Exp Brain Res
DOI 10.1007/s00221-012-3048-1

RESEARCH ARTICLE

Multisensory integration in the estimation of walked distances

Jennifer L. Campos · John S. Butler ·
Heinrich H. Bülthoff

Received: 27 September 2011 / Accepted: 21 February 2012

Abstract When walking through space, both dynamic visual information (optic flow) and body-based information (proprioceptive and vestibular) jointly specify the magnitude of distance travelled. While recent evidence has demonstrated the extent to which each of these cues can be used independently, less is known about how they are integrated when simultaneously present. Many studies have shown that sensory information is integrated using a weighted linear sum, yet little is known about whether this holds true for the integration of visual and body-based cues for travelled distance perception. In this study using Virtual Reality technologies, participants first travelled a predefined distance and subsequently matched this distance by adjusting an egocentric, in-depth target. The visual stimulus consisted of a long hallway and was presented in stereo via a head-mounted display. Body-based cues were provided either by walking in a fully tracked free-walking space (Exp. 1) or by being passively moved in a wheelchair (Exp. 2). Travelled distances were provided either through optic flow alone, body-based cues alone or through both cues combined. In the combined condition, visually specified distances were either congruent (1.0×) or incongruent (0.7× or 1.4×) with distances specified by body-based cues. Responses reflect a consistent combined effect of both visual and body-based information, with an overall higher influence of body-based cues when walking and a higher influence of visual cues during passive movement. When comparing the results of Experiments 1 and 2, it is clear that both proprioceptive and vestibular cues contribute to travelled distance estimates during walking. These observed results were effectively described using a basic linear weighting model.

Keywords Optic flow · Proprioception · Vestibular · Multisensory integration · Distance estimation · Self-motion

Introduction

Our most common daily activities involve moving within and throughout our environment. Whether navigating to acquire resources, avoiding dangerous situations or tracking one's position in space relative to important landmarks, accurate self-motion perception is critically important. Self-motion perception is typically experienced when an observer is physically moving through space including, self-propelled movements such as walking, and also when being passively moved (e.g. in a vehicle). In order to

J. L. Campos · J. S. Butler · H. H. Bülthoff (✉)
Department of Human Perception, Cognition and Action,
Max Planck Institute for Biological Cybernetics,
Spemannstr. 38, 72076 Tübingen, Germany
e-mail: Heinrich.buelthoff@tuebingen.mpg.de

J. L. Campos (✉)
Toronto Rehabilitation Institute, 550 University Ave.,
Toronto, ON M5G 2A2, Canada
e-mail: campos.jennifer@torontorehab.on.ca

J. L. Campos
Department of Psychology, University of Toronto,
Toronto, ON, Canada

J. S. Butler
Cognitive Neurophysiology Laboratory,
Department of Pediatrics, Children's Evaluation
and Rehabilitation Center (CERC), Albert Einstein College
of Medicine, Bronx, NY, USA

H. H. Bülthoff
Department of Brain and Cognitive Engineering,
Korea University, Seoul 136-713, Korea

effectively perceive self-motion, several different sensory and motor systems provide redundant information. Of particular importance are the dynamic visual cues (optic flow) and body-based cues (proprioceptive, efference copy and vestibular cues) that are available during self-motion. While much is now known about the extent to which visual cues and body-based cues can be used independently to estimate travelled distance, less is known about how they are combined when both are available, and specifically about the extent to which each is relied upon. Therefore, the objective of the current study was to systematically evaluate the relative contributions of visual, proprioceptive and vestibular cues during locomotion using a travelled distance estimation task.

Using optic flow and body-based cues to estimate travelled distance

During locomotion, spatio-temporal relations can be defined by the pattern of dynamic retinal information that is available during self-motion through space (optic flow) (Gibson 1950). Optic flow is known to play a key role in many aspects of self-motion processing. For instance, it has been shown to inform predictive processes such as estimating heading direction (Warren and Hannon 1998; Warren et al. 2001; Fetsch et al. 2009; Butler et al. 2010, 2011) and time to collision (Lee 1976; Sun and Frost 1998), as well as online locomotor control such as maintaining postural stability (Lee and Lishman 1975), modulating gait (Prokop et al. 1997; Mohler et al. 2007a; Guerin and Bardy 2008) and estimating travelling speed (Larish and Flach 1990; Sun et al. 2003; Banton et al. 2005).

Investigators have more recently begun to assess the role of optic flow in distance estimation and path integration (Bremmer and Lappe 1999; Redlick et al. 2001; Riecke et al. 2002; Frenz et al. 2003; Sun et al. 2004a, b; Frenz and Lappe 2005; Lappe et al. 2007; Campos et al. 2010). Such investigations have demonstrated that humans are generally proficient at using optic flow to make relative distance judgments, including discriminating between and reproducing visually simulated self-motion profiles within the same environment. Optic flow alone without a calibrating scale factor, however, is uninformative about absolute distance. Discrete changes in eye height, lateral distances of surrounding surfaces and texture density can all change the subjective flow velocity and thus the perceived magnitude of self-motion (Frenz and Lappe 2005, 2006). Indeed, when required to produce absolute distance judgments based on previously learned self-motion trajectories experienced exclusively with optic flow, participants consistently underestimate their distance travelled by approximately 25–30% (Frenz and Lappe 2005, 2006). Further, it has been shown that the spatial integration of optic flow is "leaky" such that error accumulates with increasing distance (Lappe et al. 2007). Therefore, it is not clear the extent to which this optic flow is used when other sensory information is available.

Traditionally, studies investigating humans' ability to use optic flow as a cue to distance have presented participants with isolated, computer simulated optic flow (Bremmer and Lappe 1999; Redlick et al. 2001; Frenz et al. 2003; Sun et al. 2003, 2004a; Frenz and Lappe 2005). In most of these studies, the context in which optic flow is experienced is often void of many visual depth cues and/or locomotor cues that are naturally coincident with self-motion. While there are indeed circumstances under which understanding the specific role of optic flow alone is relevant (e.g. fixed-base driving simulators), most natural human interactions occur in the context of multiple cues to travelled distance.

Indeed, body-based cues have also been shown to be sufficient for performing a number of different spatial behaviours. Everyday experiences highlight the fact that, even in the absence of reliable visual feedback (e.g. during darkness, foggy conditions or when moving past an obstruction), body-based cues can be used to update one's position in space. Specifically, these cues include the vestibular information that is generated during changes in velocity (Israël and Berthoz 1989; Harris et al. 2000), proprioceptive information that is provided by the muscles and joints (Mittelstaedt and Mittelstaedt 2001) and motor efference signals representing the commands of these movements.

It has been well documented that humans are highly accurate when asked to view a static visual target in the distance (<20 m) and subsequently walk to it blindfolded (Thomson 1983; Elliott 1986; Rieser et al. 1990; Loomis et al. 1992; Mittelstaedt and Mittelstaedt 2001; Sun et al. 2004b; although see Souman et al. 2009 for longer distances). Participants can also continuously point to a previously viewed target when walking past it blindfolded on a straight, forward trajectory (Loomis et al. 1992; Campos et al. 2009) and to some extent when passively moved along simple trajectories (Siegle et al. 2009; Frissen et al. 2011). While these tasks have been used extensively to demonstrate that adults are highly skilled at judging static visual distances, they also highlight the importance of body-based cues associated with large-scale, goal-directed movements. Others have also demonstrated that individuals are able to reproduce distance information when both learning and responding through blindfolded walking (Klatzky et al. 1998; Mittelstaedt and Mittelstaedt 2001; Sun et al. 2004b) and when being passively transported (Israël and Berthoz 1989; Harris et al. 2000).

Cue integration during travelled distance estimation

While establishing that individuals can use optic flow and body-based cues independently to estimate travelled distances is an important first step, this does not directly address issues related to how the brain uses and integrates these cues when both are available. When walking under natural conditions, travelled distance specified by dynamic visual information is tightly coupled with that specified by body-based cues. For instance, when walking one metre forward, the information that is received both visually and motorically specifies the same distance. Rarely does this relationship change. One real-world example of where this relationship does change is when walking on a moving sidewalk (e.g. at the airport). In this case, the information received visually specifies a greater travelled distance than would be expected given the simultaneously specified distance provided via proprioceptive cues.

Somewhat analogous to the moving sidewalk example, similar approaches have been experimentally implemented as a way of creating incongruencies between naturally coincident cues. This has been achieved by presenting participants with simultaneously available optic flow and body-based cues, while systematically changing the relation between the two (i.e. the extent of visual motion associated with a particular extent of movement). While this has been achieved in real-world settings (e.g. Rieser et al. 1995; Campos et al. 2010), advanced Virtual Reality (VR) technologies are now allowing researchers to more easily manipulate different sources of sensory information dynamically and independently (e.g. Chance et al. 1998; Harris et al. 2000; Kearns et al. 2002, 2003; Proffitt et al. 2003; Sun et al. 2004a; Durgin et al. 2005; Mohler et al. 2007a, b; see also Campos and Bülthoff 2011 for a review). For instance, Harris et al. (2000) examined the relative contributions of vestibular cues and optic flow for estimating the extent of linear self-motion through a virtual hallway. When conflicting vestibular cues and optic flow were presented simultaneously, participants' responses more closely approximated the distances specified by vestibular cues than those specified by optic flow. It has also been shown using a triangle completion task (path integration) that, while participants are able to use optic flow alone, the introduction of body-based cues (proprioceptive and vestibular) led to a decreased variability in performance (Kearns et al. 2002). Sun et al. (2003, 2004a) also created incongruencies between optic flow and proprioceptive cues (experienced by pedalling a stationary bicycle) by manipulating the optic flow gain. In this case, when the relation between the two cues was constantly varied, participants' estimates more closely approximated the visually specified distance. However, although optic flow appeared to be weighted higher in general, the mere presence of proprioceptive information (albeit incongruent) improved visually specified distance estimation, demonstrating a unique form of nonlinear cue integration. While these few approaches have been used to address the multisensory nature of self-motion perception, in general, very little work has been done to understand the relative contributions of visual and body-based cues for travelled distance perception during one of the most natural forms of self-motion—walking (Campos et al. 2010).

The objectives of the current experiments were to use travelled distance estimation as a tool for understanding the relative contributions of visual and body-based cues during walking and during passive transport. In order to achieve this, travelled distance information was provided either through simulated optic flow alone (via a head-mounted display), body-based cues alone (i.e. blindfolded movement) or through both cues combined. This allowed for an assessment of whether any effects of cue combination would be observed. Further, in the combined condition, a subtle visual gain was implemented during self-motion, thereby causing the visually specified distances to be incongruent with distances specified by body-based cues (i.e. either 70 or 140% of the physically travelled distance). This revealed the relative contributions of individual cues when they were combined. A simple mathematic model was used to describe relative cue weighting in the combined cue trials.

Experiment 1 investigated cue integration during full-scale walking behaviours when all body-based cues were present (i.e. proprioceptive and vestibular information), while Experiment 2 investigated cue integration during passive transport (i.e. mainly vestibular information). A comparison between Experiments 1 and 2 also allowed for an indirect assessment of the relative contributions of proprioceptive and vestibular cues during walking. In general, it is important to note, that while travelled distance was used here as a tool for studying multisensory integration during self-motion, the aim was not to understand the accuracy of travelled distance estimation per se, but rather to quantify the relative difference observed between estimates across the different sensory conditions.

Experiment 1: full-scale walking

Methods

Participants

Thirteen participants (5 women) between the ages of 21–26 years completed all conditions across 2–3 different experimental sessions (approximately 1–1.5 h per session).

All participants had normal or corrected-to-normal visual acuity and were naïve to the purposes of the experiment. Participants were recruited through the Max Planck Institute Subject Database and were compensated at a rate of 8 Euros per hour. All participants provided informed written consent before beginning the experiment. This research was performed in accordance with the ethical standards specified by the 1964 Declaration of Helsinki.

Stimuli and apparatus

This experiment took place in a large, fully tracked, free-walking space, 12 m × 15 m in size. Participants' positions were tracked using an optical tracking system (16 Vicon MX13 cameras) through the monitoring of helmet-mounted, reflective markers (See Fig. 1a). Each Vicon camera has a resolution of 1,280 × 1,024, and the tracking system has a maximum frame rate of 484 Hz. In addition to updating the visual environment as a function of the participant's own movements, the tracking system also allowed for the capturing of several motion parameters including walking speed.

Relative position information was used to update the visual information presented via a head-mounted display (HMD). The lightweight HMD used here (eMagin, Z800) has a resolution of 800 × 600 with a refresh rate of 60 Hz and a 40° diagonal FOV per eye. Participants' positional information was sent from the optical trackers, via a wireless connection, to a backpack-mounted laptop worn by the participant. This information was then used to update participants' position and facing direction within the virtual environment (VE). This set-up thus allowed participants to move freely throughout the walking space without being constrained or tethered. A thick black fabric was wrapped around the HMD and the lights in the room were extinguished to ensure that none of the external laboratory environment was visible to participants at any time during the experiment.

The VE was rendered using veLib, a customized VR communications and rendering library. The 3D model of the VE was developed using 3DMax and consisted of a seemingly infinite, empty hallway approximately 5 m wide and 3 m high. The walls and the floor of the hallway were mapped with a completely random texture with no repeating patterns or identifiable landmarks (See Fig. 1a). The target used for the response task consisted of an unfamiliar red and white striped sphere approximately 1 m in diameter. This target was always positioned in the centre of the hallway floor 2 m in front of the participant at the start of the response phase of each trial.

Procedure

In this task, participants were first required to travel down the hallway until reaching a particular, predefined reference distance (4, 6, 8 or 10 m), at which point the screen went blank and a visual prompt instructed them to stop and turn around (See Fig. 2). Subsequently, participants turned 180° and the target appeared in the hallway on the ground in front of them. Participants were then asked to adjust the target using a game controller until the self-to-target distance matched the distance they had just walked (i.e. place the target back to start). Once satisfied with the position of the target, responses were made via a button press. Each of the four distance intervals was repeated six times per condition in a pseudo-randomized order.

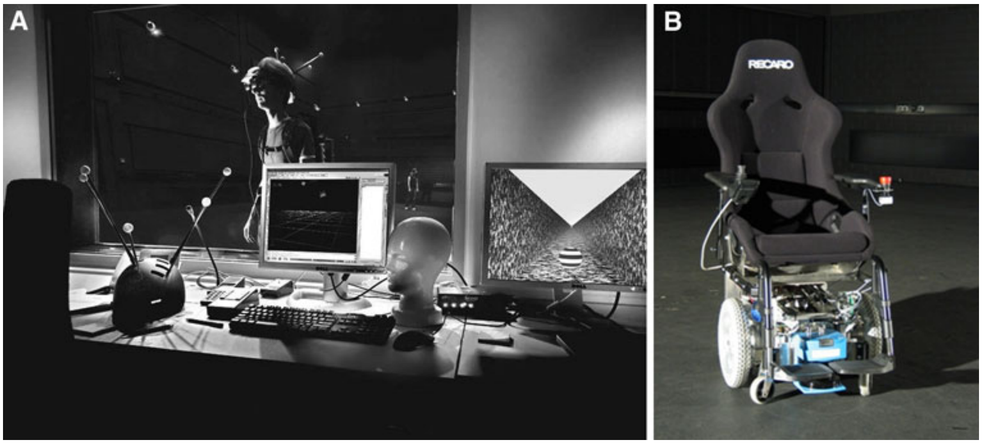

Fig. 1 The stimuli and apparatus used for Experiments 1 and 2. **a** Experiment 1 took place in a large, fully tracked walking space (12 m × 15 m). Participants wore a helmet outfitted with reflective markers that were tracked using an optical tracking system (16 Vicon MX13 cameras). Information about the participant's position and orientation was sent from the optical trackers, via a wireless connection, to a backpack-mounted laptop worn by the participant. *Photo courtesy of Manfred Zentsch.* **b** Experiment 2 took place in the same tracking hall, but participants were passively transported in a wheelchair (robotic wheelchair shown here). The stimuli in both studies consisted of a seemingly infinite hallway with random texture on the walls and floor and the target consisted of a striped sphere

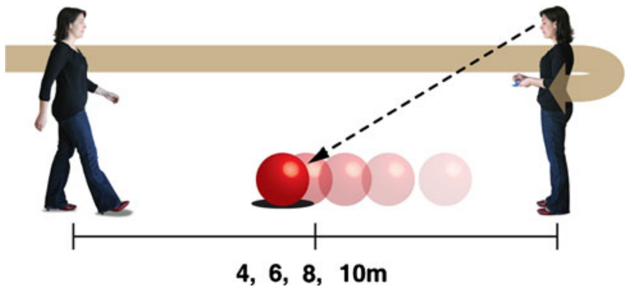

Fig. 2 The procedure for both experiments was the same (Exp. 1/walking shown here). Participants moved a particular distance (4, 6, 8, 10 m) until visually prompted to stop and turn around. They then turned 180° and used a control device to adjust a target until the self-to-target distance matched the distance they had just walked

Baseline and practice

Each participant spent 5 min on walking while wearing the HMD to become accustomed to moving within the VE. This also allowed for any necessary adaptation to occur to ensure a baseline perceptual-motor coupling. Within this time, participants also completed five practice trials without feedback to ensure that they were comfortable with the task. Both during practice and during each of the experimental trials, the experimenter walked beside the participant to increase perceived safety and to prevent extreme veering if necessary (this was rarely needed).

Experimental conditions

Four conditions were conducted as a way of evaluating the relative contributions of optic flow and body-based cues when estimating walked distances. In each of the experimental conditions, the response mode always remained the same (self-to-target distance reproduction). The sensory/motor information provided during the walked reference distance however varied as a function of condition. The order of conditions and the distance trials within conditions were pseudo-randomized across and within participants, respectively.

Visual and body-based cues congruent In the congruent condition (*CON*), during the walked distance, the optic flow and body-based information were combined and congruent throughout the entire session. This meant that there was a 1:1 relationship between the motorically specified distance and the visually specified distance at all times during walking, just as there would be during natural walking.

Visual and body-based cues incongruent In the incongruent condition (*INCON*), during the walked distance, the relation between optic flow and body-based cues was varied by manipulating the visual gain. Three gain values of 0.7×, 1.0× (CON) and 1.4× were used. By manipulating the visual gain, we were able to discretely provide participants with a visually specified travelled distance that was either longer or shorter than the motorically specified travelled distance (for a total of 12 visual distances). For instance, for a visual gain of 1.4×, if participants travelled a real-world, physical distance of 10 m, they would experience a motorically specified distance of 10 m, but a visually specified distance of 14 m. These gain values were determined during pilot testing to ensure that most participants were not consciously aware of the conflict and are within the range of gains that went undetected by participants in past studies (e.g. Steinicke et al. 2010). Across trials, the order of the gains was randomized in such a way that adaptation effects would be highly unlikely. Each of the four distances was repeated six times for each gain level. The 1.0× gain trials were also included here to ensure that the estimates in congruent trials did not change as a function of being embedded among trials with changing gain values (e.g. adaptation or carry over effects). Results demonstrated that there were no differences between distance estimates for the 1.0× gain trials in the *INCON* condition and estimate in the *CON* condition.

Vision alone In the vision alone condition (*VIS*), participants stood stationary while viewing a visual trajectory of the movement down the hallway. The parameters of the linear movement trajectory (i.e. average velocity, acceleration, etc.) were calculated using the participant's own walking parameters during their last few practice trials. Information from the y and z dimensions (i.e. head bob, lateral sway, etc.) was not reproduced, but was continuously tracked and updated online as the participant viewed the translational movement. In other words, if participants looked around during the visually simulated travelled distance, these head movements would be updated appropriately in the display. Three visual speeds were included in the *VIS* condition and were used to represent travelled distances that were equivalent to the visual distances resulting from the three gains introduced in the *INCON* condition (i.e. 0.7×, 1.0× and 1.4× of 4, 6, 8 and 10 m; for a total of 12 distances).

Body-based cues alone In the body-based cue alone (*BODY*) condition, participants walked the distances in the complete absence of vision. In this case, the HMD presented a blank screen during walking other than the prompts to commence and terminate the walked distance prior to responding. The hallway and target appeared again once the participant turned around to make their self-to-target estimate.

Results

Multisensory versus unisensory conditions

As mentioned earlier, of main interest to the current study was the relative difference in estimates across the four different conditions rather than the accuracy of absolute distance estimates per se. In fact, because absolute visual distance estimates in Virtual Reality are often mis-estimated (e.g. Witmer and Kline 1998; Knapp and Loomis 2004; Thompson et al. 2004; Mohler et al. 2007b; Waller and Richardson 2008), the final position of the target does not necessarily provide an accurate indicator of perceived distance travelled. However, assuming that the perceived scale of the response stimuli did not change throughout the four conditions, these estimates provide a good indication of *relative* differences across conditions. Therefore, for each participant we calculated the proportion difference between their average estimates for each distance in the *CON* condition and their average estimates for each distance in each of the *VIS* and *BODY* conditions (i.e. *VIS/CON* and *BODY/CON*). These values were then averaged across participants.

Results demonstrated that travelled distance estimates for the *BODY* condition (M for each distance for 4, 6, 8, 10 m = 8.25, 11.95, 15.63, 18.08) were on average 18% longer than the *CON* condition (M = 7.31, 10.62, 13.32, 16.86), while distance estimates for the *VIS* condition (M = 3.98, 5.64, 6.10, 7.33) were 35% shorter than the *CON* condition (See Fig. 3). A three (Condition: *VIS/CON, CON/CON, BODY/CON*) × four (Distance: 4, 6, 8, 10 m) repeated measures ANOVA was conducted on proportion difference scores and demonstrated a main effect of Condition (F (2, 11) $= 21.68$, $p < 0.001$), no main effect of Distance and no interaction effect. Planned comparisons demonstrated significant differences between the *CON* and *VIS* conditions (F (1, 12) $= 22.77$, $p < 0.001$) and the *CON* and *BODY* conditions (F (1, 12) $= 6.97$, $p < 0.05$).

Effect of changing visual gain

In the *INCON* condition, distance estimates in the high visual gain trials (1.4×) (M = 8.62, 12.08, 15.69, 20.13) were on average 4% longer than in the congruent trials (1.0×) (M = 8.77, 12.48, 15.20, 19.50), while distance estimates in the low visual gain trials (0.7×) (M = 8.42, 11.59, 15.19, 18.72) were 6% shorter than the congruent trials (See Fig. 4). A three (Gain; 1.4×/1.0×, 1.0×/1.0× and 0.7×/1.0×) × four (Distance) repeated measures ANOVA conducted on proportion difference scores demonstrated a significant main effect of Gain (F (2, 11) $= 5.27$, $p < 0.05$), no main effect of Distance and no interaction effects. Planned comparisons demonstrated

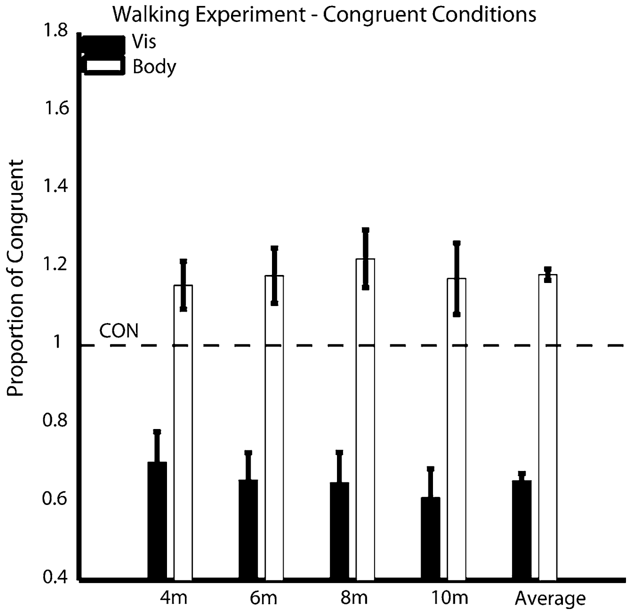

Fig. 3 Experiment 1 results illustrating the average proportion difference scores between the two unisensory conditions (*BODY* and *VIS*) and the combined and congruent condition (*CON represented by the dotted horizontal line*). Averaged data are shown for each distance and also collapsed across distances. *Error bars* represent standard errors

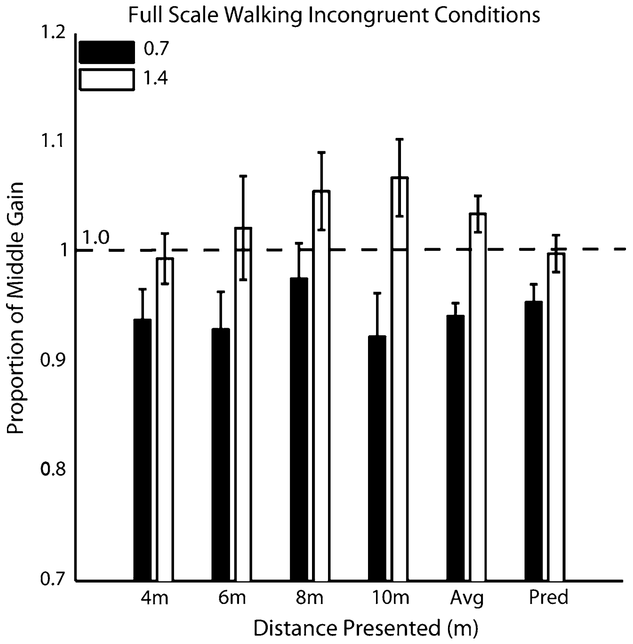

Fig. 4 Experiment 1 results illustrating the average proportion difference scores between the low gain trials (0.7×) versus the congruent trials (1.0×) and between the high gain trials (1.4×) and the congruent trials (1.0×) during walking. Averaged data are shown for each distance and collapsed across distances. Predictions based on MLE are also plotted. *Error bars* represent standard errors

significant differences between the low gain and congruent trials (F (1, 12) = 7.33, $p < 0.05$), but no significant difference between the high gain and congruent trials.

When considering the *VIS* control trials in which the visual speed/distances were equivalent to those for each of the gain levels in the *INCON* condition (but in the absence of physical movement), a significant effect was observed. Specifically, the distance estimates in the trials equivalent to the high gain trials were on average 19% longer than those equivalent to the congruent trials, and the trials equivalent to the low gain trials were 14% shorter than those equivalent to the congruent trials. These values represent the maximum effect of the gain manipulation that could be expected in the *INCON* condition should vision have been used exclusively. A three ("Gain"; 1.4×/1.0×, 1.0×/1.0×, and 0.7×/1.0×) × four (Distance) repeated measures ANOVA on proportion difference scores demonstrated a significant main effect of Gain (F (2, 11) = 39.87, $p < 0.001$), no main effect of Distance and no interaction effects. Planned comparisons demonstrated significant differences between the low gain (0.7×) and middle gain (1.0×) trials (F (1, 12) = 86.92, $p < 0.001$), and between the high gain (1.4×) and middle gain (1.0×) trials (F (1, 12) = 23.84, $p < 0.001$).

Model predictions assuming a linearly weighted summation

If the unimodal cues (*VIS, BODY*) are integrated using a weighted average, then the combined condition (CON) can be expressed as a linear sum

$$CON = w_{VIS}VIS + w_{BODY}BODY, \quad w_{VIS} + w_{BODY} = 1, \tag{1}$$

where w_{VIS}, w_{BODY} are the weights for the unimodal visual and body conditions. As participants only performed four repetitions per distance, the mean data for each distance were used to calculate the groups' weights.

For the unimodal and congruent conditions, the observed visual and body weights were calculated from

$$\begin{aligned} w_{VIS} &= \frac{CON - BODY}{VIS - BODY}. \\ w_{BODY} &= \frac{CON - VIS}{BODY - VIS} = 1 - w_{VIS} \end{aligned} \tag{2}$$

The observed weights indicate the extent to which each modality was relied upon during walking. The closer each modality's weight is to 1.0, the more it was relied upon in the *CON* condition. The observed group average weights are $w_{VIS} = 0.22$ and $w_{BODY} = 0.78$, thus suggesting a higher reliance upon body-based cues.

The unimodal observed weights were then used to make predictions for the incongruent conditions using

$$Pred^{Gain} = w_{VIS}\,VIS^{Gain} + (1 - w_{VIS})BODY. \tag{3}$$

where $Pred^{Gain}$ is the predicted distance estimate of the combination of the visual and body-based cue estimates for each of a given distance and a given gain factor. Thus, 12 predictions were calculated for each participant, one for each of the four distances at each of the three gain levels (See Fig. 4). To test whether these predictions are consistent with the observed data in the incongruent conditions, a two (Observed vs. Predicted) × three (Gain: 0.7×, 1.0×, 1.4×) × four (Distance: 4, 6, 8, 10 m) repeated measures ANOVA was conducted on the proportion difference scores. The results demonstrated no significant difference between Observed and Predicted values (F (1, 12) = 0.03, $p = 0.87$), a significant main effect of gain (F (2, 11) = 13.09, $p < 0.01$), and no significant main effect of distance (F (3, 10) = 1.04, $p = 0.418$.).

Discussion

The results of Experiment 1 demonstrate that combining visual and body-based cues leads to distance estimates that fall somewhere between the two unisensory conditions. This indicates that neither cue is used exclusively when both sources of information are available (i.e. these findings are not consistent with a "winner-takes-all" interpretation), nor does one cue "capture" the other, but rather both sources of information contribute to the final estimate. The combined cue estimates did, however, more closely approximated estimates in the *BODY* condition compared to those in the *VIS* condition. Further, when changing the visual gain during walking, even though some differences in distance estimates were observed in the predicted direction as a function of gain, these differences were only significant when comparing the low gain trials with the middle gain trials. If participants were only using vision during the combined cue conditions, one would expect to see greater differences for both the high and low gain manipulations. Finally, the observed incongruent conditions were effectively predicted assuming that participants used a weighted linear sum.

Experiment 2: passive transport

In Experiment 1, all body-based cues were available during walking (i.e. proprioceptive and vestibular) and a higher weighting of these combined non-visual cues was observed. In Experiment 2, proprioceptive information related to walking was removed by passively transporting participants in a wheelchair. This limited the body-based information to predominantly vestibular cues. Vestibular information is mainly provided by structures in the inner ear, including the semicircular canals, which detect angular accelerations, and the otoliths, which detect linear accelerations (see Angelaki and Cullen 2008; Angelaki et al. 2009, for reviews). Some findings have been interpreted to indicate that head velocity and displacement can be accurately perceived by integrating the linear acceleration information detected by the otolith system (Berthoz et al. 1995). Others indicate that the influence and/or effectiveness of vestibular information in this respect is somewhat limited, particularly when other non-visual information such as vibrations are no longer available (Seidman 2008), when moving along trajectories with more complex velocity profiles (Siegle et al. 2009) or during larger scale navigation (Waller and Greenauer 2007).

Therefore, Experiment 2 investigated whether the weighting of visual and body-based cues changes with the removal of the proprioceptive inputs from the legs that are typically available during walking. This was evaluated by again comparing unisensory conditions (vestibular alone and visual alone) to combined cue conditions and also by introducing a visual gain manipulation during passive movements. Although we are assuming that the main source of non-visual information provided during passive movements is vestibular in this case, it remains possible that other information such as vibrations, pressure from the seat during accelerations, wind, and other inertial cues may also provide information about self-motion in this context (Yong et al. 2007).

Methods

Participants

Twelve participants (4 women) between the ages of 19–36 years completed all conditions across 2–3 different sessions (approximately 1–1.5 h each session). Two participants did not have a complete dataset for the *VIS*, *VES* and *CON* conditions, and therefore their data were not included in any analyses comparing across conditions (however, their data were included when comparing across gains). All participants had normal or corrected-to-normal visual acuity and were naïve to the purposes of the experiment. They were recruited through the Max Planck Institute Subject Database and were compensated at a rate of 8 Euros per hour. All participants provided informed written consent before beginning the experiment. This research was performed in accordance with the ethical standards specified by the 1964 Declaration of Helsinki.

Stimuli and apparatus

The tracking laboratory and visualization set-up (including the Virtual Environment and the HMD) were identical to that used in Experiment 1. For the passive movement condition in this experiment, 2/3 of participants were transported by a fully programmable robotic wheelchair (BlueBotics, Lausanne, Switzerland) modified with an ergonomic seat (Recaro, Kirchheim unter Teck, Germany) (Fig. 1b), while the other 1/3 were pushed in a standard manual wheelchair. The speed of the robotic wheelchair was programmed to be equivalent to participants' average walking speed during their practice trials, and the speed of the manual wheelchair was controlled by the Research Assistant who attempted to maintain a similar speed (as was confirmed through the tracking data). There were no differences between the estimates in these two groups and therefore the data were collapsed.

Procedure

The same four conditions were used in Experiment 2 as were used in Experiment 1, with the exception that participants never walked, but were passively transported in a wheelchair during actual movement conditions, thus limiting the body-based cues mainly to vestibular information. These conditions were, therefore, *CON, INCON, VIS* and *VEST*. When performing the *VIS* condition, participants were also seated in the wheelchair even though it did not move and only the distances equivalent to the congruent condition were tested. The eye height was adjusted to be equal for all conditions in Experiments 1 and 2 so that even when participants were seated, the optic flow generated in the HMD was appropriate for a standing eye height. The same four distances (4, 6, 8, 10 m) were repeated four times per condition and four times for each gain in the *INCON* condition.

Results

Multisensory versus unisensory conditions

Distance estimates for the *VEST* condition (M = 11.64, 14.06, 15.75, 21.95) were on average 48% longer than the *CON* condition (M = 8.22, 9.87, 12.05, 14.46), while distance estimates for the *VIS* condition (M = 5.32, 7.13, 7.81, 9.00) were 35% shorter than the *CON* condition (See Fig. 5). A three (Condition: *VIS/CON, CON/CON, VEST/CON*) × four (Distance: 4, 6, 8, 10 m) repeated measures ANOVA was conducted on proportion difference scores and demonstrated a main effect of Condition (F (2, 8) = 29.88, $p < 0.001$), no main effect of Distance and no interaction effect. Planned comparisons demonstrated

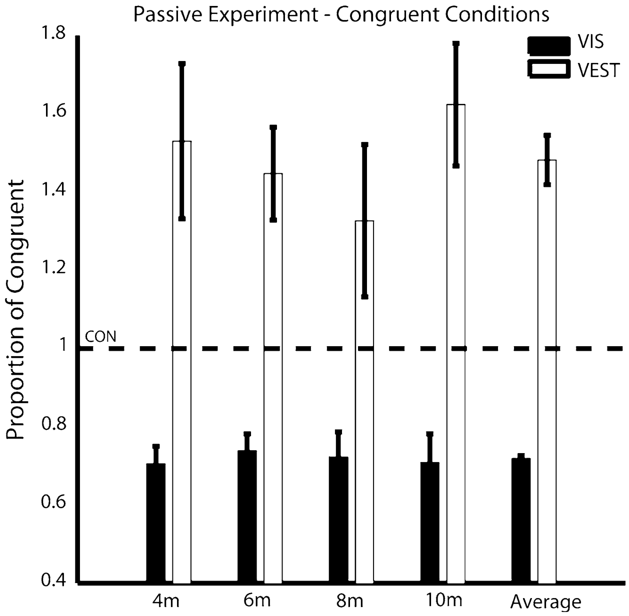

Fig. 5 Experiment 2 results illustrating the average proportion difference scores between the two unisensory conditions (*VEST* and *VIS*) and the combined and congruent condition (*CON represented by the dotted horizontal line*). Averaged data are shown for each distance and also collapsed across distances. *Error bars* represent standard errors

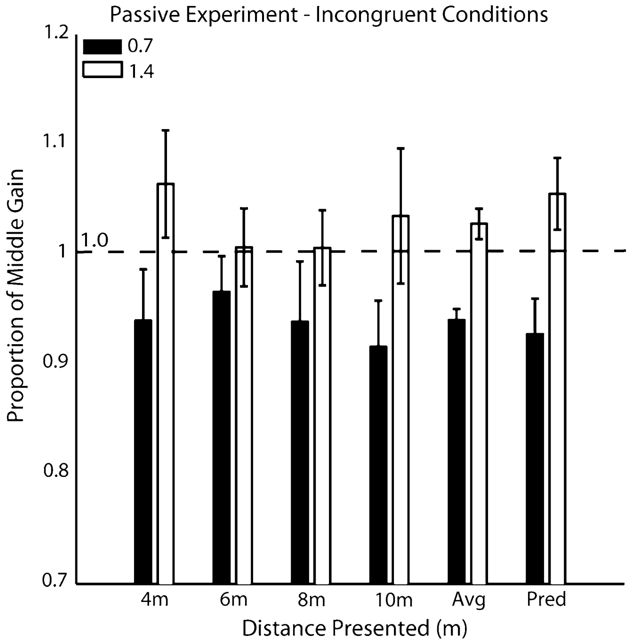

Fig. 6 Experiment 2 results illustrating the average proportion difference scores between the low gain trials (0.7×) versus the congruent trials (1.0×) and between the high gain trials (1.4×) and the congruent trials (1.0×) during passive transport. Averaged data are shown for each distance and collapsed across distances. Predictions based on MLE are also plotted. *Error bars* represent standard errors

significant differences between the *CON* and *VIS* conditions ($F\,(1, 9) = 33.10, p < 0.001$) and the *CON* and *VEST* conditions ($F\,(1, 9) = 18.22$, $p < 0.01$).

Effect of changing visual gain

In the *INCON* condition, the average distance estimates across all 12 participants in the high visual gain trials (1.4×) (M = 9.28, 11.21, 13.91, 15.84) were 3.5% longer than in the congruent trials (1.0×), (M = 8.8, 10.70, 13.21, 15.52), while distance estimates in the low visual gain trials (0.7×) (M = 8.2, 10.75, 12.71, 14.78) were 6% shorter than the congruent trials (See Fig. 6). A three (Gain; 0.7×/1.0×, 1.0×/1.0×, 1.4×/1.0×) × four (Distance) repeated measures ANOVA on proportion difference scores demonstrated a significant main effect of Gain (F (2, 10) = 5.49, $p < 0.05$), no main effect of Distance and no interaction effect. Planned comparisons demonstrated significant differences between the high gain trials and the low gain trials (F (1, 11) = 10.16, $p < 0.01$).

Model predictions assuming a linearly weighted summation

If the unimodal cues (*VIS, VEST*) are integrated using a weighted average, then we can reformulate Eq. 1 to investigate visual-vestibular integration.

$$CON = w_{VIS} VIS + w_{VEST} VEST, \quad w_{VIS} + w_{VEST} = 1, \quad (4)$$

where w_{VIS} and w_{VEST} are the weights for the unimodal visual and vestibular conditions. The group average observed weights (across the ten participants with complete datasets) were calculated from the unimodal and congruent means

$$w_{VIS} = \frac{CON - VEST}{VIS - VEST}, \quad w_{VEST} = \frac{CON - VIS}{VEST - VIS} = 1 - w_{VIS} \quad (5)$$

From the above equation, the observed group average weights were calculated to be $w_{VIS} = 0.54$ and $w_{VEST} = 0.46$, suggesting that visual and vestibular cues were approximately equally relied upon in the combined condition.

As in Experiment 1, the observed weights were used to make predictions for the incongruent conditions.

$$Pred^{Gain} = w_{VIS} VIS^{Gain} + (1 - w_{VIS}) VEST. \quad (6)$$

For each participant, 12 predictions were calculated, one for each of the three gain factors and the four distances (See Fig. 6). It should be noted that because participants did not complete the full set of visual "gains" in the *VIS* condition again in Exp. 2, we simply calculated the VIS^{Gain} by multiplying each participants *VIS* by the gain factor calculated from Experiment 1 such that $VIS^{0.7} = 0.86*VIS$ and $VIS^{1.4} = 1.19*VIS$. To test whether the model predictions were significantly different from the observed data, a two (Observed vs. Predicted) × four (Distance: 4, 6, 8, 10 m) × three (Gain: 0.7×, 1.0×, 1.4×) repeated measures ANOVA was conducted on the proportion difference scores. The results demonstrated no significant difference between Observed and Predicted values (F (1, 9) = 0.421, $p = 0.533$), a significant main effect of gain (F (2, 8) = 14.90, $p = 0.002$) and no significant main effect of distance (F (3, 7) = 0.583, $p = 0.65$).

Vestibular versus proprioceptive cue weighting: comparing across Experiments 1 and 2

Comparing the data in Experiments 1 and 2 provides indirect insight into the approximate weights of vestibular and proprioceptive cues for travelled distance perception. An independent samples t test was conducted to compare the individual participants' weighting of body-based cues in Experiment 1 ($PROP + VEST$) and the weighting of vestibular cues in Experiment 2 for each of the four distances, and demonstrated a significant difference ($t(6) = -7.04$, $p < 0.001$). This exploratory analysis suggests that the proprioceptive information from the legs during walking contributes beyond that of only passive motion cues. If vestibular and proprioceptive cues are linearly weighted, then the body-based cues can be roughly expressed as

$$\begin{aligned} BODY &= w_{PROP}PROP + w_{VEST}VEST \\ w_{PROP} &+ w_{VEST} = 1, \end{aligned} \tag{7}$$

where w_{PROP} and w_{VEST} are the weights for the vestibular and proprioceptive cues.

From Experiments 1 and 2, we have travelled distance estimates from combined body-based cues [$BODY$ ($PROP + VEST$) in Exp. 1] and mainly vestibular cues ($VEST$ in Exp. 2). Given that the average distance estimates for the $BODY$ condition and $VEST$ condition are not statistically different ($p > 0.05$), there are two possible interpretations regarding the contributions of proprioceptive information during walking: (1) proprioception plays no role when combined with vestibular inputs and (2) proprioception yields similar distance estimates to those of vestibular cues and combined body-based cue estimates. The latter assumption is more parsimonious considering that the visual weights are different between Experiments 1 and 2, suggesting that the removal of proprioceptive cues do in fact affect distance estimates. Furthermore, if the proprioceptive distance estimates were much longer or shorter than the vestibular estimates, one would expect the $VEST$ distance estimates (Exp. 1) to be different than the $BODY$ distance estimates (Exp. 2). Based on these assumptions, an estimate of the weight of the proprioceptive information (w_{PROP}) can be roughly calculated using the visual weight (w_{VIS}) from Experiment 1 and the vestibular weight (w_{VEST}) from Experiment 2

$$w_{VIS} + w_{BODY} = w_{VIS} + (w_{PROP} + w_{VEST}) = 1, \tag{8}$$

$$w_{PROP} = 1 - (w_{VIS} + w_{VEST}). \tag{9}$$

From Eq. 8, the group average weights are $w_{PROP} = 0.32$ and $w_{VEST} = 0.68$.

Notably, these results provide only a rough estimate in this context given that the proper controls may not be in place to make a direct comparison. Interestingly, however, these weights are strikingly similar to those calculated in a recent study specifically designed to compare the contributions of vestibular and proprioceptive inputs during blindfolded curvilinear walking using a within-subjects design (Frissen et al. 2011). In that study, the calculated weights were $w_{PROP} = 0.38$ and $w_{VEST} = 0.62$ (See also "General discussion").

Discussion

The results of Experiment 2 demonstrated that combining visual and vestibular cues led to distance estimates that fell somewhere between the estimates in the two unisensory conditions. This indicates that neither cue was used exclusively when both sources of information were available, but rather both sources of information contributed to the final estimate. That said, the combined cue estimates more closely approximated estimates in the vision alone condition compared to those in the vestibular alone condition, thus suggesting that visual information was weighted higher during passive transport. Further, when changing the visual gain during passive movements, a significant effect of gain was observed for both high and low gain trials. This indicates that participants were indeed using vision during the combined cue conditions. Further, the observed data were not statistically different from model predications indicating that it is likely that visual and vestibular cues were integrated in a manner consistent with a linearly weighted summation. Taken together with the results of Experiment 1, these data also suggest that during walking, both proprioceptive and vestibular inputs contribute to non-visual distance perception, with higher weights attributable to vestibular cues.

General discussion

Integration of visual and body-based cues during walking

Overall, the results of this study indicate that during walked distance perception, while both visual and body-based cues contributed to the overall estimate, body-based cues appeared to be particularly important. Further, by comparing the results of Experiments 1 (walking) and 2 (passive movement), it is clear that individual sources of body-based cues (i.e. proprioceptive and vestibular) each contribute to walked distance perception.

The specific contributions of proprioceptive inputs to self-motion perception could reflect the use of walking parameters such as step length, which have been shown to provide a stable metric by which to estimate the extent of self-motion (Durgin et al. 2009). This holds true even when explicit cognitive strategies such as step counting are not permitted. Notably, in the current experiment if participants were using a step counting strategy, there should be no differences between estimates in the *BODY* and *CON* conditions in Experiment 1, yet clear differences were observed. A collection of recent studies have also shown that other terrestrial animals such as ants rely heavily on a type of "step integrator", which integrates information about step length, step rate and load (Wittlinger et al. 2006). This proprioceptive integrator is highly robust and continues to be effective even when walking behaviour is disturbed by introducing highly uneven ground surfaces (Steck et al. 2009) or when estimating straight-line distance travelled when the elevation of the ground surface has changed over the trajectory (i.e. by subtracting out the vertical components; Wohlgemuth et al. 2001).

It is possible that the lower weighting of optic flow during walked distance estimation in humans is due to the fact that not only does optic flow need to be appropriately and continuously scaled based on the perceived layout of our constantly changing environment, but it must also take into account retinal flow based on eye movements and rotational head movements in addition to the translatory movements of the body. Therefore, it is possible that noise in the visual integration system accumulates as a result, perhaps causing the stable metric of proprioceptive inputs combined with vestibular inputs to be considered more reliable. There is also now evidence to suggest that the integration of self-motion information is "leaky". This has been shown in the context of travelled distance estimation for optic flow (Lappe et al. 2007; Lappe and Frenz 2009) and during passive self-motion perception in the absence of vision (e.g. Siegle et al. 2009). It may be the case that proprioceptive information provides an additional mechanism by which visual and vestibular information remains continuously calibrated with actual movement through space.

The results of Experiment 1 are also supported by a series of studies addressing a similar question in a natural, real-world environment. Specifically, Campos et al. (2010) used two complementary techniques to dissociate visual and body-based cues during walked distance perception in a real-world environment. First, lenses were used to magnify or minify the visual environment during walking. Second, two walked distances were presented in succession and were either the same or different in magnitude; vision was either present or absent in each. A computational model was developed based on the results of both experiments and demonstrated that body-based cues were weighted about twice as high as optic flow; the combination of the two cues being additive. This indicates that the lower weighting of visual information in the current experiment was likely not an artefact of using VR (e.g. smaller FOV, less realistic visuals, lack of a scaling features, the conscious awareness of the simulation). The ecological validity of using VR in this context is also supported by findings reported by Lappe and Frenz (2009) who have demonstrated that the consistently observed underestimation of travelled distance based on simulated optic flow is similar to the errors observed when walking in the real world with vision (Lappe and Frenz 2009; see also Sun et al. 2004b; Campos et al. 2010).

Importantly, VR is a valuable tool to effectively dissociate the relative contributions of optic flow and body-based cues when walking in a way that is difficult or impossible to achieve in the real world. Specifically, it is very difficult to completely isolate optic flow in the real world, either from other visual cues or from body-based information. Further, it is very challenging to attempt to dissociate these two very tightly linked sources of sensory information without creating noticeable conflicts. Therefore, the VR paradigm in the current study provides unique advantages to the aforementioned real-world studies due to the ability to isolate visual flow from body-based cues and through the ability to introduce subtle visual gain manipulations.

Integration of visual and vestibular cues during passive movement

By removing important sources of proprioceptive inputs during passive movements, this provided further insights into the specific contributions of vestibular inputs during self-motion. First, the results from Experiment 2 demonstrated that both visual and vestibular inputs contribute to travelled distance perception when both are available. This supports previous work demonstrating that combined visual-vestibular conditions result in performances that differ from conditions under which either visual or vestibular cues are presented alone. For instance, clear combined cue effects have been observed for tasks including egocentric heading estimation (Butler et al. 2010, 2011; Fetsch et al. 2009, 2010), perceived self-orientation following a rotation (Klatzky et al. 1998) and during the reproduction of a travelled trajectory (Bertin and Berthoz 2004).

Second, when specifically assessing relative cue weighting during combined cue conditions, the current results demonstrated a higher reliance on visual cues than on vestibular cues during passive movements (although this difference was relatively small with visual and vestibular weights of 0.54 and 0.46, respectively). These results are consistent with studies demonstrating that visual cues are weighted higher during steering tasks (Wilkie and Wann 2005) and during egocentric spatial updating (Riecke et al. 2006). However, others have demonstrated a dominant role for vestibular cues during passive self-motion perception. For instance, Harris et al. (2000) evaluated the ability of participants to perceive linear trajectories using either visual information provided through a HMD and/or vestibular sources when passively moved on a cart. In their case, they report that when visual and vestibular inputs were concurrently available, vestibular cues captured self-motion perception. There are several possible reasons for the inconsistency between their results and those of the current study. First, the gains that were used in their task were higher than those used in the current study (0.5 and 2.0 vs. 0.7 and 1.4 respectively). Indeed, other studies have demonstrated that the magnitude of a cue conflict can affect the characteristics of sensory integration, with large conflicts possibly causing the cues to not be perceived as a single percept and hence disrupting integration (Wallace et al. 2004; Gepshtein et al. 2005; Körding et al. 2007). It is possible, for instance, that because a larger gain was placed on the visuals, these cues were interpreted as being less reliable than the vestibular inputs and therefore, vestibular cues were weighted higher. Second, in the Harris et al. (2000) study, the visual display was non-stereoscopic, whereas the visual stimuli in the current display were presented stereoscopically. Recent findings have indicated that visual-vestibular integration is facilitated for most participants when stereo cues are present (Butler et al. 2011).

Integration of proprioceptive and vestibular cues during walking

The vast majority of research investigating multisensory integration during self-motion has evaluated the role of visual versus body-based cues, whereas far fewer studies have systematically evaluated the weighting of specific body-based cues to self-motion (i.e. vestibular vs. proprioceptive). Several previous studies have compared estimates during unisensory proprioceptive and unisensory vestibular conditions for simple behavioural tasks, such as judging displacement during forward linear movements through space (Mittelstaedt and Mittelstaedt 2001; Marlinsky 1999) and estimating angular displacement (Bakker et al. 1999; Becker et al. 2002; Jürgens and Becker 2006). For instance, Mittelstaedt and Mittelstaedt (2001) reported that participants could accurately estimate the length of a travelled path when walking in place (proprioception) or when being passively transported (vestibular). Similarly, others have shown that for the estimation of angular displacement when stepping in place and/or being passively rotated on a rotating platform, both proprioceptive and vestibular information can be used independently (Becker et al. 2002; Jürgens and Becker 2006).

More recently, others have introduced conflicts between proprioceptive and vestibular inputs while participants either stepped around their earth-vertical body axis (Bruggeman et al. 2009) or when walking curvilinear paths through space (Frissen et al. 2011). For instance, Frissen et al. (2011) used a large circular treadmill, which featured a motorized handlebar that could move independently of the treadmill disc. Consequently, this allowed for the manipulation of proprioceptive and vestibular inputs independently during walking given that the disc and handlebar could be moved at different rates (i.e. changing the relationship between the rate of walking through space and the rate of walking in place). The results of this study demonstrated that when conflicts were introduced between the vestibular and proprioceptive cues, spatial updating was based on a weighted average of the two inputs, with a higher weighting attributable to vestibular cues. These results are highly consistent with those of the current study such that the proprioceptive and weights reported by Frissen et al. were $w_{PROP} = 0.38$ and $w_{VEST} = 0.62$ respectively compared to the similar weights of $w_{PROP} = 0.32$ and $w_{VEST} = 0.68$ observed in the current study. This is particularly interesting given that in the current study we were only able to make rough estimates given that we did not explicitly introduce conflicts between these two sources of non-visual information.

Quantifying relative cue weighting during self-motion perception

In general, attempting to evaluate cue integration when studying self-motion perception presents unique challenges compared to other types of cue integration. For instance, one of the assumptions of popular models of multisensory integration (i.e. maximum likelihood estimation) is that the each sensory channel can be assessed independently. There is, however, a tight relationship between visual, proprioceptive and vestibular information during self-motion, which may introduce difficulties in obtaining unbiased unisensory estimates through which to base predictive models (see also Frissen et al. 2011). For instance, when visual self-motion is simulated in the absence of proprioceptive and vestibular inputs (i.e. the *VIS* condition here), the proprioceptive and vestibular systems continue to send information to the brain specifying a stationary position. In other words, because, in healthy adults the vestibular and proprioceptive systems cannot be "turned off", there might be a concern that by providing visual self-motion cues in the absence of movement, this may cause a sensory conflict. For several reasons, however, we do not feel that this concern strongly affects the interpretation of the current results. First, there are clearly common, real-life scenarios in which visual motion is experienced in the absence of robust body-based information (e.g. when riding in a vehicle at a constant velocity) and yet the brain can typically reconcile this with ease. There is also empirical evidence to suggest that purely visual self-motion experiences may not be biased due to "conflicting" vestibular information indicating no movement. Specifically, Gu et al. (2008) demonstrated that in non-human primates there was no difference in visual heading estimates between labyrinthectomized animals and animals with normal vestibular systems. In other words, the removal of this potentially conflicting vestibular information did not change visual heading estimates. This evidence, combined with the fact that the predicted data for the combined cue conditions in the current study was not statistically different from the observed data in both Experiments 1 and 2, provides support for the idea that the approach used here is valid. Further, Frissen et al. (2011) also provides support for the idea that vestibular inputs are also not biased during vestibular alone conditions due to the potentially "conflicting" proprioceptive inputs that are indicating a lack of self-motion. This also makes intuitive sense given that there are everyday scenarios in which vestibular cues are experienced without related proprioceptive information from the legs, including whenever we move our head or when moving in a vehicle.

A great deal of recent research investigating cue integration in other modalities has provided support for the idea that, for several different sensory systems (e.g. visual-auditory, visual-haptic), cues are often combined in a "statistically optimal" manner, such that cue combination leads to the most reliable estimate given the available unisensory cues (e.g. Bülthoff and Yuille 1991; Ernst and Banks 2002; Knill and Saunders 2003; Körding and Wolpert 2004; Ernst and Bülthoff 2004; Alais and Burr 2004; Cheng et al. 2007; MacNeilage et al. 2007). Specifically, the maximum likelihood estimation (MLE) model of sensory integration specifies that information from two or more modalities is combined using a weighted average and the relative weights are based on the relative reliabilities (i.e. inverse of the variance) of the unisensory cues. Consequently, the variance observed in the combined cue conditions will be lower than either unisensory estimate alone. While in the current study, a linearly weighted average was observed, whether integration was "optimal" could not be assessed using the current data for several reasons. For instance, the distance estimates for the unisensory visual and unisensory body-based distance estimates in this study were highly biased (i.e. underestimated and overestimated relative to combined cue conditions respectively), and consequently, the variance scores were also biased (a finding consistent with past studies). Indeed, variable errors are known to increase with increasing distances (Sun et al. 2004b; Lappe and Frenz 2009), which is consistent with Weber's law. Therefore, future experiments should implement designs that will be able to more carefully consider whether sensory integration during self-motion is also "optimal".

Summary and conclusion

Overall, this study highlights the fact that even when cues to self-motion provide redundant information about distance travelled, each contributes to the final estimate. It is also clear that body-based information provides a particularly important role in estimating travelled distance during walking and that both visual and vestibular information contribute to passive self-motion perception with a slightly higher weighting of vision. Finally, both sources of body-based cues contribute to walked distance estimation with a higher weighting attributable to vestibular inputs. More broadly, these collective findings could have significant implications for several basic and applied research areas including the development of multisensory strategies for locomotor rehabilitation, driving and pilot training and assessment, and the optimal development of VR and motion simulator systems.

Acknowledgments We would like to thank Hong-Jin Sun for many helpful discussions and input on the experimental design. We also thank Betty Mohler and Michael Weyel for their technical assistance and informative input, Brian Oliver for his assistance with the experimental set-up and pilot testing, Ilja Frissen for comments on earlier versions of the manuscript, and Simon Musall for his invaluable assistance in collecting the data. The research was supported by funding from the Max Planck Society and by the World Class University (WCU) program funded by the Ministry of Education, Science and Technology through the National Research Foundation of Korea (R31-10008).

References

Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. Curr Biol 14:257–262

Angelaki DE, Cullen KE (2008) Vestibular system: the many facets of a multimodal sense. Annu Rev Neurosci 31:125–150

Angelaki DE, Gu Y, DeAngelis GC (2009) Multisensory integration: psychophysics, neurophysiology, and computation. Curr Opin Neurobiol 19:452–458

Bakker NH, Werkoven PJ, Passenier PO (1999) The effects of proprioceptive and visual feedback on geographical orientation in virtual environments. Pres Teleop Virt Environ 8:36–53

Banton T, Stefanucci J, Durgin F, Fass A, Proffitt D (2005) The perception of walking speed in a virtual environment. Pres Teleop Virt Environ 14(4):394–406

Becker W, Nasios G, Raab S, Jürgens R (2002) Fusion of vestibular and podokinesthetic information during self-turning towards instructed targets. Exp Brain Res 144:458–474

Berthoz A, Israël I, Georges-François P, Grasso R, Tsuzuku T (1995) Spatial memory of body linear displacement: what is being stored? Science 269:95–98

Bertin RJV, Berthoz A (2004) Visuo-vestibular interaction in the reconstruction of travelled trajectories. Exp Brain Res 154:11–21

Bremmer F, Lappe M (1999) The use of optical velocities for distance discrimination and reproduction during visually simulated self motion. Exp Brain Res 127:33–42

Bruggeman H, Piuneu VS, Rieser JJ, Pick HL Jr (2009) Biomechanical versus inertial information: stable individual differences in perception of self-rotation. J Exp Psychol Hum Percept Perform 35:1472–1480

Bülthoff HH, Yuille A (1991) Bayesian models for seeing shapes and depth. Comment Theor Biol 2(4):283–314

Butler JS, Smith S, Campos JL, Bülthoff HH (2010) Bayesian integration of visual and vestibular signals for heading. J Vis 10(12):1–13

Butler JS, Campos JL, Bülthoff HH, Smith ST (2011) The role of stereo vision in visual-vestibular heading estimation. See Perceiv 24(5):453–470

Campos JL, Bülthoff HH (2011) Multisensory integration during self-motion in virtual reality. In: Wallace M, Murray M (eds) Frontiers in the neural bases of multisensory processes. Taylor and Francis Group, London, pp 603–627

Campos JL, Siegle J, Mohler BJ, Loomis JM, Bülthoff HH (2009) Imagined self-motion differs from actual self-motion. PLoS ONE 4(11):e7793. doi:10.1371/journal.pone.0007793

Campos JL, Byrne P, Sun HJ (2010) Body-based cues trump vision when estimating walked distance. Euro J Neurosci 31(10): 1889–1898

Chance SS, Gaunet F, Beall AC, Loomis JM (1998) Locomotion mode affects the updating of objects encountered during travel: the contribution of vestibular and proprioceptive inputs to path integration. Pres Teleop Virt Environ 7(2):168–178

Cheng K, Shettleworth S, Huttenlocher J, Rieser JJ (2007) Bayesian integrating of spatial information. Psychol Bull 133(4):625–637

Durgin FH, Pelah A, Fox LF, Lewis J, Kane R, Walley KA (2005) Self-motion perception during locomotor recalibration: more than meets the eye. J Exp Psychol Hum Percept Perform 31:398–419

Durgin FH, Akagi M, Gallistel CR, Haiken W (2009) The precision of locomotor odometry in humans. Exp Brain Res 193(3):429–436

Elliott D (1986) Continuous visual information may be important after all: a failure to replicate Thomson. J Exp Psychol Hum Percept Perform 12:388–391

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429–433

Ernst MO, Bülthoff HH (2004) Merging the senses into a robust percept. Trends Cogn Sci 8:162–169

Fetsch CR, Turner AH, DeAngelis GC, Angelaki DE (2009) Dynamic reweighting of visual and vestibular cues during self-motion perception. J Neurosci 29(49):15601–15612

Fetsch CR, DeAngelis GC, Angelaki DE (2010) Visual-vestibular cue integration for heading perception: applications of optimal cue integration theory. Euro J Neurosci 31:1721–1729

Frenz H, Lappe M (2005) Absolute travel distance from optic flow. Vision Res 45(13):1679–1692

Frenz H, Lappe M (2006) Visual distance estimation in static compared to moving virtual scenes. Span J Psych 9(2):321–331

Frenz H, Bremmer F, Lappe M (2003) Discrimination of travel distances from “situated” optic flow. Vision Res 43(20):2173–2183

Frissen I, Campos JL, Souman JL, Ernst MO (2011) The relative contributions of proprioceptive and inertial information for self-motion perception. Exp Brain Res 212:163–176

Gepshtein S, Burge J, Ernst MO, Banks MS (2005) The combination of vision and touch depends on spatial proximity. J Vis 5(11): doi:10.1167/5.11.7

Gibson JJ (1950) Perception of the visual world. Houghton Mifflin, Boston

Gu Y, Angelaki DE, DeAngelis GC (2008) Neural correlates of multisensory cue integration in macaque MSTd. Nat Neurosci 11(10):1201–1210

Guerin P, Bardy BG (2008) Optical modulation of locomotion and energy expenditure at preferred transition speed. Exp Brain Res 189:393–402

Harris LR, Jenkin M, Zikovitz DC (2000) Visual and non-visual cues in the perception of linear self-motion. Exp Brain Res 135:12–21

Israël I, Berthoz A (1989) Contributions of the otoliths to the calculation of linear displacement. J Neurophysiol 62(1):247–263

Jürgens R, Becker W (2006) Perception of angular displacement without landmarks: evidence for Bayesian fusion of vestibular, optokinetic, podokinesthetic and cognitive information. Exp Brain Res 174:528–543

Kearns MJ (2003) The roles of vision and body senses in a homing task: the visual environment matters. Dissertation, Brown University

Kearns MJ, Warren WH, Duchon AP, Tarr MJ (2002) Path integration from optic flow and body senses in a homing task. Perception 31:349–374

Klatzky RL, Loomis JM, Beall AC, Chance SS, Golledge RG (1998) Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. Psychol Sci 9(4):293–298

Knapp JM, Loomis JM (2004) Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. Pres Teleop Virt Environ 13(5):572–577

Knill DC, Saunders JA (2003) Do humans optimally integrate stereo and texture information for judgments of surface slant? Vis Res 43:2539–2558

Körding KP, Wolpert DM (2004) Bayesian integration in sensorimotor learning. Nature 427(15):244–247

Körding KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L (2007) Causal inference in multisensory perception. PLoS ONE 2(9):e943. doi:10.1371/journal.pone.0000943

Lappe M, Frenz H (2009) Visual estimation of travel distance during walking. Exp Brain Res 199:369–375

Lappe M, Jenkin M, Harris LR (2007) Travel distance estimation from visual motion by leaky path integration. Exp Brain Res 180(1):35–48

Larish JF, Flach JM (1990) Sources of optical information useful for perception of speed of rectilinear self-motion. J Exp Psychol Hum Percept Perform 16:295–302

Lee DN (1976) Theory of visual control of braking based on information about time-to-collision. Perception 5(4):437–459

Lee DN, Lishman JR (1975) Visual proprioceptive control of stance. J Hum Mov Stud 1:89–95

Loomis JM, Da Silva JA, Fujita N, Fukusima SS (1992) Visual space perception and visually directed action. J Exp Psychol Hum Percept Perform 18:906–921

MacNeilage PR, Banks MS, Berger DR, Bülthoff HH (2007) A Bayesian model of the disambiguation of gravitoinertial force by visual cues. Exp Brain Res 179:263–290

Marlinsky VV (1999) Vestibular and vestibule-proprioceptive perception of motion in the horizontal plane in blindfolded man: I. Estimations of linear displacement. Neuroscience 90:389–394

Mittelstaedt ML, Mittelstaedt H (2001) Idiothetic navigation in humans: estimation of path length. Exp Brain Res 13:318–332

Mohler BJ, Thompson WB, Creem-Regehr SH, Pick HL, Warren WH (2007a) Visual flow influences gait transition speed and preferred walking speed. Exp Brain Res 181(2):221–228

Mohler BJ, Thompson WB, Creem-Regehr SH, Willemsen P, Pick HL, Rieser JJ (2007b) Calibration of locomotion due to visual motion in a treadmill-based virtual environment. ACM Trans Appl Percept 4(1):20–32

Proffitt DR, Stefanucci J, Banton T, Epstein W (2003) The role of effort in perceiving distance. Psychol Sci 14(2):106–112

Prokop T, Schubert M, Berger W (1997) Visual influence on human locomotion. Exp Brain Res 114:63–70

Redlick FP, Jenkin M, Harris LR (2001) Humans can use optic flow to estimate distance of travel. Vision Res 41:213–219

Riecke BE, van Veen HAHC, Bülthoff HH (2002) Visual homing is possible without landmarks—A path integration study in virtual reality. Pres Teleop Virt Environ 11(5):443–473

Riecke BE, Cunningham DW, Bülthoff HH (2006) Spatial updating in virtual reality: the sufficiency of visual information. Psychol Res 71(3):298–313

Rieser JJ, Ashmead DH, Talor CR, Youngquist GA (1990) Visual perception and the guidance of locomotion without vision to previously seen targets. Perception 19(5):675–689

Rieser JJ, Pick HL, Ashmead DH, Garing AE (1995) Calibration of human locomotion and models of perceptual motor organization. J Exp Psychol Hum Percept Perform 21(3):480–497

Seidman SH (2008) Translational motion perception and vestibooc-ular responses in the absence of non-inertial cues. Exp Brain Res 184:13–29

Siegle J, Campos JL, Mohler BJ, Loomis JM, Bülthoff HH (2009) Measurement of instantaneous perceived self-motion using continuous pointing. Exp Brain Res 195(3):429–444

Souman JL, Frissen I, Sreenivasa M, Ernst MO (2009) Walking straight into circles. Curr Biol 19(18):1538–1542

Steck K, Wittlinger M, Wolf H (2009) Estimation of homing distance in desert ants, Cataglyphis fortis, remains unaffected by disturbances of walking behaviour. J Exp Biol 212(18):2893–2901

Steinicke F, Bruder G, Jerald J, Frenz H, Lappe M (2010) Estimation of detection thresholds for redirected walking techniques. IEEE Trans Vis Comp Graph 16(1):17–27

Sun HJ, Frost BJ (1998) Computation of different optical variables of looming objects in pigeon nucleus rotundus neurons. Nat Neurosci 1(4):296–303

Sun HJ, Lee AJ, Campos JL, Chan GSW, Zhang DH (2003) Multisensory integration in speed estimation during self-motion. Cyberpsychol Behav 6(5):509–518

Sun HJ, Campos JL, Chan GSW (2004a) Multisensory integration in the estimation of relative path length. Exp Brain Res 154(2):246–254

Sun HJ, Campos JL, Chan GSW, Young M, Ellard C (2004b) The contributions of static visual cues, nonvisual cues, and optic flow in distance estimation. Perception 33:49–65

Thompson WB, Willemsen P, Gooch AA, Creem-Regehr SH, Loomis JM, Beall AC (2004) Does the quality of the computer graphics matter when judging distances in visually immersive environments? Pres Teleop Virt Environ 13(5):560–571

Thomson JA (1983) Is continuous visual monitoring necessary in visually guided locomotion? J Exp Psychol Hum Percept Perform 9:427–443

Wallace MT, Roberson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA (2004) Unifying multisensory signals across time and space. Exp Brain Res 158:252–258

Waller D, Greenauer N (2007) The role of body-based sensory information in the acquisition of enduring spatial representations. Psychol Res 71(3):322–332

Waller D, Richardson AR (2008) Correcting distance estimates by interacting with immersive virtual environments: effects of task and available sensory information. JEP Appl 14(1):61–72

Warren WH, Hannon DJ (1998) Direction of self-motion is perceived from optical-flow. Nature 336:162–163

Warren WH, Kay BA, Zosh WD, Duchon AP, Sahuc S (2001) Optic flow is used to control human walking. Nat Neurosci 4:213–216

Wilkie RM, Wann JP (2005) The role of visual and nonvisual information in the control of locomotion. J Exp Psychol Hum Percept Perform 31(5):901–911

Witmer BG, Kline PB (1998) Judging perceived and traversed distance in virtual environments. Pres Teleop Virt Environ 7:144–167

Wittlinger M, Wehner R, Wolf H (2006) The ant odometer: stepping on stilts and stumps. Science 312:1965–1967

Wohlgemuth S, Ronacher B, Wehner R (2001) Ant odometry in the third dimension. Nature 411(6839):795–798

Yong NA, Paige GD, Seidman SH (2007) Multiple sensory cues underlying the perception of translation and path. J Neurophysiol 97:1100–1113