not in qualified controllers the timing of maximal head rotations had a small but significant correlation to their RTs, suggesting that the causal relation between spontaneous movements and the concurrent cognitive process was even tighter in apprentices.

Why are the task-related rotations specifically about the naso-occipital axis, or, in other words, why are they rolls rather than yaw or pitch rotations? We have referred to our task as one involving spatial perspective taking, but this is simply how we described the task to our participants. In reality, the task could have performed using a number of different strategies. One set of strategies would involve mental transformations in the world represented by the stimulus map. For example, the subject could imagine him- or herself, or an avatar, rotating from their current orientation (on the ground) to the facing direction, also projected onto the ground plane. Such transformations would involve rotations about the vertical or yaw axis. Alternatively, participants could have imagined their bodies, or a clockface, rotating in the frontal plane from the upright to the facing direction on the map itself, or could have mentally rotated the visual stimulus in order to align the facing direction with the physical upright. Such transformations involve rotations about the naso-occipital or roll axis. In this study, we have found that spontaneous task-related rotations are approximately rolls, in other words rotations about the naso-occipital axis. This supports the hypothesis that mental operations were with respect to the map, rather than in the world represented by the map. Interestingly, in our older study (Popescu & Wexler, 2012) we found that spontaneous rotations were sometimes about a vertical axis—suggesting that participants mentally projected the vertical map onto the ground plane—and sometimes about the naso-occipital axis. The strategy of mentally rotating the map could give rise to spontaneous hand movements, given the interaction between hand movements and mental rotation (Wexler, Kosslyn, & Berthoz, 1998; Wohlschläger & Wohlschläger, 1998). By also measuring spontaneous hand movements in addition to head movements, we could in principle identify the strategy used by each participant.

Even for participants who are highly trained on spatial perspective taking, the task is demanding. This claim is supported by the fact that the RTs increase significantly with absolute deviation angle, and that the error rate is quite high, about 6.7 ± 2.1% (median ± between-subject median absolute deviation). The task is more difficult than a related binary left-right task we used in Popescu & Wexler (2012). In the current task, the error rate for air traffic controllers and apprentices is higher than the laypersons' median error rate of 1.2 ± 0.6% in our previous study, and this difference is statistically significant (Mann-Whitney test statistic = 288, $p < 0.001$). As for spontaneous movements, in contrast to our previous result, the amplitude of head rotations does not increase with the amplitude of the deviation angle. On the average, the head rotation amplitudes observed in the current study were larger than in our previous study: the medians across participants were 2.64 ± 0.90 (± between-subject median absolute deviation) and 0.65 ± 0.26 deg, respectively. It is possible that the roughly constant-amplitude movements served to trigger a simulation process, with the overt movement inhibited after a fixed duration, leading to a fixed-amplitude rotation.

Mental rotations, when accompanied by corresponding rotations of the body, and even in absence of visual and auditory cues, are performed easily, with a nearly perfect level of task performance (Klatzky et al., 1998; Rieser, 1989). What exactly—in the continuum from action intention to its execution—is facilitating performance? Is facilitation happening before overt action and due to motor planning upstream, or does it occur during motor execution and is it rather due to its vestibular and proprioceptive feedback downstream? Several studies of spatial orientation found that active movements, executed in order to align one's actual perspective with the imagined one, improve task performance more than passive movements (in infants, Bai & Bertenthal, 1992; in adults, Chrastil & Warren, 2013; Wang & Simons, 1999; for reviews, see Chrastil & Warren, 2012; Wexler & van Boxtel, 2005). In these studies, passive movements are usually imposed whole-body motions (such as being turned in a chair), removing both efference copy and most proprioceptive cues to self-motion. These studies rule out vestibular input as a major contributor to performance improvement, because passive movements preserve most vestibular cues. Hence, the better performance as a result of active movements can be due to the two remaining sources of information about motion, present in active movements but absent in passive ones: motor efferent and proprioceptive reafferent cues. What is the specific contribution of each of these cue sources? In principle, there could be a continuum of weighted contributions of the two sources, depending on task and individual parameters. On one extreme, the exclusive contribution of perceptual feedback cues would necessitate the execution of full movements. On the other extreme, the exclusive contribution of motor efference would not require movement execution but instead its active inhibition. In support of the role of motor efference, Brooks et al. (2017) found that it was motor planning that contributed to task performance in experts performing mental calculations with an imaginary abacus: disturbing the motor planning via a dual motor interference task (participants were instructed to tap on a keyboard as the numbers to be added were presented) decreased performance while disturbing proprioceptive feedback (participants were either blindfolded or instructed to keep their hands flat on a table) did so only marginally; in non-experts, the three types of disturbances had the same effect. This study suggests that experts rely more on motor planning than other inputs as compared to non-experts (see also Beilock, Lyons, Mattarella-Micke, Nusbaum, & Small, 2008). The contribution of motor planning rather than that of movement execution is also suggested by Wohlschläger (2001) who showed that planning a hand rotation without executing it interferes with the performance of a mental rotation task. This interpretation is further supported by studies of dancers rehearsing new choreographies: the dancers often use partial movements—termed marking—instead of full movements, and marking is more efficient than the execution of full movements (Kirsh, Caballero, & Cuykendall, 2012). The contribution of motor efference to task performance is also indirectly supported by the finding that supplementary motor area involved in motor planning actively inhibits the activity if the primary motor area in charge of movement execution (Kasess et al., 2008).

How and why can motor planning lead to anticipation of action consequences, before movement execution? One of the possible predictive neural mechanisms is an *internal forward model* (Wolpert & Flanagan, 2001; see also the review by Davidson & Wolpert, 2005). According to these authors, "it is *internal* to the CNS, *models* the behaviour of the body and captures the *forward* or causal relationship between actions and their consequences" (emphasis ours, Wolpert & Flanagan, 2001, p. 729). In the first place, such a mechanism is necessary for accurate motor control: the current state of the body cannot rely on sensory feedback for this information, because it is too slow and noisy, particularly in the case of accurate or fast movements. The prediction of the model can be also used to reduce the sensory effects of self-motion (reafference) in order to highlight the other inputs (exafference), which can affect motor control. Attenuation of reafferent signals by efference copy has been supported by neurophysiological studies in insects and lower vertebrates (Bell & Grant, 1989; Poulet & Hedwig, 2006; Sperry, 1950; von Holst & Mittelstaedt, 1950), as well as studies of the stability of visual perception during eye movements in non-human primates (Cavanaugh, Berman, Joiner, & Wurtz, 2016; Duhamel, Colby, & Goldberg, 1992; M. A. Sommer & Wurtz, 2002; Sommer & Wurtz, 2008; for a review see Wurtz, 2008), and by psychophysical studies on humans (Blakemore, Frith, & Wolpert, 1999; Cardoso-Leite, Mamassian, Schütz-Bosbach, & Waszak, 2010; Desantis, Mamassian, Lisi, & Waszak, 2014). In our study, we found spontaneous low-amplitude head rotations that were related to concurrent task parameters: participants turn the head as if to assume the task-relevant viewpoint, but the movement is 10-100 times too small. The very partial nature of head rotations that we observed is consistent with the role of motor efference in perspective-taking performance via the operation of an internal forward model. The motor prediction by such an internal model could be used to anticipate the sensory consequences of a rotation of the point of view: the initial motor activation would be similar to that of a full body rotation. The efference copy or corollary discharge from this motor activation, activating forward models, could be used to anticipate the egocentric directions from the new perspective. Subsequent motor activation be inhibited farther downstream without loss of predictive power, greatly reducing the overt movement's amplitude. The temporal relation that we have found between spontaneous movement and the cognitive task—that spontaneous movements only prior to response are geometrically correlated to the task—further supports this efference copy explanation. Finally, the small but significant correlation between maximum rotations and RTs, but only in apprentices and not in experienced air traffic controllers, supports the idea that the nature of computations underlying spatial perspective taking changes during training. SPT is perhaps carried out via internal forward models in less-experienced participants, while relying more on memorized representations in participants with greater experience (Frank & Macnamara, 2017; Logan, 1988).

A separate research current, gesture studies, has also analyzed the spontaneous movements produced during cognitive tasks. Most of these studies have dealt specifically with hand movements, usually investigated in parallel to speech production, analyzing the effect of gestures either on the speaker or the listener (Goldin-Meadow, 2017; McNeill, 1992; Novack & Goldin-Meadow, 2017). Hand gestures have been found to improve performance of various tasks (Beilock & Goldin-Meadow, 2010; Goldin-Meadow et al., 2001; Hegarty, 2004; Schwartz & Black, 1999). Gestures are also more frequent and ample for more difficult tasks (Brooks et al., 2017; Chandrasekharan, Athreya, & Srinivasan, 2010). Encouraging gestures may improve performance (Chu & Kita, 2011). Spatial direction of gestures may be related to spatial parameters of the task (Chandrasekharan, Athreya, & Srinivasan, 2010; De Maeght & Prinz, 2004; Knuf, Aschersleben, & Prinz, 2001). Gestures occur more frequently with descriptions of an action, which elicits motor imagery, than with descitpions of a visual scene, which elicits visual imagery (Feyereisen & Havard, 1999). Among a number of theoretical models explaining how gestures arise, the one that explains both co-speech and co-though gestures, hence including motor imagery, is the *gesture-as-simulated-action* (GSA) framework elaborated by Hostetter & Alibali (2008). It posits that gestures originate in perceptual and motor simulations that are run during speech production and mental imagery. Action and perception simulations activate the neural areas involved in motor planning and perception, respectively. To prevent an overt action that may result from the motor activation, further movements needs to be inhibited downstream. When, during the simulation, the motor activation reaches certain strength, beyond an individual threshold, the gestures cannot be any longer inhibited and are physically executed. The interpretation of our results is consistent with the GSA framework and extends its application to the incipient head rotations occurring in absence of speech production, mostly invisible to human eye, that we measured during a pure reasoning task.

We believe that our methodology makes several advances in the study of spontaneous gestures. Traditional gesture studies use a time-consuming, qualitative methodology: human observers typically code gesture frequency and a few other visually perceptible variables, as the direction or configuration of the hand. To be visually perceptible and unambiguously categorized, the gestures studied in this way need to attain a minimum amplitude (e.g., at least 30 deg hand tilt in Chandrasekharan, Athreya, & Srinivasan, 2010; at least 10 deg head tilt in Risko, Medimorec, Chisholm, & Kingstone, 2014). Our automatic motion-tracking method offers a new way to study spontaneous movements, including those invisible to human eye. Further, our analysis methodology allows us to condense the high degrees of freedom in a motion trajectory to a manageable number of summary variables, such as an axis and angle of rotation. These summary variables allow us to study the relations between spontaneous gestures and the cognitive task being performed in an objective way.

Spontaneous movements during cognitive tasks are a potential channel of useful information that is too often discarded. Our study demonstrates an automatic motion-tracking methodology applied to spontaneous head movements in experts performing a spatial perspective-taking task, and shows that small head rotations are correlated to the geometry of each task performed. Furthermore, there are strong indications that the movements play a causal role in task performance, possibly through a mechanism of sensorimotor anticipation.

Acknowledgments

We are grateful to Thales Group SA and especially to Alain Léger, Kees Nieuwenhuis, and Philippe Chopin for support of this research. We are also grateful to the air traffic controllers of Cinq-Mars-la-Pile, and to their commanding officers, for their participation in this study. We thank Michel-Ange Amorim and Klaus Kessler for useful comments on a draft of this manuscript.

References

Bai, D. L., & Bertenthal, B. I. (1992). Locomotor status and the development of spatial search skills. *Child Development*, *63*(1), 215–26.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. http://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). _lme4: Linear mixed-effects models using Eigen and S4_. R package version 1.1-7.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. http://doi.org/10.18637/jss.v067.i01

Beilock, S. L., & Goldin-Meadow, S. (2010). Gesture Changes Thought by Grounding It in Action. *Psychological Science*, *21*(11), 1605–1610. http://doi.org/10.1177/0956797610385353

Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C., & Small, S. L. (2008). Sports experience changes the neural processing of action language. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(36), 13269–73. http://doi.org/10.1073/pnas.0803424105

Bell, C. C., & Grant, K. (1989). Corollary discharge inhibition and preservation of temporal information in a sensory nucleus of mormyrid electric fish. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *9*(3), 1029–44.

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Perceptual modulation of self-produced stimuli: The role of spatio-temporal prediction. *J Cogn Neurosci*, *11*, 551–9.

Brooks, N. B., Barner, D., Frank, M., & Goldin-Meadow, S. (2017). The Role of Gesture in Supporting Mental Representations: The Case of Mental Abacus Arithmetic. *Cognitive Science*. http://doi.org/10.1111/cogs.12527

Cardoso-Leite, P., Mamassian, P., Schütz-Bosbach, S., & Waszak, F. (2010). A new look at sensory attenuation. Action-effect anticipation affects sensitivity, not response bias. *Psychological Science*, *21*(12), 1740–5. http://doi.org/10.1177/0956797610389187

Cavanaugh, J., Berman, R. A., Joiner, W. M., & Wurtz, R. H. (2016). Saccadic Corollary Discharge Underlies Stable Visual Perception. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *36*(1), 31–42. http://doi.org/10.1523/JNEUROSCI.2054-15.2016

Chandrasekharan, S., Athreya, D., & Srinivasan, N. (2010). Investigating prospective control: When and how actions supporting cognition are generated. *International Journal of Mind, Brain, & Cognition*, *1*(1).

Chisholm, J. D., Risko, E. F., & Kingstone, A. (2014). From gestures to gaming: visible embodiment of remote actions. *Quarterly Journal of Experimental Psychology (2006)*, *67*(3), 609–24. http://doi.org/10.1080/17470218.2013.823454

Chrastil, E. R., & Warren, W. H. (2012). Active and passive contributions to spatial learning. *Psychonomic Bulletin & Review*, *19*(1), 1–23. http://doi.org/10.3758/s13423-011-0182-x

Chrastil, E. R., & Warren, W. H. (2013). Active and passive spatial learning in human navigation: acquisition of survey knowledge. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(5), 1520–37. http://doi.org/10.1037/a0032382

Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General*, *137*(4), 706–23. http://doi.org/10.1037/a0013157

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General*, *140*(1), 102–16. http://doi.org/10.1037/a0021790

Davidson, P. R., & Wolpert, D. M. (2005). Widespread access to predictive models in the motor system: a short review. *Journal of Neural Engineering*, *2*(3), S313–S319. http://doi.org/10.1088/1741-2560/2/3/S11

De Maeght, S., & Prinz, W. (2004). Action induction through action observation. *Psychological Research*, *68*(2–3), 97–114. http://doi.org/10.1007/s00426-003-0148-3

Desantis, A., Mamassian, P., Lisi, M., & Waszak, F. (2014). The prediction of visual stimuli influences auditory loudness discrimination. *Experimental Brain Research*, *232*(10), 3317–3324. http://doi.org/10.1007/s00221-014-4001-2

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science (New York, N.Y.)*, *255*(5040), 90–2.

Easton, R. D., & Sholl, M. J. (1995). Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 500–483.

Feyereisen, P., & Havard, I. (1999). Mental imagery and production of hand gestures while speaking in younger and older adults. *Journal of Nonverbal Behavior*, *23*(2), 153–171. http://doi.org/10.1023/A:1021487510204

Frank, D. J., & Macnamara, B. N. (2017). Does the Acquisition of Spatial Skill Involve a Shift From Algorithm to Memory Retrieval? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(12), 1845–1856. http://doi.org/10.1037/xlm0000429

Goldin-Meadow, S. (2017, January). Using our hands to change our minds. *Wiley Interdisciplinary Reviews: Cognitive Science*. http://doi.org/10.1002/wcs.1368

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining Math: Gesturing Lightens the Load. *Psychological Science*, *12*(6), 516–522. http://doi.org/10.1111/1467-9280.00395

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, *8*(6), 280–5. http://doi.org/10.1016/j.tics.2004.04.001

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: gestures as simulated action. *Psychonomic Bulletin & Review*, *15*(3), 495–514.

Jeannerod, M. (2001). Neural Simulation of Action: A Unifying Mechanism for Motor Cognition. *NeuroImage*, *14*(1), S103–S109. http://doi.org/10.1006/nimg.2001.0832

Kasess, C. H., Windischberger, C., Cunnington, R., Lanzenberger, R., Pezawas, L., & Moser, E. (2008). The suppressive influence of SMA on M1 in motor imagery revealed by fMRI and dynamic causal modeling. *NeuroImage*, *40*(2), 828–837. http://doi.org/10.1016/j.neuroimage.2007.11.040

Kirsh, D. (1995a). Complementary strategies: why we use our hands when we think. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (pp. 212–217).

Kirsh, D. (1995b). The intelligent use of space. *Artificial Intelligence*, *73*(1–2), 31–68. http://doi.org/10.1016/0004-3702(94)00017-U

Kirsh, D., Caballero, R., & Cuykendall, S. (2012). When doing the wrong thing is right. *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, 1786–1791.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*(4), 513–549. http://doi.org/10.1016/0364-0213(94)90007-8

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial Updating of Self-Position and Orientation During Real, Imagined, and Virtual Locomotion. *Psychological Science*, *9*(4), 293–298. http://doi.org/10.1111/1467-9280.00058

Knuf, L., Aschersleben, G., & Prinz, W. (2001). An analysis of ideomotor action. *Journal of Experimental Psychology: General*, *130*(4), 779–798. http://doi.org/10.1037//0096-3445.130.4.779

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492–527. http://doi.org/10.1037/0033-295X.95.4.492

McNeill, D. (1992). *Hand and mind : what gestures reveal about thought*. University of Chicago Press.

Moreau, D. (2013). Motor expertise modulates movement processing in working memory. *Acta Psychologica*, *142*(3), 356–61. http://doi.org/10.1016/j.actpsy.2013.01.011

Novack, M. A., & Goldin-Meadow, S. (2017). Gesture as representational action: A paper about function. *Psychonomic Bulletin and Review*, *24*(3), 652–665. http://doi.org/10.3758/s13423-016-1145-z

Pinheiro, J., & Bates, D. (2010). *Mixed-Effects Models in S and S-PLUS*. Springer Science & Business Media.

Popescu, S. T., & Wexler, M. (2012). Spontaneous body movements in spatial cognition.

Frontiers in Psychology, *3*, 136. http://doi.org/10.3389/fpsyg.2012.00136

Poulet, J. F. A., & Hedwig, B. (2006). The cellular basis of a corollary discharge. *Science (New York, N.Y.)*, *311*(5760), 518–22. http://doi.org/10.1126/science.1120847

Presson, C. C., & Montello, D. R. (1994). Updating after rotational and translational body movements: coordinate structure of perspective space. *Perception*, *23*(12), 1447–1455. http://doi.org/10.1068/p231447

R Core Team. (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1157–1165. http://doi.org/10.1037/0278-7393.15.6.1157

Risko, E. F., & Gilbert, S. J. (2016, September). Cognitive Offloading. *Trends in Cognitive Sciences*. http://doi.org/10.1016/j.tics.2016.07.002

Risko, E. F., Medimorec, S., Chisholm, J., & Kingstone, A. (2014). Rotating with rotated text: a natural behavior approach to investigating cognitive offloading. *Cognitive Science*, *38*(3), 537–64. http://doi.org/10.1111/cogs.12087

Roskos-Ewoldsen, B., McNamara, T. P., Shelton, A. L., & Carr, W. (1998). Mental representations of large and small spatial layouts are orientation dependent. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(1), 215–226. http://doi.org/10.1037/0278-7393.24.1.215

Schwartz, D. L., & Black, T. (1999). Inferences through imagined actions: Knowing by simulated doing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 116–136. http://doi.org/10.1037/0278-7393.25.1.116

Shepard, R. N., & Hurwitz, S. (1984). Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition*, *18*(1–3), 161–193. http://doi.org/10.1016/0010-0277(84)90024-6

Shorrock, S. T., & Isaac, A. (2010). Mental Imagery in Air Traffic Control. *The International Journal of Aviation Psychology*, *20*(4), 309–324. http://doi.org/10.1080/10508414.2010.487008

Sommer, M. A., & Wurtz, R. H. (2002). A Pathway in Primate Brain for Internal Monitoring of Movements. *Science*, *296*(5572), 1480–1482. http://doi.org/10.1126/science.1069590

Sommer, M. A., & Wurtz, R. H. (2008). Brain circuits for the internal monitoring of movements. *Annual Review of Neuroscience*, *31*, 317–38. http://doi.org/10.1146/annurev.neuro.31.060407.125627

Sperry, R. W. (1950). Neural Basis of the Spontaneous Optokinetic Response Produced by Visual Inversion. *Journal of Comparative and Physiological Psychology*, *43(6)*(6), 482–489. http://doi.org/10.1016/S0950-5601(54)80044-X

Trafton, J. G., Trickett, S. B., Stitzlein, C. A., Saner, L., Schunn, C. D., & Kirschenbaum, S. S. (2006). The Relationship Between Spatial Transformations and Iconic Gestures. *Spatial Cognition & Computation*, *6*(1), 1–29. http://doi.org/10.1207/s15427633scc0601_1

Vingerhoets, G., De Lange, F. P., Vandemaele, P., Deblaere, K., & Achten, E. (2002). Motor imagery in mental rotation: An fMRI study. *NeuroImage*, *17*(3), 1623–1633. http://doi.org/10.1006/nimg.2002.1290

von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip - Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Die Naturwissenschaften*, *37*(20), 464–476. http://doi.org/10.1007/BF00622503

Wang, R. F., & Simons, D. J. (1999). Active and passive scene recognition across views. *Cognition*, *70*(2), 191–210. http://doi.org/10.1016/S0010-0277(99)00012-8

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*, *68*(1), 77–94. http://doi.org/10.1016/S0010-0277(98)00032-8

Wexler, M., & van Boxtel, J. J. A. (2005). Depth perception by the active observer. *Trends in Cognitive Sciences*, *9*(9), 431–438. http://doi.org/10.1016/j.tics.2005.06.018

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications.

Wohlschläger, A. (2001). Mental object rotation and the planning of hand movements. *Perception & Psychophysics*, *63*(4), 709–718. http://doi.org/10.3758/BF03194431

Wohlschläger, A., & Wohlschläger, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(2), 397–412. http://doi.org/10.1037/0096-1523.24.2.397

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), 729–732.

Wraga, M., Creem, S. H., & Proffitt, D. R. (2000). Updating displays after imagined object and viewer rotations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 151–168. http://doi.org/10.1037/0278-7393.26.1.151

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research*, *48*(20), 2070–2089. http://doi.org/10.1016/j.visres.2008.03.021

Zacks, J. M. (2008). Neuroimaging studies of mental rotation: a meta-analysis and review. *Journal of Cognitive Neuroscience*, *20*(1), 1–19. http://doi.org/10.1162/jocn.2008.20013