Cognitive Influences on Self-Rotation Perception

Daniel R. Berger

Max Planck Institute
for Biological Cybernetics
Spemannstr. 38
72076 Tübingen, Germany
daniel.berger@tuebingen.mpg.de

Markus von der Heyde

Bauhaus-Universität Weimar
Steubenstraße 6a
99421 Weimar, Germany
markus.von.der.heyde
@scc.uni-weimar.de

Heinrich H. Bülthoff

Max Planck Institute
for Biological Cybernetics
Spemannstr. 38
72076 Tübingen, Germany
heinrich.buelthoff@tuebingen.mpg.de

Abstract

In this study we examined the types of information that can influence the perception of upright (yaw) rotations. Specifically, we examined the influence of stimulus magnitude, task-induced attention and awareness of inter-sensory conflicts on the weights of visual and body cues.

Participants had to reproduce rotations that were presented as simultaneous physical body turns (via a motion platform) and visual turns displayed as a rotating scene. During the active reproduction stage, conflicts between the body and visual rotations were introduced by means of gain factors. Participants were instructed to reproduce either the visual scene rotation or the body rotation. After each trial participants reported whether or not they had perceived a conflict.

We found significant influences of the magnitude of the rotation, attention condition (instruction to reproduce platform or scene rotation), and reported awareness of a sensory conflict during the reproduction phase. Attention had a larger influence on the response of the participants when they noticed a conflict compared to when they did not perceive a conflict. Attention biased their response towards the attended modality.

Our results suggest that not only the stimulus characteristics, but also cognitive factors play a role in the estimation of the size of a rotation in an active turn reproduction task.

1 Introduction

When moving in the world, we need to generate reliable estimates of our own location and movement in space from the available sensory signals. The amount of influence each sensory modality has on the resulting estimate can depend on the characteristics of the stimulus itself, but also on cognitive factors such as attention, task, and awareness of conflicts between different sensory signals.

The different senses provide signals with different reliabilities depending on the velocity and acceleration range of self-motion. Very slow or long-lasting movements cannot be sensed accurately by the vestibular system because of its high-pass transduction characteristics. The visual sense, on the other hand, may be more accurate for small and slow movements than for fast and large movements. For an optimal estimation of the movement, given the available sensory inputs, the brain should weight the different modalities according to their reliabilities, giving more weight to more reliable cues (Ernst & Bülthoff, 2004). Effects of sensory reliability on the weights have for example been found in the integration of visual and haptic modalities (Ernst, Banks & Bülthoff, 2000; Ernst & Banks 2002; Gepshtein & Banks, 2003). Current models of the perception of self-motion contain filters to model the dynamic response of the different sensors. By such filtering the amount of influence each cue has on the resulting percept in the model is also dependent on movement magnitude (Mergener, Schweigart, Kolev, Hlavacka & Becker, 1995; Zupan, Merfeld & Darlot, 2002).

While the influence of stimulus characteristics on perception has been addressed in many multisensory integration studies, few have looked at the influence of cognitive factors such as attention (Calvert & Thesen, 2004). The normal psychophysical procedure is to limit the influence of cognitive factors in careful experimental designs, instead of using them as an experimental variable. A few studies have studied cognitive effects on the perception of self-motion. Kitazaki and Sato (2003) investigated the influence of guided attention on the illusory perception of self-motion induced by visual stimuli (vection). A study by Lambrey & Berthoz (2004) recently addressed the influence of conflict awareness on multisensory integration in self-motion in a navigation task.

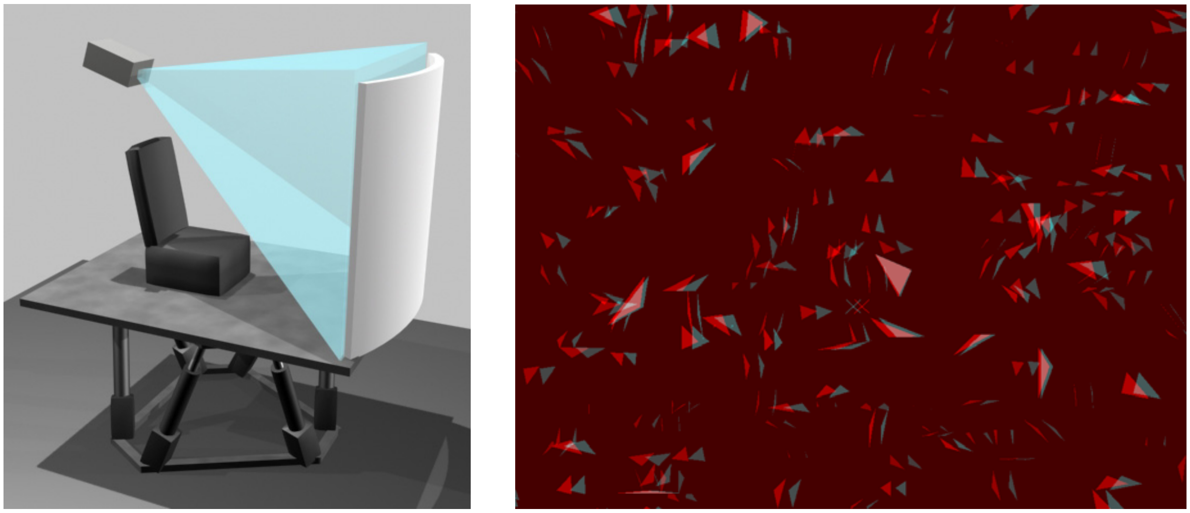

Figure 1: Left: Schematic view of the motion platform with projection screen. In the experiments described in this paper, a flat projection screen was used. Right: Visual pattern used in the experiments. The scene was viewed through red-cyan glasses. The motion of the scene was displayed with moving triangles with limited lifetime to avoid tracking of individual elements.

The aim of this study was to examine the influence of stimulus-related and cognitive factors on the integration of visual and body cues of self-motion in the perception of upright (yaw) rotations. We used different rotation magnitudes as a means to manipulate the reliabilities of visual and body cues for self-motion in a natural way. To investigate the influence of task-related attention, we used two different tasks, to "turn back the visual scene" and to "turn back the platform", which guided the attention of the participant to either one of the modalities. After each turn, the participant indicated whether a conflict had been perceived or not. This helped us to examine the influence of conflict perception on the responses.

2 Materials and Methods

Thirteen participants took part in this experiment (7 female and 6 male), randomly drawn from the MPI subject database. They gave their informed consent and were paid for participation.

Participants were seated on a Stewart platform equipped with a projection screen (see Figure 1, left), which had a projected visual field of 86° horizontal × 63° vertical. They viewed a field of random triangles with limited lifetime of 1.2 seconds, presented in 3D by using red-cyan colour anaglyphic viewers (Figure 1, right). The colours of background and triangles were adjusted to minimize inter-ocular crosstalk. The 3D triangles appeared further away than the screen (up to 2-3 meters). Triangle brightness was attenuated with distance to enhance the impression of depth. A fixation cross at eye height was drawn approximately 10 cm in front of the screen. Participants were told to remain fixated during all turns.

Physical body rotations were performed by rotating the motion platform. To rotate the observer in the visual scene, the virtual camera was rotated in this random triangle field in the direction opposite to the platform rotation. The screen then appeared as a window which turned with the observer, and through which a non-rotating outside scene could be observed. The rotation did not introduce any parallax motion in the image.

Previous studies have shown that the impression of self-motion is enhanced by both stereoscopic cues (Palmisano, 2002) and the fact that the moving scene is perceived as further away than fixation point and projection frame (Kitazaki & Sato, 2003). We chose to use random triangles instead of random dots to reduce the correspondence problems when viewing a random dot field in stereo with dots that all look the same.

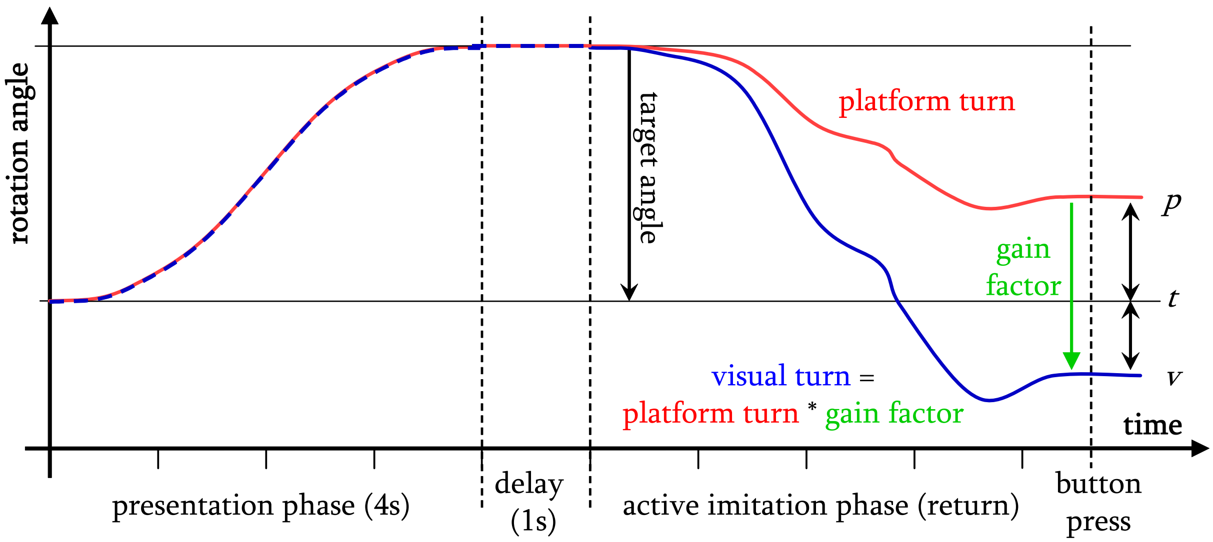

Figure 2: Example rotation trajectory of one trial. Red line is platform rotation, blue line visual scene rotation. During the presentation of a rotation (left), platform and visual scene turn equally far, using the same motion profile. During the active return (right), a gain factor causes the visual scene to turn more or less than the platform. The visual turn profile is equal to the platform turn profile multiplied by a constant gain value.
From the size of the active rotation, the weights of the visual cue and the body cues can be computed:
visual_weight = (t - p) / (v - p), body_cue_weight = 1 – visual_weight.

Every trial of the experiment began with the presentation of a yaw rotation. The participant was rotated passively both on the platform and within the visual scene, by 10°, 12.5°, 15°, 20°, 25° or 30° within 3 seconds. The visual and body turns had consistent profiles (same speed and final angle). We used a cosine velocity profile, which ensured smooth accelerations and decelerations (Figure 2, *presentation phase*). The duration of each turn was kept constant to prevent subjects from estimating the size of a rotation from its duration. Thus, larger rotations also involved higher velocities and accelerations.

After the presentation phase, a short delay of 1 second was inserted before the participant was allowed to turn back. During the delay the instruction to "turn back the visual scene" or "turn back the platform" was given aurally to remind the participant of the current task condition. We kept the delay short to prevent decay of the memory of the presented rotation.

After the presentation phase the participant was given control of the movement of platform and scene via a joystick. The task was to turn the platform (and/or scene, see below) for the same angle in the opposite direction of the previously presented rotation (Figure 2, *active imitation phase*). After completing the rotation, the participant pressed one of three joystick buttons (see below). The next rotation was then presented.

The experiment consisted of two blocks of 108 trials (six rotation angles × nine gain factors (see below), each condition measured twice). In one block participants were told to return the visual scene ("RVis" condition), in the other to return the platform ("RBody" condition). These instructions guided attention to either the visual cue or the (inertial) body cues of self-motion. Note that the only difference between the RVis and RBody conditions was the focus of attention. The presented stimuli were the same in both conditions.

During the active return, a gain factor (factorial difference in rotation velocity) was introduced between the visual rotation and the body rotation. This factor was chosen from 0.35, 0.5, 0.71, 0.84, 1.0, 1.19, 1.41, 2.0, or 2.83. Every combination of a presented rotation angle and a gain factor during reproduction was tested twice. Participants were explicitly told about the gain factors. After each return movement, the participant had to press one of three joystick buttons, depending on whether the visual scene had been perceived as faster, the body rotation had been perceived as faster, or no difference between the two had been noticed. This allowed analysis of the influence of perceiving a conflict on the responses. As participants had to compare the rotation magnitudes of both modalities, they could not completely ignore the unattended modality.

The yaw rotation range of the motion platform was limited to 100°. Since we did not want to reposition the platform between trials, an online trial selection algorithm was used. Depending on the current position of the platform, one trial was selected from the subset of feasible trials, which maximized the balance of trials presented so far, balanced over rotation angles and gain factors. If several trials were equally good, the choice among them was made randomly. Only if no possible trial was left, the platform was repositioned. This happened approximately one or two times per experimental block of 108 trials, depending on the participant's behavior. Repositioning was for example necessary if a participant had turned way too far and the platform had reached the limit of its motion range.

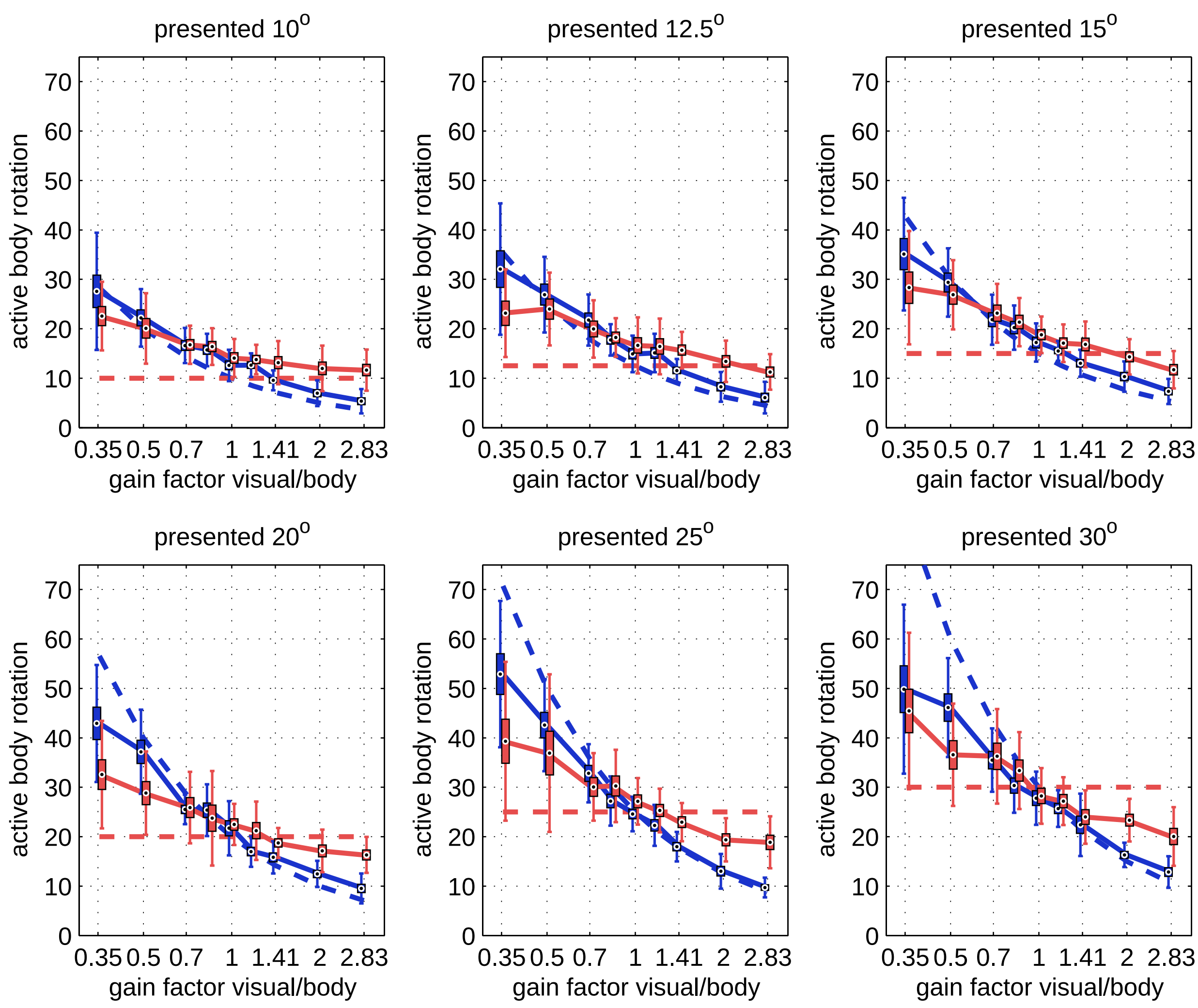

Figure 3: Response curves for RBody (red) and RVis (blue) attention conditions, for all 13 participants. Responses are plotted in active platform rotation angles over gain factors. Dashed red (horizontal) line: target angle when correctly turning back the platform; dashed blue (curved) line: target angle when correctly turning back the visual scene. For a gain factor of 0.35, for example, the platform has to be over-turned by a factor of 2.83 to return the visual scene accurately.

3 Results

For gain factors of 0.5 or less, or 2.0 or more, the responses in the RVis and RBody conditions differed from each other (Figure 3). Sensory conflicts caused by those gain factors were also often recognized by the participants.
Three-way analysis of variance (ANOVA) calculated on the raw data with attention condition (RVis or RBody), presented rotation angle, and gain factor as within-subject factors revealed a significant interaction of attention condition with the gain factors ($F(8,96) = 10.69$, $p < .001$***). Thus attention to one or the other cue – expressed in the task to either ”turn back the platform” or ”turn back within the visual scene” – has a significant influence on the response. In Figure 3, this shows in the difference of slopes of the red and blue curves. The ANOVA also reported significant main effects for angle and gain – responses were significantly different for different presented angles, and for different gain factors. The interactions between angle and gain and the three-way interaction of block, angle and gain were also significant. For the complete ANOVA results, see Table 1.

Table 1: 3-way ANOVA results on raw responses. 'Attention' denotes the two attention conditions (RVis coded as 1, RBody coded as 2 for the ANOVA). 'Gain' is the set of gain factors, 'Angle' the set of target angles.

Condition	***SS***	***df***	***MS***	***F***	***p***
Attention	175.2	1, 12	175.2	0.49	.498
Angle	40437.6	5, 60	194.0	193.96	<.001***
Attention, Angle	60.5	5, 60	12.1	0.45	.813
Gain	79212.2	8, 96	9901.5	88.37	<.001***
Attention, Gain	7036.2	8, 96	879.5	10.69	<.001***
Angle, Gain	4705.0	40, 480	117.6	5.82	<.001***
Attention, Angle, Gain	1093.1	40, 480	20.5	1.33	.089

3.1 Analysis of Visual Weight and Offset

To assess the relative weights of visual cue and body cues of self-motion in the different conditions, linear functions over gain factors were fitted to the responses. By this method we calculated cue weights and response offsets for the different participants, rotation angles and attention conditions separately. Two free parameters were fitted by using least-squares fitting (via Matlab *fminsearch*): a cue weighting parameter w_v (visual weight) to interpolate linearly between the RVis and the Rbody target curves, and a constant offset c. For a given presented target angle α, visual target rotation angle t_v, body rotation angle t_b, gain factors g_i and response angle r, the visual weight w_v and the offset c were derived by minimizing

$$\sum_i ((w_v \cdot t_v(\alpha, g_i) + (1 - w_v) \cdot t_b(\alpha, g_i) + c) - r(\alpha, g_i))^2$$

Figure 4 shows the resulting best-fitting parameters in different conditions for individual participants.

Both for the visual weight and for the offset 2-way ANOVA analyses were computed, with attention condition (RBody, RVis) and presented rotation angle as within-subject factors.
We found a mean visual weight for the RVis condition of 0.73, whereas it was 0.39 for the RBody condition, and this difference was highly significant ($F(1,12)=15.24$, $p=.002$**). The visual weight also depended significantly on stimulus magnitude (target angle) ($F(5,60)=6.03$, $p<.001$***), with a higher visual weight for small rotations (Figure 4, left). In the RVis condition, the mean visual weight over all participants dropped from 0.89 for rotations of 10° to 0.51 for rotations of 30°, and in the RBody condition, it dropped from 0.47 for rotations of 10° to 0.33 for rotations of 30°.
The offset also depended significantly on stimulus magnitude ($F(5,60)=45.61$, $p<.001$***). Small rotations were over-estimated and large rotations under-estimated (Figure 4, right). For rotations of 10°, participants turned on average too far by 2.5° in the RVis and by 4.6° in the RBody condition. For rotations of 30°, they did not turn far enough (-3.5° in the RVis and -1.5° in the RBody condition). Complete ANOVA results are shown in Table 2.

Table 2: 2-way ANOVA results of model parameters 'visual weight' and 'offset', in dependency of stimulus parameters block and angle. 'Attention' denotes the two attention conditions (RVis coded as 1, RBody coded as 2 for the ANOVA). 'Angle' is the set of target angles.

variable	***condition***	***SS***	***df***	***MS***	***F***	***p***
Visual weight	Attention	4.64	1, 12	4.64	15.24	.002**
	Angle	1.11	5, 60	0.22	6.03	<.001***
	Attention, Angle	0.31	5, 60	0.06	1.72	.144
offset	Attention	144.2	1, 12	144.2	4.89	.047*
	Angle	725.1	5, 60	145.0	45.61	<.001***
	Attention, Angle	5.44	5, 60	1.09	0.67	.648

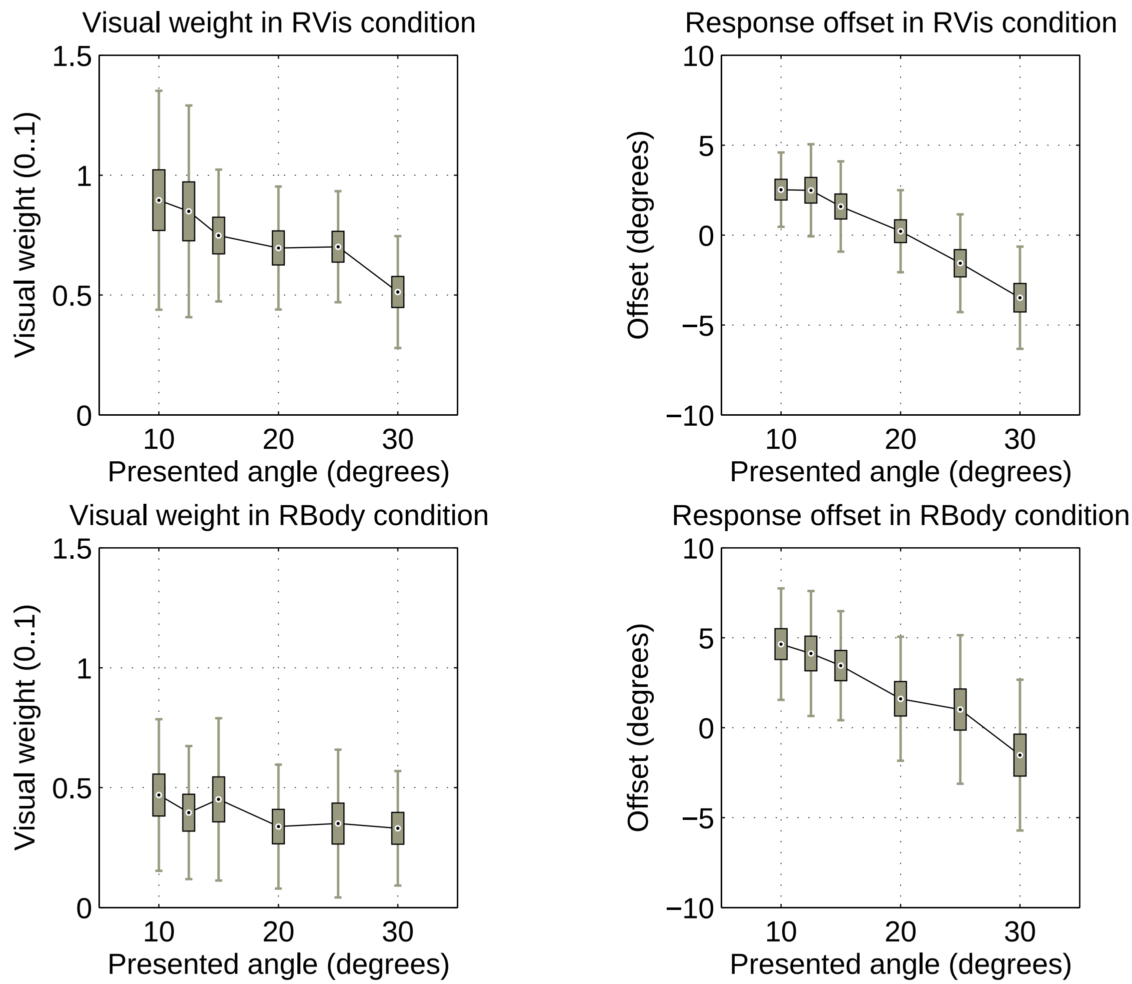

Figure 4: Offsets and visual weights of best fit response functions for all thirteen participants. The upper two subfigures show the results of the "turn back within the visual scene" condition, the lower two show the results of the "turn back platform" condition. A visual weight larger than 1 can result if the response over gain factors is steeper than the visual target curve (blue dashed curve in Figure 3). Visual weights larger than 1 have been measured for some participants particularly when they were instructed to turn back in the visual scene (RVis condition), and might happen if participants try to over-compensate for the influence of the body rotation on their response.

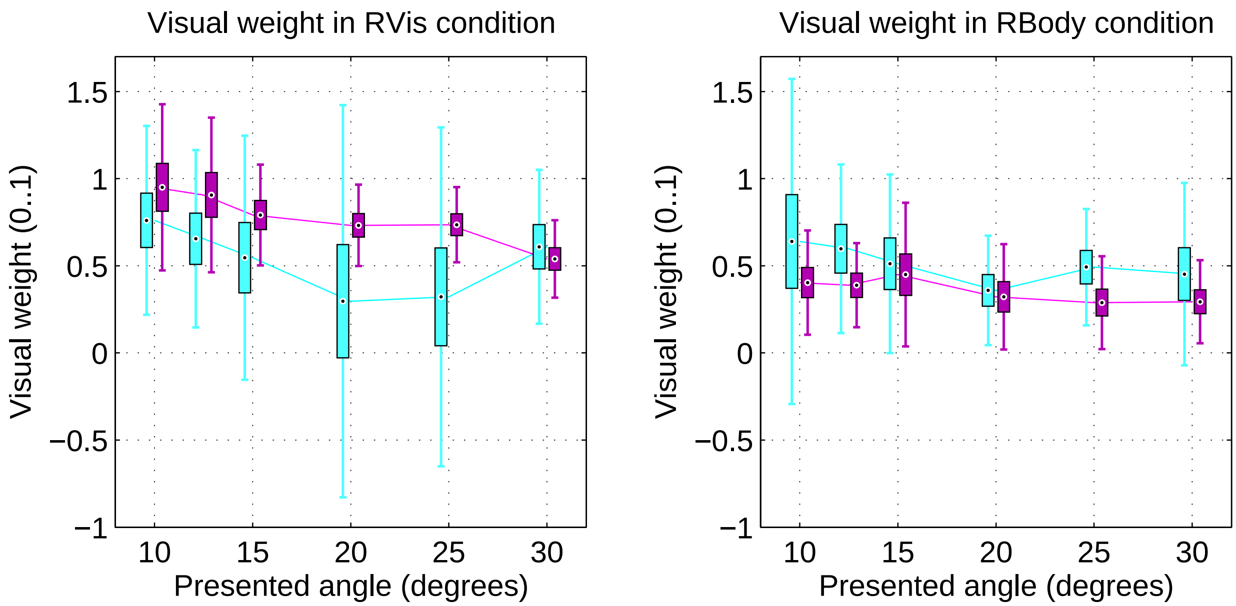

Figure 5: Visual weights of the participants' responses for different presented angles in the two attention conditions, and whether or not they noticed a conflict. Light cyan: no conflict detected, Dark magenta: conflict detected.

3.2 Analysis of the Influence of Conflict Awareness

After each trial, the participants had to press one of three joystick buttons, depending on whether during active return the rotated angle of the ignored modality (the visual rotation in the RBody condition, and the body rotation in the RVis condition) had appeared to them smaller, equal (no conflict) or larger than the angle of the returned modality.

To analyze the influence of the awareness of a conflict on the responses, we split the data in two sets, depending on whether or not the participant had perceived a conflict between visual and body rotation during active return.

The splitting causes an accumulation of "conflict detected" trials for large or small gain factors, and an accumulation of "no conflict detected" trials for gain factors near 1, because large conflicts are easier to detect than small conflicts. This makes statistical comparisons between the two sets difficult. Therefore we again fitted linear functions over gain factors, this time for "conflict detected" and "no conflict detected" trial subsets individually, and retrieved visual weights and offset values. We then used three-way ANOVA with attention condition (RBody, RVis), presented rotation angle, and detection of a conflict as within-subject factors.

One participant had to be excluded from this analysis because she had a strong bias to respond "conflict detected", which made curve fitting impossible in some conditions.

We found a significant influence of conflict awareness on the cue weights. If participants were not aware of the conflict, the visual weights were close to 0.5 in both RBody and RVis conditions (0.51 in RBody and 0.53 in RVis), whereas they were clearly separated as soon as a conflict was noticed (0.36 in RBody and 0.78 in RVis condition). This effect is significant in the ANOVA as interaction between attention condition and perception of a conflict ($F(1,11) = 5.6$, $p=.037$*). Figure 5 shows the visual weight response distributions for different presented angles.
We did not find any significant effects of conflict perception on the response offset.

4 Discussion

This experiment showed that the weights of visual and vestibular/proprioceptive modalities in the sensor integration process for the perception of self-rotation were influenced both by stimulus characteristics (magnitude of the rotation) and by cognitive factors (task-induced attention on one modality and awareness of conflicts).

A dependency of the visual weight on the rotation magnitude – higher visual weight for small rotations – is expected by a sensor integration model in which the weight of each cue depends on its reliability (Ernst & Bülthoff, 2004), as the reliability of the vestibular sense is lower for small and slow rotations close to threshold than for rotations of larger magnitude. The smallest of our rotations (3.3°/sec average) were quite close to vestibular threshold compared to the largest rotations (10°/sec average). The threshold for such cosine yaw rotations of about 3 seconds duration is around 1.5°/sec in darkness and 0.55°/sec in the presence of a visual target (Benson, Hutt & Brown, 1989), but the variance between participants is high. The threshold for visual motion is lower – it is in the range of 0.3°/s (as used in the model by (Mergener et al., 1995)).
Attention is thought to bias the competition of different sensory inputs to reach higher brain areas (Desimone, 1998), and may thus also gate the inputs from lower brain areas processing single modalities to higher multimodal areas. It is therefore plausible that guiding attention to one of several modalities could have an effect on the influence this modality has on an area which integrates signals from several modalities. The brain could even use this attentional gating for the selection of appropriate inputs to perform a certain task. Our findings support this hypothesis, as attending to one cue significantly increased the weight of that cue in the resulting percept.

One important issue in multisensory integration is the identification of corresponding information. Only stimuli which are representing the same variable should be integrated. For example, the decision whether a given sound and a given visual stimulus belong together can often only be made if the stimulus has been identified to belong to a certain object category. Studies on multisensory integration in the superior colliculus provide evidence that the decision to integrate or not to integrate signals of different modalities is controlled by inputs from multimodal cortical areas (Wallace, 2004). These signals could be related to top-down attention, which would depend on the identification of the stimulus identity in higher cortical areas which are also involved in awareness. If a conflict is detected, top-down attention could be responsible for selecting the signals appropriate for the current task. Discordant stimuli may give rise to conflicting neural representations, inducing competition. It is known that attention has a large effect in biasing the competition (Desimone, 1998). Such processes would explain why we found a stronger integration of the multimodal signals when no conflict was detected, compared to trials in which participants were aware of a conflict between the two cues.

The effect of the awareness of a conflict on the response could also be interpreted differently. It could be that detection of a conflict and amount of multisensory integration are correlated because they depend on a common underlying mechanism. A possible model is shown in Figure 6. Assume that the perceived rotation angle in each modality is taken from a Gaussian distribution around the actual angle. Then, in trials in which the angles in the two modalities are perceived more alike (dotted lines), both the reproduced angle will be more towards the centre of the two distributions, and the conflict is detected less often, because it is very small. In trials in which the angles in the two modalities are perceived as very different (dashed lines), the reproduced angle would be more biased to one side (possibly determined by attention) and at the same time the difference between the two perceived angles would be large and typically above threshold for conflict awareness.

Because of this alternative interpretation we can not count our results as unequivocal evidence for a direct influence of conflict awareness on multisensory integration. Further experiments are needed to address this issue.

In conclusion, the results of the self-rotation perception experiment support models of multisensory integration in which the weight of each modality depends on its reliability. The experiment also showed that the processes of sensory integration for the perception of self-rotation in the human brain are not only governed by the stimulus characteristics alone, but also by task-related (top-down) attention, which can bias the influences of the relevant modalities. We found that the amount by which the influence of the modalities on the response is modulated by attention is correlated with the awareness of a conflict. The results are consistent with the idea that becoming aware of a conflict might trigger sensory biasing by means of top-down attention to generate robust behaviour under conflicts.

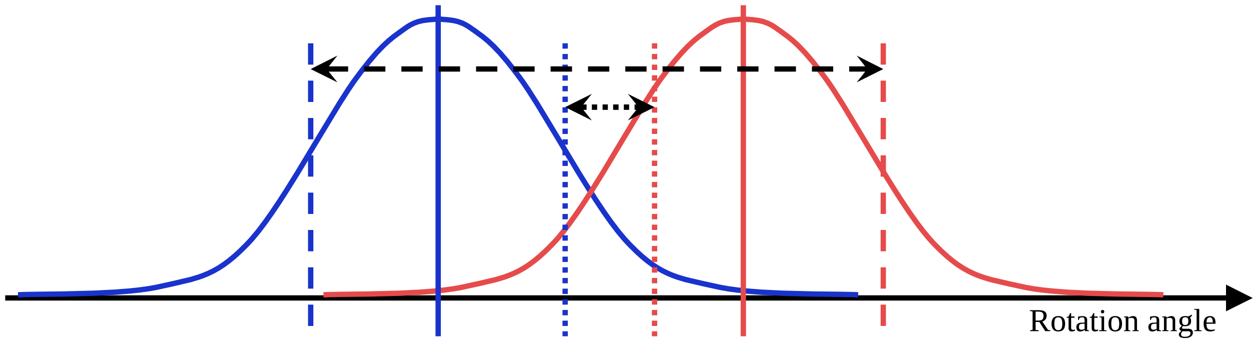

Figure 6: Hypothetical process for responses and awareness of conflicts. Blue: visual rotations, Red: body rotations. For two concurrently presented rotations with different angles (blue and red solid lines) the participant perceives rotation angles in the vicinity, presumably with a certain probability distribution. For a case in which the perceived rotation angles are similar (dotted lines), one would expect both the conflict to be less perceptible, and the response to be closer to the centre of the two distributions. If the perceived rotation angles are further apart (dashed lines), both the perceived conflict would be larger, and easier to detect, and the response would be biased to one of the sides. Attention could determine which side is selected.

5 References

Benson, A. J., Hutt, E. C. B. & Brown, S. F. (1989). Thresholds for the perception of whole body angular movement about a vertical axis. *Aviation, Space and Environmental Medicine*, 60, 205-213.

Calvert, G. A. & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology – Paris*, 98:191-205.

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Phil. Trans. R. Soc. Lond. B* 353:1245-1255.

Ernst, M. O. & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences* 8(4):162-169.

Ernst, M. O. & Banks, M. S. (2002). Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion. Nature 415, 429-433.

Ernst, M. O., Banks M. S. & Bülthoff H. H. (2000). Touch can change visual slant perception. Nature Neuroscience 3 (1), 69-73.

Gepshtein, S. & Banks, M. S. (2003). Viewing geometry determines how vision and haptics combine in size perception. *Current Biology*, 13:483-488.

Kitazaki, M. & Sato, T. (2003). Attentional modulation of self-motion perception. *Perception*, 32:475-484.

Lambrey, S. & Berthoz, A. (2003). Combination of conflicting visual and non-visual information for estimating actively performed body turns in virtual reality. *International Journal of Psychophysiology*, 50:101-115.

Mergener, T., Schweigart, G., Kolev, O., Hlavacka, F., & Becker, W. (1995). Visual-vestibular interaction for human ego-motion perception. In: T. Mergener and F. Hlavacka (Eds.), *Multisensory control of Posture*, Plenum Press, New York.

Palmisano, S. (2002). Consistent stereoscopic information increases the perceived speed of vection in depth. *Perception*, 31:463-480.

Wallace, M. T. (2004). The development of multisensory processes. *Cognitive Processes*, 5: 69-83.

Zupan, L. H., Merfeld, D. M., & Darlot, C. Using sensory weighting to model the influence of canal, otolith and visual cues on spatial orientation and eye movements. *Biological Cybernetics*, 86:209-230.

Acknowledgements

We would like to thank Ian Thornton, Manuel Vidal, Marc Ernst, Astros Chatziastros and Bernhard Riecke for valuable comments on the experiment, data analysis and an earlier draft of the manuscript.
This work has been supported by the Max Planck Society and the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 550).