Biases in the Visual Perception of Heading

Master’s Thesis
to obtain the academic title
Master of Science (M.Sc.) Psychology
at the University of Twente

Max Lukas Kurtz
m.l.kurtz@student.utwente.nl
s1879596
Submitted: 11.09.2017

External supervisor:
Dr. Ksander de Winkel
Max Planck Institute for Biological Cybernetics

First supervisor:
Prof. Dr. Ing. Willem Verwey
University of Twente

Second supervisor:
Dr. Rob van der Lubbe
University of Twente
Abstract

Earlier research has shown that heading can be accurately estimated from visual cues. Nevertheless, a multitude of studies has shown that people commit constant errors when estimating heading from visual cues. The exact origin of these biases remains unclear despite many research efforts. The visual stimuli used in past research on heading estimation varied considerably – a factor which might have systematically influenced heading accuracy. Therefore, we examined the possible effect of divergent stimuli characteristics on heading estimation in the current study. Measurements of twenty participants (12 females) with a mean age of 25.80 years (SD = 3.53) were obtained during the study. Participants were shown stimuli depicting horizontal linear motion in the ground plane and were asked to judge the direction of heading as well as to provide vection ratings. Stimuli shown differed in characteristics related to Field of View (FOV), disparity, motion profile, and layout of the scene. A model explaining hypothesized relationships and effect on estimation bias with a partial mediating effect of vection was proposed and estimated via path analysis. Findings indicated that bias away from the fore-aft axis found in this study could not be explained by visual properties alone. FOV, disparity and scene had only in combination with certain stimuli headings a marginal effect on magnitude of bias, but not its direction. Vection had no mediating effect. However, all characteristics had an effect on vection ratings. Visual properties thus marginally affect heading estimation, but exact source(s) of biases remain an objective for future research.
Biases in the Visual Perception of Heading

Heading is defined as the direction of horizontal linear self-motion. Heading estimation is essential to everyday tasks such as locomotion and navigation. The Central Nervous System (CNS) is able to estimate heading from inertial cues, by means of the vestibular system and several other sensory cells distributed throughout the body, but also from visual cues (Howard, 1982).

Visually, we estimate heading from optic flow, which was first described by Gibson in 1950 and has since been subject of a large number of studies (see Lappe, Bremmer, & Van Den Berg, 1999, for a review). Optic flow refers to the pattern of motion generated as reflections of objects move across the retina when we move ourselves (Gibson, 1950). As we move on a straight path, this pattern of visual motion radially expands from a singular point along the direction of heading, the Focus of Expansion (FOE). Heading can be estimated by localizing this point. When there is perfectly lateral movement, the direction of heading cannot be derived from the FOE, but can be estimated from direction of vectors represented in the optic flow.

Although errors in heading estimation found in research are generally not more than a few degrees (Telford, Howard, & Ohmi, 1995; Warren & Kurtz, 1999; Warren, Morris, & Kalish, 1988), systematic biases are found across studies (Crane, 2012; Cuturi & MacNeilage, 2013; De Winkel, Katliar, & Bülthoff, 2015, 2017). The exact source(s) of these biases remain unclear, but it has been proposed that these biases have a neurophysiological cause: Visual optic flow cues are mainly processed in the dorsal medial superior temporal (MSTd) area of the brain and neural populations in this area show an overrepresentation of lateral preferred directions (Gu, Fetsch, Adeyemo, DeAngelis, & Angelaki, 2010; Saito et al., 1986). It has been shown that such a distribution of preferred directions can lead to a bias away from the fore-aft axis and to a more sensitive discrimination of headings laying near
straight ahead (Gu et al., 2010). From an evolutionary standpoint, this sensitivity could be functionally relevant because we mostly move forward in our daily environments (Cuturi & MacNeilage, 2013). However, this hypothesis cannot explain all of the biases found across past studies, because the nature of the biases reported in the literature is highly variable. Some studies show biases towards the fore-aft axis, i.e., underestimation, (D'Avossa & Kersten, 1996; De Winkel et al., 2015; Van Den Berg & Brenner, 1994a) while other studies show biases away from the fore-aft axis, i.e., overestimation, (Cuturi & MacNeilage, 2013; Telford & Howard, 1996; Warren, 1976). Therefore, these biases might have extraneous origins. De Winkel and colleagues (2015) suggest that the observed variability could stem from divergent characteristics of the visual stimuli used in different studies. In the following section, we list properties of the visual stimuli used across studies and discuss how these characteristics may affect bias.

**Field of View**

A salient difference between studies is the variability in Field of View (FOV). The equipment used in past studies varied between head-mounted displays (Hummel, Cuturi, MacNeilage, & Flanagin, 2016), projection screens (De Winkel et al., 2015; Gu, DeAngelis, & Angelaki, 2007), LCD screens (Crane, 2014; Cuturi & MacNeilage, 2013) and plain monitors (D'Avossa & Kersten, 1996; Warren & Kurtz, 1999), varying in resolution and display size. Due to these differences, the Field of View also ranged from 40° by 32° (Warren & Kurtz, 1999) up to 160° by 100° (Johnston, White & Cumming, 1973), horizontal by vertical respectively, with most current studies employing a FOV of around 100° by 80° (Butler, Campos, & Bülthoff, 2014; Crane, 2014; Cuturi & MacNeilage, 2013). The FOV may affect bias in heading estimation indirectly by means of the FOE: When the FOV is large enough, the FOE is visible. However, when the FOE is outside the FOV, heading cannot be estimated by locating the FOE. In this case the CNS can still estimate heading by
triangulation of vectors based on reference points in the optic flow (Koenderink & Van Doorn, 1987), but it has been shown that heading estimates tend to become less precise (Crowell & Banks, 1993; Gu et al., 2007; MacNeilage, Banks, Berger, & Bülthoff, 2010). For studies with smaller FOV, this means that physically, the range of FOEs lying within the bound of the FOV will become smaller, thus leading to more triangulation errors and to biased perceptions. Very small FOVs have already been shown to affect accuracy and to bias heading judgements towards the fore-aft axis (Li, Peli, & Warren, 2002; Warren & Kurtz, 1999). Consequently, it appears that FOV could affect bias in heading estimates by forcing the CNS to apply different estimation strategies, although the effect of FOV on nature of the bias has not yet been systematically studied.

**Binocular Disparity**

A second difference between studies is the availability of (binocular) disparity cues when displaying optic flow stimuli. Some studies presented stimuli without disparity either monocular (Banks, Ehrlich, Backus, & Crowell, 1996; Crowell & Banks, 1993) or binocular with the same image for both eyes (De Winkel, Weesie, Groen, & Werkhoven, 2010; Johnston et al., 1973). While other studies presented stimuli with disparity cues, using active stereo glasses (De Winkel et al., 2015; Fetsch, Turner, DeAngelis, & Angelaki, 2009) or anaglyph glasses (Crane, 2012; Butler, Smith, Campos, & Bülthoff, 2010). The availability of disparity cues in optic flow stimuli generates more information for the observer, especially in terms of depth order. The depth order gives information about the relative distance of objects in the environment and can be used to distinguish between self-motion and eye movement or random movements in the optic flow. The distinction between eye or random movement and self-motion can then improve localisation of the FOE and can thus be used to improve heading. In fact, it has been shown that stereoscopic depth cues can improve accuracy of heading estimation in noisy environments (Grigo & Lappe, 1998; Van Den Berg & Brenner,
BIASES IN THE VISUAL PERCEPTION OF HEADING

1994b). Interestingly, most studies reporting overestimation of heading used stereoscopic stimuli (Crane, 2012; Cuturi & MacNeilage, 2013; Telford & Howard, 1996) whereas studies without disparity primarily report underestimation (Johnston et al., 1973; Li et al., 2002; Warren & Kurtz, 1999). However, the availability of stereoscopic depth cues is not a perfect predictor for the direction of the bias (De Winkel et al., 2015; Warren, 1976), and its precise effect on bias in heading estimation still needs to be assessed experimentally.

Visual Scene

Various visual scenes have been used throughout studies, differing in content and layout. Most studies used a variation of random dot clouds or star fields as visual stimuli (Butler et al., 2010; Crane, 2012; Warren & Kurtz, 1999), but differed considerably in the specific objects that these clouds consisted of, using either Gaussian blobs (Butler et al., 2010), frontoparallel triangles (Crane, 2012; Cuturi & MacNeilage, 2013; Gu et al., 2010), circles (De Winkel, Grácio, Groen, & Werkhoven, 2010), round particles (De Winkel et al., 2015, 2017), or single pixels (Warren & Kurtz, 1999). Likewise, some studies included a ground plane (Royden, Banks, & Crowell, 1992; Warren & Kurtz, 1999; De Winkel et al., 2015, 2017) while others did not (Crane, 2012; Cuturi & MacNeilage, 2013; Gu et al., 2007).

Early research has shown that basic optic flow patterns are sufficient for perceiving heading (Warren & Hannon, 1988) and that heading judgements are largely independent of 3D-layout and density of dots (Warren et al., 1988). However, similar to the presence or absence of binocular disparity cues, the presence or absence of a ground plane affects the type and amount of information from which the observer can judge heading. Specifically, Koenderink and Van Doorn (1987) point out that reference points in the visual scene support distance judgement of self-motion when knowledge about the shape of the points is available. This is for example the case when the points lie in a plane. In this case, more reference points help to resolve uncertainty in the motion parallax and could possibly affect heading judgements. Li
and colleagues (2002) found that textured ground planes enhance accuracy in the case of very small FOVs. Moreover, heading estimation could be affected by ground planes because the horizon adds a prominent cue to the scene. For instance, Van Den Berg and Brenner (1994a) found a bias towards the fore-aft axis when the depth of the horizon was reduced. Most studies reporting a bias away from the fore-aft axis used clouds of dots (Crane, 2012; Cuturi & MacNeilage, 2013; Hummel et al., 2016) while studies reporting a bias towards the fore-aft axis mostly relied on displays with a ground plane (De Winkel, et al., 2015; Li et al., 2002; Warren & Kurtz, 1999). This dichotomy appears to be mostly independent of the availability of binocular disparity cues, underlining that the content of the visual scene could affect bias (but see: De Winkel et al., 2017; Warren, 1976).

**Motion Profile**

Another difference between studies is the type of motion profile used. The studies vary between profiles with constant velocities (Johnston et al., 1973; Li et al., 2002; Warren & Kurtz, 1999) and profiles with a variable velocity (Crane, 2012; Cuturi & MacNeilage, 2013; Telford & Howard, 1996). The latter, featuring acceleration in the beginning and deceleration towards the end of the display, may be considered more naturalistic, as self-motion in the real world shows similar characteristics (Butler, Campos, & Bülthoff, 2015). Butler and colleagues (2015) report that heading estimates are more precise for raised cosine bell velocity profiles than for constant velocity profiles. This difference appears to be also present in the further literature on heading biases. Most studies reporting overestimation use raised cosine bell velocity profiles (Crane, 2012; Cuturi & MacNeilage, 2013; Hummel et al., 2016) whereas most studies reporting underestimation use constant velocity profiles (Johnston et al., 1973; Li et al., 2002; Warren & Kurtz, 1999). However, there are again exceptions to this pattern (De Winkel et al., 2015, 2017; Warren, 1976), and the precise role of motion profile should be investigated experimentally.
Vection

Taken together, the factors mentioned above suggest another possible source of bias. Whereas FOV, disparity, and motion profile might have direct individual effects on heading perception, these factors have also been found to affect vection, which is the visually induced sensation of self-motion. When vection is rated higher, observers feel more present in the displayed scene and the scene feels more compelling to them (Riecke, Schulte-Pekum, Avraamides, Von Der Heyde, & Bülthoff, 2006). With more compelling vection, participants are more likely to perceive displays as self-motion instead of object motion (Palmisano, 1996). Stereoscopic displays have been shown to create more compelling vection (Palmisano, 1996, 2002; Ziegler & Roy, 1998), and motion profiles with variable velocity have been shown to more effectively induce vection than constant velocity profiles (Palmisano, Allison, & Pekin 2008). Moreover, stimulation in the visual periphery and therefore the size of the FOV strongly affect vection (Habak, Casanova, & Faubert, 2002). FOVs of around 60° have already been shown to suffice for the perception of self-motion (Andersen & Braunstein, 1985; Pretto, Ogier, Bülthoff, & Bresciani, 2009), but displays with larger FOVs or even floor projection were shown to improve vection nevertheless (Trutoiu, Mohler, Schulte-Pekum, & Bülthoff, 2008). The effect of different scene layouts on the experience of vection is not so clear: Scenes with a ground plane offer a clearer contrast between foreground and background stimuli, which was shown to affect vection (Brandt, Wist, & Dichgans, 1975; Nakamura & Shimojo, 1999), and more realistic scenes were shown to increase vection (Riecke et al., 2006). Because observers feel more present in the scene through more compelling vection, they could be able to better distinguish the optic flow as self-motion instead of object motion. This in turn could improve heading judgements and give vection a role as a partial mediator.
**Present study**

The review of the literature highlights several aspects: A great deal of research has been done in the field of heading perception, yielding a considerable variability in reported characteristics of bias in heading estimates. We identified several factors that may affect the nature of the visual bias: size of the FOV, presence or absence of binocular disparity, motion profile, and the layout of the scene. However, none of the identified factors can by itself explain the observed variability, as there are exceptions in all cases. We hypothesize that the size of the FOV, the presence or absence of binocular information, the motion profile, and also the layout of the scene may affect a potentially mediating variable: vection. Taken together, these factors yield a network of potential relations (Figure 1). In the present study, we aim to empirically assess the relations between these factors as well as the role of vection as a possible mediating variable. In an experiment, we systematically vary the above-mentioned factors while registering observers’ heading estimates and vection ratings for optic flow stimuli with various headings. Using path-analysis, we assess weights of and interactions between these factors. Ultimately, we aim to achieve a better understanding of the mechanisms underlying heading perception.

![Path model of hypothesized relations between the four visual properties and vection with the role of a partial mediator.](image-url)
Methods

Participants and design

The study was conducted in accordance with the Declaration of Helsinki. Participants were employees of the Max Planck Institute in Tübingen, Germany, or were recruited from the institute’s database. Written informed consent was obtained from all participants prior to participation. The experimental protocol and consent forms were approved by the ethical committee of the faculty of Behavioural, Management, and Social Sciences of the University of Twente in Enschede, the Netherlands (Request number 17217). Measurements of 20 participants (12 female) were obtained in the study. Participants had a mean age of 25.80 years (SD = 3.53), ranging from 22 to 38 years. All subjects had normal or corrected-to-normal vision. A repeated measures design was chosen, with participants experiencing all possible combinations of stimuli. Four properties with two levels each were manipulated, resulting in sixteen experimental conditions, the four properties were: (a) size of the FOV (restricted vs. unrestricted), (b) presence or absence of binocular disparity information, (c) motion profile (constant velocity vs. variable velocity) and (d) content of the visual scene (cloud vs. ground). Additionally, heading direction was varied as a covariate with sixteen different heading angles. Heading estimates as well as vection ratings were collected from participants.

Equipment

The experiment was conducted in the virtual environment of the BackproLab at the Max Planck Institute in Tübingen, Germany. The setup consisted of a large back projection screen, spanning over 2.16 m horizontally and 1.62 m vertically, and a single SXGA+ projector (Christie Mirage S+3K DLP) with a resolution of 1,400 by 1,050 pixels (Figure 2). Participants were seated at an eye-height of around 1.33 m and a distance of around 1.10 m in front of the projection screen (Figure 3). Stereoscopic displays were presented using active
3D-glasses (NVIDIA 3D Vision Pro LCD shuttered glasses). The resulting setup filled a field of view of around 90° horizontally by 50° vertically. During the experiment, participants’ heads were stabilized using a chin rest. The view of the participants was masked so that only the screen and devices for response collection were visible. A fixation cross was not implemented because it would lead to relative motion. As a result, people would be able to track certain elements of the scene, thus altering the nature of the task.

Figure 2. Experimental setup with stimuli on screen. During actual experiment lights were turned off to achieve better visibility of stimuli presented on screen.

Heading estimates of participants were collected using a physical pointer device. The pointer device consisted of a stainless steel rod of about 15 cm, which was connected to a potentiometer. Participants held the rod at one end and aligned the other end with the
perceived heading (Figure 3). The rod could be freely rotated on the horizontal plane and provided a < 0.1° resolution. Vection ratings were collected using a numerical keypad (Figure 3).

![Figure 3. Experimental setup with participant. Dial and keyboard for response collection can be seen in the front left.](image)

**Stimuli**

The stimuli simulated linear translation (i.e., motion in a straight line) in the horizontal plane with different headings. Heading direction and properties of the visual stimuli varied across trials.

FOV varied between full view and a restricted view. For the restricted view, the edges of the simulation were masked resulting in a FOV of 45° horizontal by 50° vertical. The stimuli were either presented monoscopically or stereoscopically using active 3D glasses.
Participants wore the glasses throughout the whole experiment. Translation either had a motion profile with constant velocity or a variable motion profile featuring acceleration and deceleration. The constant motion profile had a velocity of 0.075 m/s. The variable motion profile had a raised cosine bell in velocity which can be specified as

$$v = \frac{v_{\text{max}}}{2} \left(1 - \cos \omega t\right)$$

with $\omega = 2\pi f$, $f = 0.5$ Hz, and a maximum velocity $v_{\text{max}} = 0.15$ m/s. Both motion profiles lasted 2 s and both had a displacement of 0.15 m. With regard to the motion profiles, the stimuli were similar to the ones used by De Winkel and colleagues (2015, 2017). The content of the visual scenes varied between a cloud of random dots or a random dot ground (Figure 4). Both the random dot cloud and the random dot ground consisted of limited-lifetime particles with a lifetime of 1 s, randomly distributed in either a rectangular volume or in a plane around the observer. The rectangular volume had dimensions of 10 m x 10 m x 10 m and the plane spanned over 10 m x 10 m. Dot density in the cloud was 10 dots/m³ and dots had a size of 5 cm. Dot density in the plane was 50 dots/m² and dots had a size of 2.5 cm. Dimensions, size, and density of the dots were based on pilot sessions. They were chosen so that dot size, dot amount, and the resulting impression of depth were comparable across conditions. Dots in both scenes were bright grey on a dark grey background to improve stereoscopic display and prevent ghosting. Stimuli were generated using the Unity game engine (Unity Technologies, San Francisco, United States, version Unity 4.2.0f4) and were controlled during the experiment using MATLAB Simulink (The MathWorks Inc., Natick, United States, version R2016b).
In each experimental condition, heading was varied as a covariate. Headings were sampled from the range of possible headings (±180°) in 16 evenly spaced steps (22.5°), starting at -168.75° (11.25°, 33.75°, 56.25°, etc.). In this way, the cardinal axes were omitted from the experiment. This was necessary because bias is not defined for the cardinal axes. Positive angles (11.5°) represented clockwise or right directions and negative angles (-11.5°) represented counter-clockwise or left directions.

Figure 4. Dot cloud and dot ground shown during trials.
Procedure

Prior to participation, potential participants were informed about the experimental goals and procedures. If the participants provided their written informed consent, they were seated in the simulator and asked to place their head on a chin rest. Once seated, they were familiarized with the setup. Practice trials were conducted to ensure their understanding of the task and to familiarize them with the procedure. Order of conditions and headings in the practice trials was identical for all participants. During the actual trials, conditions and headings were presented in randomized order. Participants were tasked to provide their heading estimates and vection ratings at the end of the trial as accurately as possible. After confirming their responses by button press, they could initiate the next trial via another button press. Every combination of a particular experimental condition and heading angle was repeated 3 times. This resulted in 16 (combinations of conditions) x 16 (headings) x 3 (repetitions) = 768 trials. Trials were measured in two different sessions on different days to ensure alertness and minimize participant burden. The experiment sessions took 90 minutes each. At the end of the experiment, participants were thanked and debriefed.

Data Analysis

Bias was calculated as angular error from heading judgments and actual headings, by calculating how many degrees the heading judgements deviated from the actual headings (heading response – stimulus heading). Vection was measured on a 7-point scale (‘1’ = no feeling of self-motion, ‘7’ = very strong feeling of self-motion). The effects of the investigated factors on the heading bias were quantified using path analysis, a special case of structure equation modelling (SEM). SEM refers to a diverse family of mathematical models and statistical methods that are used to estimate and calculate relationships of measured and latent variables. Relationships hypothesized in a model can be compared to variables in an actual sample. Path analysis can be understood as a special case of SEM with a single
measured indicator for all analyzed variables. The benefit of SEM is the ability to estimate whole models with multiple dependencies in a single analysis. Path analysis was conducted in R (R Foundation for Statistical Computing, Vienna, Austria, version 3.4.1) using the lavaan package (Rosseel, 2012, version 0.5-23.1097). R script used for path analysis can be found in the appendix.

Results

Datasets of three participants were excluded from the analysis due to technical errors with the input and saving of responses during the experiment. The remaining 17 datasets went into the analysis. Heading responses varying more than 90° from stimuli headings were treated as outliers and omitted from analysis.

Figure 5. Overall heading responses. Ordinate represents the different heading angles shown during the experiment, abscissa represents the heading error. Every dot is a single response. Blue line shows the smoothed mean and shaded area shows 90% CI. A small amount of random noise was added to the location of each dot to prevent overplotting.
Overall heading responses are shown in Figure 5. Heading errors of all participants were mapped on the y-axis across stimuli headings shown during the experiment on the x-axis, this way showing the variance of heading errors for every stimulus heading. Overall, the data indicates a bias away from the fore-aft axis with participants generally overestimating heading angles. The whole bias curve is slightly offset, thus not crossing the cardinal axes. Furthermore, the bias for backward headings seems slightly stronger than the bias for forward headings. Mean error ranged from -0.6° (for 11.25°) to -8.46 (for 146.25°).

**Figure 6.** Heading responses on individual level. The panels show the responses for each participant separately. Ordinate represents the different heading angles shown during the experiment, abscissa represents the heading error. Every dot is a single response. Blue line shows the smoothed mean and shaded area shows 90% CI. A small amount of random noise was added to the location of each dot to prevent overplotting.
Heading responses on the individual level are shown in Figure 6. Heading errors for single participants were again mapped on the y-axis across stimuli headings shown during the experiment on the x-axis. Bias away from the fore-aft axis is also visible on the participant level - quite saliently for some participants (1,10,14,16,20) and with more noise for others (2,3,6,10,14,15,17,19). The remaining participants show different patterns with some considerable noise (4,5,7,8,12,18). The confidence intervals visible in Figure 6 display some substantial variance in responses of these participants.

Figure 7. Overall heading responses split across conditions. Ordinate represents different heading angles shown during the experiment, abscissa represents heading error. Every dot is a single response. Blue and red line show the smoothed mean and shaded areas show 90% CI for each level of conditions: FOV (full = blue, restricted = red), disparity (stereoscopic = blue, monoscopic = red), motion profile (constant velocity = red, raised cosine velocity = blue), and scene (dot ground = red, dot cloud = blue). A small amount of random noise was added to the location of each dot to prevent overplotting.
Figure 7 again shows overall heading errors, now split by the different conditions of visual properties: FOV, disparity, profile, and scene. Bias away from the fore-aft axis can be found across all conditions, although with some differences in magnitude. The two motion profiles (constant velocity vs. raised cosine velocity) do not differ visibly with regards to the angular error. For FOV and scene, a small effect is visible in the amplitude of the curve. When FOV is restricted or when a dot cloud is shown as a scene, the bias away from the fore-aft axis is slightly stronger in magnitude. The strongest effect seems to appear for disparity: When stimuli were shown monoscopically (i.e., without disparity) participants overestimated heading direction considerably in comparison to when stimuli were shown stereoscopically.

Path analysis was used to quantify possible effects of these factors on heading estimation bias. The model specified in the beginning (Figure 1) was adjusted for the analysis: The four visual properties, vection ratings, and angular error went into the analysis. Angular error was used as an indicator for bias. Direction of stimuli headings was added as a covariate to account for the direction sensitivity of the bias. Furthermore, the model was extended with product terms for interaction effects between the direction of stimuli headings and each of the four visual properties. Interaction terms were added to account for the sine-wave form of the bias curve: As the curve potentially inverts, effects of a factor could change depending on the direction of heading.

As the model was recursive, without feedback loops, and without correlated disturbances, identification can be assumed (Bollen, 1989). Identification is an important prerequisite for estimating a model as it implies that it is theoretically possible to derive a unique set of model parameter estimates. Model fit was calculated with maximum likelihood method and robust (Huber-White) standard errors. Global fit testing was done via a chi-square test with \( \chi^2(5, N = 12551) = 8.66, p = .124 \), which indicated an exact model fit because the test is an accept-support test with the null-hypothesis that the estimated model is correct.
That means model-implied covariances are close to the actual calculated covariances from the sample, indicating that relationships specified in the model can be found in the sample. This allows to assume that visual characteristics somehow affected bias and vection. Root Mean Square Error (RMSEA) had a confidence interval from 0 to 0.02 and a p of close fit < 0.001, thus indicating a close-fitting model. The Comparative Fit Index (CFI) had a value of 0.99 and the Standardized Root Mean Square Residual (SRMR) a value of 0.01. All measures of fit equally indicate a very good model fit, showing that on average model corresponds to the data. However, the model still explains very little variance found in the data, only 2% of the variance in bias (R² = 0.02) and 6% of the variance in vection (R² = 0.06). Subsequent local fit testing was conducted with multiple regression analyses to examine individual parts of the model (Table 1).

Table 1

Summary of local fit testing done via multiple regression analyses.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE</th>
<th>β</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>heading</td>
<td>-0.02</td>
<td>&lt; 0.01</td>
<td>-.16</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td>FOV</td>
<td>-0.25</td>
<td>0.27</td>
<td>-.01</td>
<td>.369</td>
</tr>
<tr>
<td>disparity</td>
<td>-0.05</td>
<td>0.27</td>
<td>&lt; -.01</td>
<td>.865</td>
</tr>
<tr>
<td>profile</td>
<td>0.01</td>
<td>0.27</td>
<td>&lt; .01</td>
<td>.966</td>
</tr>
<tr>
<td>scene</td>
<td>0.18</td>
<td>0.28</td>
<td>.01</td>
<td>.521</td>
</tr>
<tr>
<td>vection</td>
<td>0.04</td>
<td>0.11</td>
<td>&lt; .01</td>
<td>.749</td>
</tr>
<tr>
<td>heading * FOV</td>
<td>0.01</td>
<td>&lt; 0.01</td>
<td>.05</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td>heading * disparity</td>
<td>0.01</td>
<td>&lt; 0.01</td>
<td>.07</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td>heading * profile</td>
<td>&lt; -0.01</td>
<td>&lt; 0.01</td>
<td>-.01</td>
<td>.673</td>
</tr>
<tr>
<td>heading * scene</td>
<td>-0.01</td>
<td>&lt; 0.01</td>
<td>-.06</td>
<td>&lt; .001***</td>
</tr>
</tbody>
</table>
The regression analyses show that the effects of the visual properties alone on bias were not significant. However, the effects of stimulus heading and the interactions between stimulus heading and FOV, disparity, and scene were significant. Bias was largely dependent on stimulus heading. Follow up analyses on simple effects of interactions were not available and were thus interpreted via figures. As can be seen in Figure 7, panels a, b, d, the blue and red line for different levels of FOV, disparity and scene differed most for heading angles exactly between cardinal axes (i.e., 45°, 135°). Headings closer to straight ahead or sideways were less influenced by any visual properties. For FOV, smaller FOV seemed to cause a larger bias, indicated by the larger amplitude of the curve. Similarly, monoscopic display for disparity and dot cloud for scene led to a larger bias. Because only certain heading angles seem to be affected, regression estimates of the effects are all quite small. Additionally, vection had no significant effect on bias, but was affected by all visual properties. As expected larger FOV, variable motion profile and ground plane lead to higher vection ratings. Different than expected monoscopic display led to higher vection ratings.

**Discussion**

With regards to the findings, the model proposed in the introduction could not be supported. Although the model achieved a good fit, it just explained a minor part of the observed variance in heading estimations. Bias away from the fore-aft axis as found in this
study could thus not be explained by visual properties alone. The visual properties only marginally affected the magnitude of bias, not its direction. FOV, disparity, and scene slightly affected heading estimates. However, as bias is direction sensitive, this effect was dependent on stimulus directions. FOV, disparity and scene increased bias in certain conditions. When participants had to judge heading angles exactly between cardinal axes, visual properties had their largest effect. Participants showed a slightly larger bias when their FOV was restricted. With a smaller FOV, more FOEs fell out of the view of the participants, making it necessary for them to estimate the direction of heading by triangulation and thus potentially leading to more errors (Koenderink & Van Doorn, 1987). Although not related to FOV, larger errors during backwards motion most likely also stem from triangulation errors. As FOE is not visible during backwards motion, participants have to solely rely on flow triangulation, thus potentially making more errors.

Similarly, as FOV, disparity and scene had an influence depending on stimuli heading. When shown stereoscopic cues or a ground plane, participants showed smaller biases in their heading responses compared to monoscopic displays and dot clouds, respectively. The superiority of stereoscopic display and ground plane scenes could stem from the availability of depth cues. Disparity cues in the optic flow as well as knowledge about the layout of the scene generated more reference points and thus offered information about the depth order, potentially helping participants to resolve uncertainty in the motion parallax and supporting accurate heading estimation (Koenderink & Van Doorn, 1987). In contrast to FOV, disparity, and scene, the two different motion profiles did not affect participants differently: In both motion profiles (constant velocity vs. bell curve velocity), participants committed similar estimation errors in amount and type. Since both motion profiles had the same duration and covered the same distance, they apparently offered the same amount of information to the participants.
Despite these visible effects, there was still considerable noise in the data. Some stimuli headings or combinations of factors apparently led to substantial variance in the responses of some participants. For example, one participant (5) estimated heading direction on some trials in the direct opposite direction as stimulus heading, depending on the scene during some of the trials. This participant most likely changed strategies in heading estimation and judged dot motion instead of self-motion. As self-motion was the opposite direction as dot motion, heading responses are therefore also reversed. This led to larger heading errors and potentially skewed bias. Although no such clear pattern was visible for other participants, it is still possible that participants changed strategies between trials, leading to some variance in their responses.

As mentioned before, the proposed model has to be rejected, also because a possible mediating effect of vection on bias could not be found. However, vection itself was affected by all visual properties. Although the large sample size has to be taken into consideration here, all visual properties affected the vection ratings of the participants. A larger FOV, a monoscopic display, a variable motion profile, and a ground plane all led to higher vection ratings. A larger FOV most likely led to higher ratings because participants felt more present in the scene and motion felt more convincing (Trutoiu et al., 2008). The variable motion profile and the ground plane scene could have led the participants to perceive the motion as more natural and closer to the real world and could thus have caused higher ratings of vection (Palmisano et al., 2008). However, in contrast to earlier research (Palmisano, 1996, 2002; Ziegler & Roy, 1998) we found that monoscopic displays enhanced vection ratings. In prior research, stereoscopic displays were found to enhance vection, because they are perceived more convincing and closer to real-world experience. An explanation could be that the duration of trials was too short for the participants to adapt to stereoscopic stimuli. As a result, the participants would not have experienced the full effect of disparity. However,
BIASES IN THE VISUAL PERCEPTION OF HEADING

during piloting, the stereoscopic stimuli were reported as compelling and vection-inducing. It is also possible that the vection ratings for some participants do not directly reflect the experience of vection and some different concept was rated instead. During practice trials, some participants reported that they had erroneously rated speed instead of vection. Although these participants were corrected and the majority of participants seemingly understood the concept of vection. Additionally, as earlier research has shown, longer durations are needed for linear motion to give the impression of vection (Dichgans & Brandt, 1978). The fact that vection generally has onset times that exceed the duration of the visual motion stimuli presented in the current study may explain why vection was not found to affect bias.

Rejecting the model despite a good model fit in the analyses seems counterintuitive, but good fit has to be taken with a grain of salt because of several reasons. First, as bias is direction sensitive, adding stimuli headings to the analysis most likely largely improved model fit. Thus, visual properties potentially have only a small role in the model fit calculated. Nevertheless, stimuli headings have to be included to account for direction sensitivity of bias. Second, the model in the analysis is just one plausible model with further equivalent models. In the field of SEM, equivalent models are models with the same variables but different causal structures that achieve the same degree of freedom (Stelzl, 1986). While equivalent models achieve the same degree of overall fit, the paths defined still can be quite different. Thus, it is possible that the hypothesized direction of effects could be different than the specified one, still leading to a good model fit. Finally, since a repeated measured design with different individuals was employed, the nested character of the data obtained needs to be considered. Although the experimental conditions were kept stable and reported bias is visible for most participants, effects of nesting on the results could still be possible. As SEM is based on the assumption that data comes from independent observations and with several observations from single participants, analysis has to be seen with care. This
shortcoming is often criticized in SEM literature because nested structure of data is rarely considered (Julian, 2001). A multilevel analysis of data would be a viable approach to counter this issue but during the time of the study, it was not available in lavaan and because it can be considered an advanced technique in SEM (Marcoulides & Schumacker, 1996) it went beyond the scope of this study.

Similarly, the experimental setup could have impacted the overall bias. Using a dial to indicate heading is somewhat artificial and could thus have affected the reference points of the participants and their internal representation of the heading. Additionally, the task of pointing a dial includes a motoric and haptic component, possibly affecting bias. However, Crane (2012) compared responses to spoken heading stimuli with responses to visual and vestibular stimuli and could show that bias mostly stems from sensory stimuli rather than haptic or motor influences. The fact that visual properties affected bias to a certain degree in the present study likewise shows that bias more likely stems from sensory stimuli.

Additionally, studies using on screen arrows or probes (Hummel et al., 2016; Li et al., 2002) or head pointing (Telford et al., 1995) as response mode also found similar biases. Thus, we assume that the experimental setup did not affect bias systematically.

Despite these shortcomings, the present study is comparable to studies with similar setups investigating heading estimation and the bias away from the fore-aft axis found here is in line with earlier studies (Crane, 2012; Cuturi & MacNeilage, 2013; Telford & Howard, 1996). Nevertheless, the network of potential relations between the visual properties andvection outlined in the introduction could not be found in the results. Thus, the true source of bias seems to lie elsewhere. As mentioned before, bias away from the fore-aft axis as found in this study could have a neurophysiological origin: The overrepresentation of neural populations with preference for lateral directions in the MSTd could lead to a more sensitive discrimination of headings near straight ahead and a bias away from the fore-aft axis (Gu et
al., 2010). It is further possible that the distribution of preferred directions of neural populations in the MSTd varies across individuals, leading to individually different biases (De Winkel et al., 2017).

In summary, although FOV, scene, and disparity have a marginal effect on the magnitude of the biases in heading estimation, there is neither evidence that the direction of these biases is affected by either of these characteristics, nor that vection bears any relation to the nature of the bias. Based on other literature, we conclude that the direction of biases in visual heading estimation is likely to be an idiosyncratic property.
References


doi:10.1145/1166087.1166091


of direction signals of image motion in the superior temporal sulcus of the macaque

doi:10.1207/s15327906mbr2103_3

Telford, L., & Howard, I. P. (1996). Role of optical flow field asymmetry in the perception of
doi:10.3758/BF03211881


and curvilinearvection in a large-screen virtual environment with floor projection.

doi:10.1016/0042-6989(94)90324-7

Van Den Berg, A. V., & Brenner, E. (1994b). Why two eyes are better than one for
judgements of heading. *Nature, 371*, 700–702. doi:10.1038/371700a0


Appendix

R Script

Import dataset

```r
# Import dataset from csv file
data <- read.csv(file = "D:/Dokumente/Studium/Universiteit Twente/Thesis/Analysis/Data/dataset.csv")
```

Prepare dataset

```r
# Add session counter
data$sess <- ifelse(data$trial <= 384, 1, 2)

# Add trial counter
data <- data %>%
  group_by(par, FOV, disparity, profile, scene, theta) %>%
  mutate(rep = row_number())
```

Filter excluded datasets

```r
# Filter out participants
data <- data %>% filter(par != 11)  # Error saving responses of second session

data <- data %>% filter(par != 13)  # Error during pointer initialization

data <- data %>% filter(par != 9)   # Error saving vection ratings of first session

# Clean up vection ratings
```
data <- data %>% filter(vection > 0)

Outlier handling

# Filter heading judgements
data <- data %>% filter(error <= 90 & error >= -90)

Create dummy variables for interaction

# Interaction between factors and theta
data$"thetaFOV" = data$FOV*data$theta
data$"thetadisparity" = data$disparity*data$theta
data$"thetaprofile" = data$profile*data$theta
data$"thetascene" = data$scene*data$theta

Define path models

# Full model with interactions
path.full <-'
error ~ theta + FOV + disparity + profile + scene + vection +
thetaFOV + thetadisparity + thetaprofile + thetascene
vection ~ FOV + disparity + profile + scene'

Calculate model fit

library(lavaan)
## This is lavaan 0.5-23.1097
## lavaan is BETA software! Please report any bugs.
model.full <- sem(path.full, data, warn = TRUE, estimator = "mlr", meanstructure = TRUE)
summary(model.full, standardized = TRUE, fit.measures = TRUE, rsquare = TRUE)
## lavaan (0.5-23.1097) converged normally after 79 iterations
##
## Number of observations 12551
##
## Estimator ML Robust
## Minimum Function Test Statistic 8.447 8.656
## Degrees of freedom
5       5
## P-value (Chi-square)
0.133   0.124
## Scaling correction factor
for the Yuan-Bentler correction
0.976
## Model test baseline model:
## Minimum Function Test Statistic 1035.871 1031.718
## Degrees of freedom 19       19
## P-value 0.000  0.000
## User model versus baseline model:
## Comparative Fit Index (CFI) 0.997  0.996
## Tucker-Lewis Index (TLI) 0.987  0.986
## Robust Comparative Fit Index (CFI) 0.996
## Robust Tucker-Lewis Index (TLI) 0.987
## Loglikelihood and Information Criteria:
## Loglikelihood user model (H0) -454911.236 -454911.236
## Scaling correction factor 1.050
## for the MLR correction
## Loglikelihood unrestricted model (H1) -454907.013 -454907.013
## Scaling correction factor 1.034
## for the MLR correction
## Number of free parameters 18       18
## Akaike (AIC) 909858.472  909858.472
## Bayesian (BIC) 909992.348  909992.348
## Sample-size adjusted Bayesian (BIC) 909935.146  909935.146
## Root Mean Square Error of Approximation:
## RMSEA 0.007  0.008
## 90 Percent Confidence Interval

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000</td>
<td>0.016</td>
<td>0.000</td>
</tr>
</tbody>
</table>

## P-value RMSEA <= 0.05

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

## Robust RMSEA

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.008</td>
</tr>
</tbody>
</table>

## 90 Percent Confidence Interval

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000</td>
<td>0.01</td>
</tr>
</tbody>
</table>

## Standardized Root Mean Square Residual:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.005</td>
<td>0.005</td>
</tr>
</tbody>
</table>

## Parameter Estimates:

<table>
<thead>
<tr>
<th>Information</th>
<th>Observed</th>
<th>Standard Errors</th>
<th>Robust.huber.white</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regressions:</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Estimate | Std.Err | z-value | P>|z|< | Std.lv | Std.a |
|----------|---------|---------|-------|-------|-------|
| theta    | -0.023  | 0.003   | -7.843| 0.000 | -0.023| -0.1  |
| FOV      | -0.246  | 0.274   | -0.898| 0.369 | -0.246| -0.0  |
| disparity| -0.046  | 0.274   | -0.170| 0.865 | -0.046| -0.0  |
| profile  | 0.012   | 0.274   | 0.043 | 0.966 | 0.012 | 0.0   |
| scene    | 0.181   | 0.282   | 0.642 | 0.521 | 0.181 | 0.0   |
| vection  | 0.036   | 0.111   | 0.320 | 0.749 | 0.036 | 0.0   |
| thetaFOV | 0.010   | 0.003   | 3.801 | 0.000 | 0.010 | 0.0   |
| thetadisparity | 0.014 | 0.003 | 5.414 | 0.000 | 0.014 | 0.0 |
| thetaprofile | -0.001 | 0.003 | -0.422 | 0.673 | -0.001 | -0.0 |
| thetascene | -0.013 | 0.003 | -5.058 | 0.000 | -0.013 | -0.0 |

## vection ~
## FOV                0.127    0.023    5.576    0.000     0.127    0.0
## disparity         -0.074    0.023   -3.218    0.001    -0.074   -0.0
## profile            0.155    0.023    6.793    0.000     0.155    0.0
## scene             -0.608    0.023  -26.612    0.000    -0.608   -0.2

## Intercepts:

|                  | Estimate  | Std.Err  | z-value | P(>|z|) | Std.lv     | Std.a     |
|------------------|-----------|----------|---------|---------|------------|-----------|
| .error           | -2.238    | 0.511    | -4.381  | 0.000   | -2.238    | -0.1      |
| .vection         | 3.627     | 0.025    | 143.581 | 0.000   | 3.627     | 2.7       |

## Variances:

|                  | Estimate  | Std.Err  | z-value | P(>|z|) | Std.lv     | Std.a     |
|------------------|-----------|----------|---------|---------|------------|-----------|
| .error           | 234.713   | 4.057    | 57.847  | 0.000   | 234.713    | 0.9       |
| .vection         | 1.639     | 0.020    | 83.340  | 0.000   | 1.639      | 0.9       |

## R-Square:

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>error</td>
<td>0.020</td>
</tr>
<tr>
<td>vction</td>
<td>0.060</td>
</tr>
</tbody>
</table>