



Published in final edited form as:

Exp Brain Res. 2011 April ; 210(2): 173–184. doi:10.1007/s00221-011-2612-4.

Perceived tilt and translation during variable-radius swing motion with congruent or conflicting visual and vestibular cues

Andrew A. Rader,

Man Vehicle Lab, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, USA

Charles M. Oman, and

Man Vehicle Lab, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, USA

Daniel M. Merfeld

Jenks Vestibular Research Laboratory, Massachusetts Eye and Ear Infirmary, Boston, MA, USA

Daniel M. Merfeld: dan_merfeld@meei.harvard.edu

Abstract

We examined the influence of dynamic visual scenes on the motion perception of subjects undergoing sinusoidal (0.45 Hz) roll swing motion at different radii. The visual scenes were presented on a flatscreen monitor with a monocular 40° field of view. There were 3 categories of trials: 1. Trials in the dark; 2. Trials where the visual scene matched the actual motion; and 3. Trials where the visual scene showed swing motion at a different radius. Subjects verbally reported perceptions of head tilt and translation. When the visual and vestibular cues differed, subjects reported perceptions that were geometrically consistent with a radius between the radii of the visual scene and the actual motion. Even when sensations did not match either the visual or vestibular stimuli, reported motion perceptions were consistent with swing motions combining elements of each. Subjects were generally unable to detect cue conflicts or judge their own visual-vestibular biases, which suggests that the visual and vestibular self-motion cues are not independently accessible.

Keywords

Visual; Vestibular; Perception; Human; Swing; Null

Introduction

Spatial orientation includes central integration of ambiguous motion cues (Lackner and DiZio 2005; Mayne 1974; Young 1984) from both sensory (e.g., vision, vestibular organs, and proprioception), and non-sensory (e.g., efferent copy, cognition) sources. As an example of ambiguous sensory cues, the otolith organs cannot intrinsically separate linear acceleration from gravity; like all graviceptors/linear accelerometers, they measure gravito-inertial force, defined as the difference between gravity and acceleration ($GIF = g - a$). Additional information is thus required in order to disambiguate head tilt from head

translation in darkness (Angelaki et al. 1999; Droulez and Darlot 1989; Glasauer 1995; Merfeld et al. 1993).

As with all studies that investigate graviceptor responses using dynamic motion stimuli (Angelaki et al. 1999; Glasauer 1995; Merfeld et al. 2005a; Merfeld et al. 2005b; Park et al. 2006), matters are complicated by the fact that the otolith organs are not our only graviceptors. Specifically, Mittelstaedt and colleagues have reported that truncal receptors influence perceived body tilt (posture) without affecting either perceived subjective visual vertical (perceived orientation of a visual line with respect to the perceived direction of gravity) or perceived head tilt (Mittelstaedt 1999; Mittelstaedt 1997). Given these reports that non-vestibular graviception has a reduced influence on head-bound perception, studies have focused primarily on vestibular perception (arising from forces and accelerations acting on the inner ear) by asking subjects to focus full attention on their head translation and head tilt (Merfeld et al. 2005a; Merfeld et al. 2005b; Park et al. 2006; Rader et al. 2009).

The most recent of these studies (Rader et al. 2009) reported that perceived head motion for roll-tilt in the absence of visual cues (i.e., in the dark) represented a geometrically consistent compromise of available sensory cues. Specifically, although subjects reported their motion with significant errors for some cue combinations, motion perception was consistent with the geometric constraints of swing motion. This finding supports earlier studies that have demonstrated a strong influence of cognition on tilt/translation resolution (Wertheim et al. 2001; Wright et al. 2006). Might similar cognitive influences be found for visual and vestibular cue combinations during swing motion, and how would visual-vestibular conflicts interact with these effects?

We designed a set of studies to answer these questions by combining dynamic visual cues with dynamic roll swing motion. While much is known about the contributions of static visual cues to tilt/translation resolution (e.g., Asch and Witkin 1948; Bagust et al. 2005; Howard 1982; Li and Martin 1995; Singer et al. 1970; Vingerhoets et al. 2008) and the contributions of dynamic visual cues with the body stationary (e.g., Benson 1978; Dichgans et al. 1972; Howard and Heckmann 1989; Wright et al. 2006; Yasui and Young 1975; Zupan and Merfeld 2003), only a few studies (Gresty et al. 2003; Wright et al. 2005; Wright et al. 2009) have examined how dynamic visual cues contribute to tilt/translation resolution during dynamic whole body motion.

This study expands upon these previous dynamic visual-vestibular interaction studies in the following ways: 1) The swing motions blend tilt and translation in a recognizable combination; 2) The stimulation frequency (0.45 Hz) was about twice that of earlier visual-vestibular studies, putting it in a range where both visual and vestibular cues are key contributors; 3) We included profiles where the inter-aural gravito-inertial force was perceptually “nulled”; 4) Even when the visual and vestibular stimuli were in conflict, the visual stimulus was compatible with motion that the motion device *could* provide at another radius. This emphasizes the “believability” of the visual cues in order to maximize their influence on motion perception (Fig. 1).

We hypothesized that subjects would report motion that was geometrically consistent with swing motion, but inconsistent with either the actual motion (Fig. 1B) or the visual stimuli (Fig. 1C) when these were inconsistent with each other (Fig. 1A). Specifically, we hypothesized that the subjects would perceive swing motion (Fig. 1D) that reflected a compromise between the visual and vestibular cues. Furthermore, because visual, vestibular, and efferent interactions occur very early - in the vestibular nuclei (Cullen 2004; Cullen and Minor 2002; Henn et al. 1980) - we hypothesized that subjects would be unable to access information regarding the lack of consistency between the cues. This would contribute to subjects being unable to distinguish if their perception was more consistent with the visual stimulus or actual motion, or determine when the cues were in conflict.

Methods

Using a servo-controlled swing (Massachusetts Eye and Ear Infirmary “tilt device”, Neurokinetics, Inc, Pittsburgh, PA), we generated roll tilt motion profiles with sinusoidal time courses. These oscillations were symmetric about the upright position and combined tilt and translation to generate a y-axis (see coordinate frame, Fig. 2a) GIF that varies with radial arm length. A description of all motion profiles tested can be found in Table 1, and the profile classes are graphically presented (along with some results) in Figure 3.

A frequency of 0.45 Hz and tilt angle of 10° were selected for all trials so results could be compared with those from our previous study. Each trial included ramp-up and ramp-down periods of about one minute each, with steady-state motion for about one and a half minutes between. The subject’s ear canal position with respect to the rotation center was varied from 20 cm above (+20 cm) to 122 cm below (-122 cm).

In trials where a subject’s ear was 122 cm below the center of rotation, their y-axis otolith GIF was effectively “nulled” below threshold. Although body graviceptors would still be expected to provide some information at this position, the trunk-level dynamic y-axis GIF was significantly reduced compared with normal swing motion. For example, the *peak* y-axis acceleration at body-level did not exceed 0.6 m/s^2 , less than half the 1.3 m/s^2 *steady-state* value shown by Mittelstaedt to influence *body* tilt (Mittelstaedt and Mittelstaedt 1996). Moreover, we specifically asked subjects to report their motion with respect to the bridge of their nose to emphasize head motion perception.

Subjects were restrained by torso side clamps, a 5 point harness, and by an MRI shoulder and head mask (Aquaplast, Inc. thermoplastic 4.8mm disposable S-frame; part number RT-1992SD) individually moulded to each subject. Breeze cues were minimized with long pants, a long-sleeve shirt, a balaclava, and gloves.

Visual scenes

Some trials were performed in complete darkness. For other trials, a visual scene was displayed on a flatscreen monitor showing a pre-recorded video – created by placing a video camera on the tilt device at subject eye position - depicting the same motion a subject would see if they were watching from that radius in a lighted room.

There were three types of experimental trials:

1. no visual scene: (Dark; D),
2. visual scene showing the actual radius (No conflict; NC), and
3. visual scene showing a different radius (Conflict; C).

The monitor was mounted in front of the subjects at a distance of approximately 38 cm from the pupil in order to maximize the angle of view (AOV). Studies have shown that AOV is a critical factor in visual scene influence, with a strong sense of vection beyond 30° (Paige et al. 1998). A black cardboard mask with a circular aperture was affixed to the screen to create a “porthole” window 28 cm in diameter, resulting in a 40° AOV (Figs. 2b&2c).

The visual scenes were recorded at each experimental radius using a Kodak Easyshare C813 camera mounted at the subject’s eye position. This camera recorded QVGA 320×240 pixel resolution images at 30 fps in QT-JPEG format. Timing between the recorded scene and actual motion was precisely controlled by using an experimentally-tuned electronic impulse trigger, and was verified using a third-person video that showed the visual scene at the same time as the actual motion (<17 msec, or less than 1 frame error).

Throughout the experiment, the subjects and the flatscreen monitor were enclosed in a shroud in the dark so that the only source of light in any trial was the visual scene projected in front of them. In order to minimize vergence cues, subjects viewed the scene monocularly with their dominant eye; an eye patch covered the other eye.

Subjects and instructions

Ten healthy subjects (4 male, 6 female; age 23–52 or 32 ± 9 yrs) were pre-screened as “normal” via standard clinical vestibular tests (caloric, rotation, EquiTest posturography, and Hallpike positional testing). All subjects provided informed consent. Six subjects (3 male, 3 female) had participated in a previous separate companion experiment (Rader et al. 2009) that used similar motion profiles without visual cues. All ten subjects performed three sessions, totalling 36 trials (3×12), for this study.

Before starting the experiment, subjects watched a presentation on the monitor that familiarized them with the range of motions of the device and the instructions on how to report their perception of motion. Specifically, three questions were asked during training and during the experiment:

1. What is the maximal side-to-side translation you are experiencing?
2. If you feel like you are tilting, what is your best estimate of the distance to the center of rotation?
3. What is the maximal head tilt to either side you are experiencing?

Subjects were told that the motion would be symmetric, and to report using any cues they have available, including the visual scene and their “internal sense”. The majority of the subjects were most familiar with estimating small distances in inches, so all subjects were

required to report all distances in inches and all angles in degrees, although distances have been converted into SI units for presentation.

For each of the questions, the subjects were also asked about their level of confidence in their answer in order to determine if a visual-vestibular conflict, if detected at all, might manifest as a reduction of subjective confidence. Confidence was the chance out of ten that their answer should be correct to within 2° (for tilt), or 5 inches (for translation/radius). Confidence was plotted as a fractional value between 0 and 1 (i.e., if subject reported 2, this was plotted as 0.2).

For comparison, correctness (0–1) was computed as a ratio of how many responses were within the specified range (2° or 5 inches) divided by the total number of responses. Note that subjects would be perfectly judging their ability to perform the task if their confidence equalled their correctness. Subjects were considered overconfident if their confidence was higher than their correctness.

In addition to verbal reports, dynamic tilt perception was recorded by having subjects keep a somatosensory tilt bar (Park et al. 2006) level with their perceived Earth horizontal at all times. The tilt bar signals were measured using potentiometer analog data at 60 Hz (angular discrimination < 0.1°). This task also served to maintain subject alertness and verify compliance.

Experiment design

The experiment consisted of three sessions of 12 randomly ordered trials. The order was changed for the second and third sessions (Table 1) in order to examine and balance out order effects. There was also a training session immediately preceding each experimental session.

Magnitude estimation and training

We focused on obtaining reliable verbal reports because we have previously found that subjects are able to reliably report their motion as long as 1) tasks are precisely-defined, and 2) a focused training regime is carefully designed not to bias the results (Rader et al. 2009). Subject reports were based on a direct magnitude estimation task using real values (inches) rather than ratios. However, without training, this method often yields non-linear scaling errors as the expansion of the perception domain increase with the increase in stimulus magnitude (Class I stimulus as defined by Stevens for his psychophysical law: Sensation = $A \times (\text{Stimulus})^n$ (Stevens 1957). In fact, no subject can be considered a blank slate, and all start with different degrees of familiarity with 1) experiencing these types of motions, 2) making these types of perceptual judgements, and 3) reporting their perceptions. Thus it is essential to train subjects to reduce inter-subject variability and familiarize them with the experimental reporting method. Furthermore, in our earlier study (Rader et al. 2009), we confirmed that untrained naïve subjects had much higher reporting variance than trained subjects, though the mean reports were not significantly different without training.

Subjects were trained using 12 representative motion profiles (6 light/6 dark) for peak tilt angles between 5° and 25°, even though only 10° was examined in the experiment (Table 1).

To avoid introducing biases, subjects were not trained with cue conflicts. During training, subjects received feedback on their answers.

The training was successful in helping subjects to accurately report their perceived motion in light and dark; while experienced subjects initially demonstrated less variability, by the end of the first training session all subjects were consistently reporting with a correctness of over 50% (up from 10% in some cases for less experienced subjects).

Post-test survey and rod and frame test

A survey was administered after the third experimental session asking subjects how well they thought they performed overall on estimation of tilt angle, translation, and radius, and asking them how much they thought they were influenced by 1) the visual scenes, and 2) their “internal sense”. Subjects also completed a computerized rod and frame test (Grabherr et al. 2008) which has demonstrated results consistent with other computerized rod and frame tests (Bagust et al. 2005; Lopez et al. 2006) that show only slightly smaller field dependence than the classic test (Asch and Witkin 1948).

Hierarchical mixed regression analysis

A two-stage hierarchical mixed regression (HMR) model (SYSTAT, version 12, SYSTAT Software, Inc.) estimated the effects of the independent variables (subject, actual radius, y-axis nulling, and viewing radius, etc.) on the dependent measures (perceived tilt angle, perceived horizontal translation, and perceived radius) reported by subjects and averaged over each experimental session. Two-Stage Least Squares (TSLS) regression was used to account for each subject’s random effect (White 1982). In the first stage, the independent variables were regressed on the subject (treated as an instrumental variable). In the second stage, the dependent variables were regressed on all the variables, including the values of the independent variables estimated in the first stage. TSLS produces heteroskedasticity-consistent standard errors for instrumental variable models (SYSTAT).

For each dependent variable, the analysis minimized the residuals at each step to fit linear models of the form

$$Y_j = [(A + S_i) + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon_{i,n}]_j,$$

where Y_j is one of the measured dependent variables (perceived tilt angle, horizontal translation, or radius), $(A + S_i)$ is the grand mean plus a random subject effect (“intercept”) from the first regression stage, the x_k are independent variables, the β_k are the corresponding estimated coefficients, and $\varepsilon_{i,k}$ is noise. Each coefficient gives the slope of the measured variable (e.g., reported tilt) against the corresponding variable, x_k .

The principal independent variables were actual radius, viewing radius, and a categorical variable (y-axis nulling: yes = 1 or no = 0) was used. In addition, gender, age, session number, trial order within session, subject weight, subject field dependence, and various answers to post-test questionnaire were also examined. We also examined regression models

examining the cross-effect terms for the independent variables (e.g., viewing radius \times y-axis nulling).

Repeated measures ANOVA

For horizontal translation and radius reports, we performed an ordinary repeated measures ANOVA on the dependent variable correctness with the two within-subject factors ‘visual condition’ (conflict, no conflict, dark), and ‘type of estimate’ (subjective, objective).

Paired t-test

We tested geometric consistency using a paired t-test, where reported radii were compared with radii calculated geometrically from reported horizontal translation and tilt angle. For this analysis, trials with reported tilt angles of less than 5 degrees (around just 5% of total cases) were not included because their computed radii became very high (8 or more standard deviations from the mean).

Results

Overview of main findings

The results for the “Dark” conditions were similar to those from our previous experiment (Rader et al. 2009). Perception of motion was generally veridical, except that y-axis nulling increased tilt perception while decreasing horizontal translation and radius perception.

When the visual scene was present, manipulation of both the visual scene (i.e. viewing radius) and vestibular cues (actual radius) had strong effects on subject reports (Fig. 3). Overall, the viewing radius had a greater influence than the actual radius on subjective reports of perceived tilt angle, horizontal translation, and radius).

Hierarchical mixed regression analysis

We found no significant effects of trial-order, gender, age, or anthropometry. There were no consistent significant effects of session or order, but some subjects performed noticeably less well in certain sessions. These “off days” were few in number and did not follow a consistent pattern.

In models of tilt angle, horizontal translation, and radius perception (Table 2) we found significant effects of actual radius, viewing radius, and y-axis nulling in 5 out of a possible 9 cases. Note that 2 of the 4 cases that were not significant were the effect of actual radius on perceived tilt and the effect of viewing radius on perceived tilt. These effects were not expected to be significant since actual tilt angle was not dependent on either radius. The random subject effects were not significant overall for perceived radius, and the model accounted for more than 95% of the variance. We did not find significance of any cross-effects of the independent variables (e.g., $p = 0.26$ for viewing radius \times y-axis nulling).

Angle of tilt

As shown in Fig. 4A, subjects overestimated their reported angle of tilt slightly; most reports from individual trials fell in the 11° to 14° range, whereas the actual angle was 10° for all

trials. Tilt angle reports were not significantly different from the 0.45 Hz data from the previous experiment in the dark (Rader et al. 2009). As shown by the HMR model, y-axis nulling significantly increased reported tilt angle ($p < 0.01$). In all but one case, reports tended to be slightly closer to the real angle of 10° when visual and vestibular radii agreed, but this effect was not significant. We found no significant effect of varying the visual radius on perceived tilt, but we would not expect to find such an effect since the visual scene showed the same tilt angle in all cases.

Neither confidence nor correctness were significantly affected by the absence of a visual scene (Dark), or whether or not the viewing radius agreed with the actual radius (Conflict/No Conflict) (Fig. 4B). Subjects were significantly overconfident (i.e, confidence greater than correctness) for all 3 visual conditions ($p < 0.05$ for each).

Horizontal translation

Subjects reported their horizontal translation accurately for dark trials lacking y-axis nulling, and for trials with the visual scene in agreement with the actual radius (Fig. 5A). The HMR presented previously showed that y-axis nulling caused subjects to underestimate their horizontal translation by 4.3 cm ($p < 0.01$). The visual scene acted as a strong attractant of subject reports as indicated by the regression analysis ($p < 0.01$) and by the fact that translation estimates were strongly skewed towards the translation of the visual scene in the “conflict” case (Fig. 5A). However, we did not find a significant effect of the actual radius on translation estimates ($p = 0.26$). Overall, when the visual scene agreed with the actual motion (“no conflict” cases), it improved reports (Fig. 5B). However, when the visual scene was different from the actual radius (“conflict”), it acted as a confound.

Confidence was not affected by the absence of a visual scene (Dark), or whether or not the viewing radius agreed with the actual radius (Conflict/No Conflict) (Fig. 5B). However, subjects were much more likely to be correct (within ± 5 inches) in the “no conflict” cases, and much less likely to be correct in the “conflict” cases (all $p < 0.01$). The ANOVA on correctness demonstrated that significant differences existed between conflict, no conflict, and dark cases in terms of their effect on subjective confidence and objective correctness ($p < 0.01$).

Roll radius

As with translation, radius reports were fairly accurate when there was no y-axis nulling for both the “dark” and “no conflict” cases (Fig. 6A). The visual scene acted as a strong attractant of subject radius estimates as indicated by the regression analysis ($p < 0.01$) and by the fact that radius estimates were strongly skewed towards the viewing radius in the “conflict” cases (Fig. 6A). By contrast, the actual radius acted as a significantly weaker (though still highly significant) attractant of subject radius estimates ($p < 0.01$). As a result, when the viewing radius agreed with the actual radius (“no conflict” cases), it improved reports. However, when the viewing radius was different from the actual radius (“conflict”), it acted as a confound.

As with angle of tilt and horizontal translation, radius confidence did not vary between “dark”, “no conflict”, and “conflict” cases (Fig. 6B). However, subjects were much more

likely to be correct (within ± 5 inches) in the “no conflict” case, and much less likely to be correct in the “conflict” case (all $p < 0.01$). The ANOVA on correctness demonstrated that significant differences existed between conflict, no conflict, and dark cases in terms of their effect on subjective confidence and objective correctness ($p < 0.01$).

Overall influence of the visual scene

In trials with visual cues, subjects reported translations and radii closer to what was displayed visually than to their actual motion: i.e., the viewing radius had significantly more influence than the actual radius on perceived horizontal translation and perceived radius (both $p < 0.05$). To demonstrate this effect graphically, Fig. 7 shows the average difference between the reported radius and the viewing radius arranged next to the average difference between the reported radius and the actual radius. A similar plot for horizontal translation is also shown in Fig. 7. Note that the differences between the reported values and visual values are significantly smaller than the differences between the reported values and actual values for both radius and horizontal translation.

Other results

Variability during initial training did not significantly correlate with later variability, which suggests that the training successfully reduced variability associated with lack of prior experience and is consistent with our earlier findings (Rader et al. 2009). As in this earlier study, the somatosensory bar results were consistent with the tilt angle verbal reports for all profiles.

During the experiment, only 4 subjects reported any motion sickness symptoms, and only 1 reported a value higher than “2” on a 0–20 scale (a “9” in one trial). This report, as well as over 70% of all reports of motion sickness, occurred during “conflict” trials. We found no significant correlation of field dependence (as determined by the rod and frame test) with either performance in making estimates, or with influence of the visual scene. On the post-test surveys, all subjects reported that their “other senses” had at least as much influence on their estimates as the visual scene, and some reported that they had far more influence (contrary to the actual data, which showed visual stimuli when present having a larger influence on average than motion cues).

Discussion

We found significant effects of y-axis nulling, the actual radius, and the viewing radius on subject motion perception. In all cases, the visual radius had significantly more influence on subject motion perception than the actual radius.

Tilt angle, translation, and radius without visual scene

Subjects overestimated their tilt angle, especially in cases where the y-axis forces were nulled. This was also noted in our previous study (Rader et al. 2009), and is consistent with swing and passive yaw studies that have highlighted a tendency for subjects to overestimate tilt and translation in the dark - possibly as a safety mechanism in an unfamiliar environment

(Guedry and Harris 1963; Israel et al. 1995; Ivanenko et al. 1997; Jongkees and Groen 1946).

In dark trials, translation and radius estimates were nearly veridical when there was no y-axis nulling, and were significantly underestimated with y-axis nulling present. This suggests that the lack of confirming y-axis otolith cues did influence the perception of motion in favour of being closer to the center of rotation and at a larger angle.

A lack of y-axis otolith information has two possible interpretations: (a) no motion in the y-axis, or (b) swing motion at exactly the correct set of radius, tilt angle, and frequency. While interpretation (b) was in fact the actual motion case in the experiment, interpretation (a) is arguably more likely in everyday life since few motions combine the exact radius, frequency, and angle to achieve near cancellation of y-axis forces. Such prior estimates of probability of different classes of motion could influence neural processing towards solutions favouring motions of lower amplitude when y-axis forces are nulled. Combined with a preferential focus on the extant cues (e.g., semicircular canal cues), this could drive perceptions towards those observed in the experiment. This suggests that while internal models are able to correctly resolve motions even if some cues (e.g. y-axis otolith signals) are absent, the implementation may be imperfect and rely to some extent on weighted processing, filtering (i.e. frequency segregation (Mayne 1974)), experience, or cognition.

Geometric consistency and cognitive influences

Even when the visual and actual radius were significantly different and subjects made large perceptual errors, their reports of tilt, translation, and radius were *internally* geometrically consistent with swing motions that the tilt device could actually generate. More specifically, using any two of the three quantities - angle of tilt, horizontal translation, and radius - one could correctly calculate the third. When questioned afterwards, subjects indicated that they did not perform this calculation overtly; and yet their reports did not deviate significantly from the geometric constraint across any of the cases (confirmed by paired t-tests on geometric *inconsistency*: dark $p = 0.41$; conflict $p = 0.14$; no conflict $p = 0.42$). Thus, subjects did not report radii significantly different from values that were geometrically consistent with their reported tilt angles and horizontal translations.

This geometric consistency was likely built up as a cognitive influence on perception through knowledge of the tilt device capabilities acquired during informed consent and in training. Similar cognitive influences have been reported in linear acceleration perception studies (Wertheim et al. 2001; Wright et al. 2006), and is particularly strong when subjects actually see the motion device before or during the experiment (Israel et al. 1993).

Influence of the visual scene

The visual scene was the strongest attractant of subject reports in all cases. This indicates that our pre-recorded visual scene was sufficiently immersive to strongly drive perception, even when combined with completely different vestibular cues. Overall, subject reports were significantly more consistent with the viewing radius than with the actual radius. These results are an extension of previous studies that have demonstrated the dominant role played

by vision during tilt or translation motion (Wright et al. 2006; Wright et al. 2009). However, this contrasts with the post-test questionnaire results, where all ten subjects rated “influence of internal sense” as equal to, or greater than, “influence of visual scene”.

Even though both y-axis nulling and the visual scene were individually highly significant, we found no cross-effect ($p = 0.26$), and thus surmise that both effects are perceived independently. In other words, the effectiveness of the visual scene was neither amplified nor mitigated by the presence of y-axis nulling, and vice-versa. This suggests that the perceptual contribution of the visual scene did not increase when it could have been used to correct perceptual errors – probably because the subjects could not distinguish trials that included y-axis nulling from those that didn't.

Moreover, influence of the visual scene was neither correlated with overall performance nor field dependence as established by the rod and frame test. Two subjects (7 & 10 in Fig. 7) did demonstrate a small tendency to report perceptions slightly more consistent with the actual radius than the visual scene, but these subjects did so for different reasons. Subject 7 had the lowest overall reporting errors across the three cases (dark, no conflict, and conflict); it is possible that this subject was actually more influenced by their vestibular and proprioceptive cues, but this is impossible to establish. Subject 10 reported a conscious effort to reject the information displayed by the visual scene, and as a result ended up with the highest overall errors of any subject because the visual scene agreed with the correct radius as often as it differed.

Visual-vestibular conflict

Confidence varied remarkably little across trials (see standard error bars in Fig. 4B, 5B, and 6B), and was not affected by the presence of cue conflict. By contrast, cue conflict dramatically reduced a subject's likelihood of being correct (as demonstrated by the ANOVAs as well as individual t-tests). This supports the idea that motion perception and the judgement of the degree to which a set of cues make up a compelling perception may be driven by dissociable processes (Wright et al. 2006).

Some previous experiments have demonstrated that subjects can detect conflict between projected visual scenes and device motion (Berger and Bulthoff 2009; Butler et al. 2006), although this was for large differences in translation gain factors, or even motion in a different direction. In the present experiment, at least the tilt angle and frequency cues were congruent even in the “conflict” case. Moreover, the baseline ability of subjects in the studies that have reported an ability to identify conflicts only correctly identify cases without conflict when none is present is only around 55% of the time (Berger and Bulthoff 2009) – near the overall confidence levels of subjects in this experiment.

Not only were subjects unable to reliably identify trials with cue conflicts, subjects were apparently not aware of their biases toward the visual scenes. Seven out of ten subjects reported that they were less influenced by the visual scene (with the remaining 3 reporting equal influence); however the opposite was in fact true. Further, subjects were not able to judge their performance in estimating horizontal translation or roll radius. We found no correlation between subjects who thought they performed well or poorly at the estimation

tasks and those who actually did perform well or poorly. In other words, subjects as a whole failed to reliably assess their ability to perform the estimation tasks. In fact, we found no reliable predictor of performance or of visual scene influence (experience, gender, age, post-test self-assessment, field dependence, etc.). This carries profound implications for situations where people may experience conflicting sensory cues without realizing it.

Summary

The results showed that subjects were influenced by both the visual scene and their actual motion, although the visual cues were more influential. Subjects generally reported perceptions consistent with radii somewhere between the radius of the visual scene and the actual radius in a manner consistent with motions that could be actually produced by the tilt device, highlighting the cognitive influences on motion.

These results suggest that subjects used internal models to combine all visual and vestibular cues, along with known geometric constraints, to estimate their overall motion. Moreover, subjects were not in general able to overtly detect cue conflicts, nor judge their biases (visual *versus* vestibular) or performance, suggesting that this information is not accessible at a conscious level.

Acknowledgments

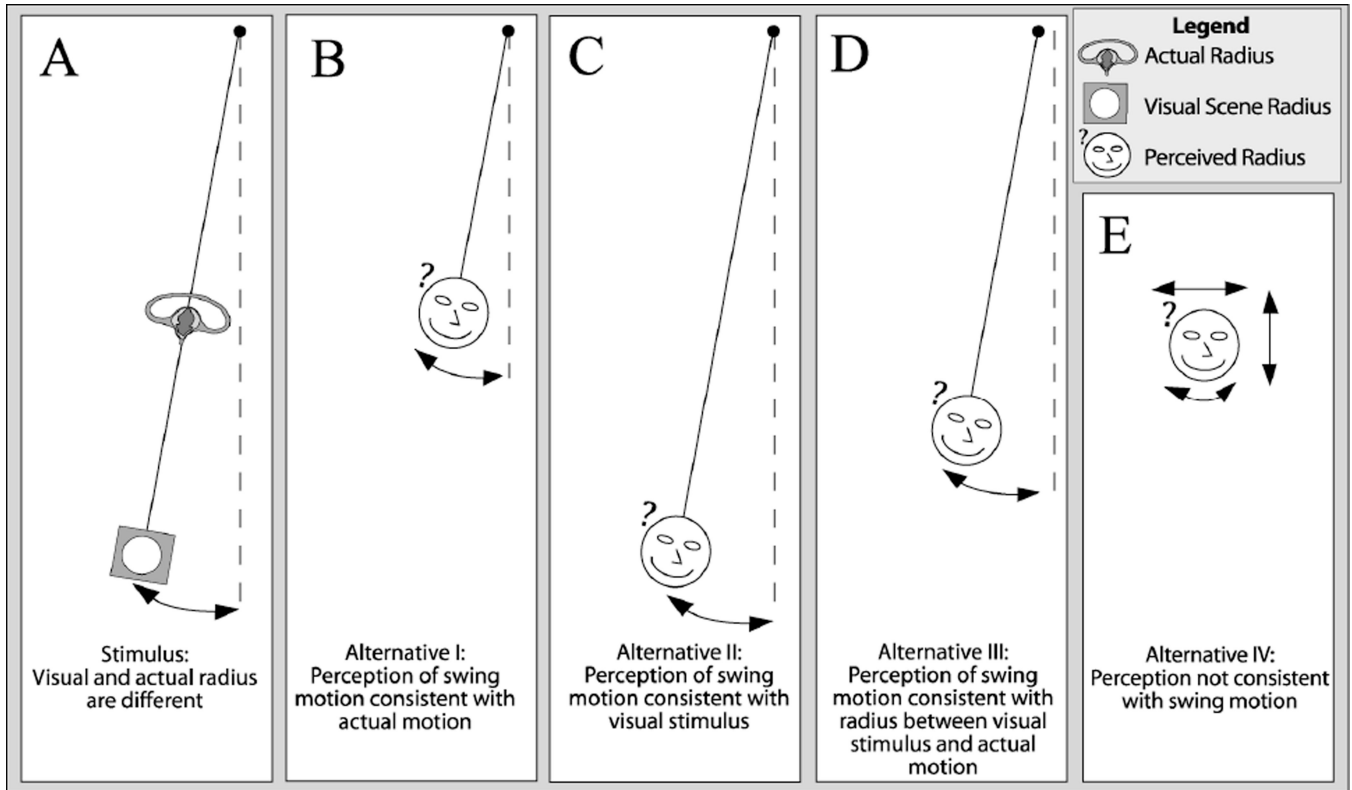
Funding for this experiment was generously provided by NASA (NNJ04HF79G) and NIH (NIDCD R01 DC04158). The authors would also like to thank the subjects who participated. This experiment protocol was approved by the MEEI and MIT IRBs.

References

- Angelaki DE, McHenry MQ, Dickman DD, Newlands SD, Hess BJM. Computation of inertial motion: neural strategies to resolve ambiguous otolith information. *J Neurosci*. 1999; 19:316–327. [PubMed: 9870961]
- Asch SE, Witkin HA. Studies in space orientation. II. Perception of the upright with displaced visual fields and with body tilted. *Journal of Experimental Psychology*. 1948; 38:455–477. [PubMed: 18874603]
- Bagust J, Rix GD, Hurst HC. Use of a Computer Rod and Frame (CRAF) Test to assess errors in the perception of a visual vertical in a clinical setting - A pilot study. *Clinical Chiropractic*. 2005; 8:134–139.
- Benson, AJ. Perceptual Illusions. In: Dhenin, G., editor. *Aviation Medicine*. London: Tri-Med Books; 1978.
- Berger DR, Bulthoff HH. The role of attention on the integration of visual and inertial cues. *Exp Brain Res*. 2009 (accepted).
- Butler, J.; Smith, ST.; Beykirch, K.; Bulthoff, HH. DSC Europe. Paris: 2006. *Visual Vestibular Interactions for Self Motion Estimation*.
- Cullen KE. Sensory signals during active versus passive movement. *Curr Opin Neurobiol*. 2004; 14:698–706. [PubMed: 15582371]
- Cullen KE, Minor LB. Semicircular Canal Afferents Similarly Encode Active and Passive Head-On-Body Rotations: Implications for the Role of Vestibular Efference. *Journal of Neuroscience*. 2002; 22:1–7. [PubMed: 11756482]
- Dichgans J, Held R, Young LR, Brandt T. Moving visual scenes influence the apparent direction of gravity. *Science*. 1972; 178

- Droulez, J.; Darlot, C. The geometric and dynamic implications of the coherence constraints in three-dimensional sensorimotor interactions. In: Jeannerod, M., editor. *Attention and performance*. New Jersey: Lawrence Erlbaum; 1989. p. 495-562.
- Glasauer S. Linear acceleration perception: frequency dependence of the hilltop illusion. *Acta Otolaryngol Suppl.* 1995; 520:37–40. [PubMed: 8749075]
- Grabherr, L.; Cuffel, C.; Guyot, J-P.; Mast, FW. FENS. Geneva, Switzerland: 2008. *Mental transformation abilities in unilateral and bilateral vestibular patients.*
- Gresty MA, Waters S, Bray A, Bunday K, Golding JF. Impairment of Spatial Cognitive Function With Preservation of Verbal Performance During Spatial Disorientation. *Current Biology.* 2003; 13:R829–R830. [PubMed: 14588257]
- Guedry, FE.; Harris, CS. *Labyrinthine function related to experiments on the parallel swing Pensacola. FL: U.S. Naval school of aviation medicine; 1963.*
- Henn V, Cohen B, Young LR. Visual vestibular interaction in motion perception and the generation of nystagmus. *Neurosci Res Program Bull.* 1980; 18:459–651.
- Howard, IP. *Human visual orientation.* New York: 1982.
- Howard IP, Heckmann T. Circular vection as a function of the relative sizes, distances, and positions of two competing visual displays. *Perception.* 1989; 18:657–665. [PubMed: 2602091]
- Israel I, Chapuis N, Glasauer S, Charade O, Berthoz A. Estimation of Passive Horizontal Linear Whole-Body Displacement in Humans. *J Neurophys.* 1993; 70:1270–1273.
- Israel I, Sievering D, Koenig E. Self-rotation estimate about the vertical axis. *Acta Otolaryngol.* 1995; 115:3–8. [PubMed: 7762380]
- Ivanenko Y, Grosso R, Isreal I, Berthoz A. Spatial orientation in humans: perception of angular whole-body displacements in two-dimensional trajectories. *Exp Brain Res.* 1997; 117:419–427. [PubMed: 9438709]
- Jongkees LBW, Groen JJ. The Nature of Vestibular Stimulus. *The Journal of Laryngology and Otology.* 1946; 38:529–541. [PubMed: 20260393]
- Lackner JR, DiZio P. Vestibular, Proprioceptive, and Haptic Contributions to Spatial Orientation. *Annu Rev Psychol.* 2005;115–147. [PubMed: 15709931]
- Li W, Matin L. Differences in influence between pitched-from-vertical and slanted-from-frontal horizontal lines on egocentric localization. *Perception & Psychophysics.* 1995; 40:311–316.
- Lopez C, Lacour M, Magnan J, Borel L. Visual field dependence-independence before and after unilateral vestibular loss. *Auditory and Vestibular Systems.* 2006; 17:797–803.
- Mayne, R. A systems concept of the vestibular organs. In: Kornhuber, HH., editor. *Handbook of Sensory Physiology Vestibular System Psychophysics, Applied Aspects and General Interpretations.* 1974. p. 493-580.
- Merfeld DM, Park S, Gianna-Poulin C, Black FO, Wood S. Vestibular Perception and Action Employ Qualitatively Different Mechanisms. I. Frequency Response of VOR and Perceptual Responses During Translation and Tilt. *J Neurophysiol.* 2005a; 94:186–198. [PubMed: 15728767]
- Merfeld DM, Park S, Gianna-Poulin C, Owen-Black F, Wood S. Vestibular Perception and Action Employ Qualitatively Different Mechanisms. II. VOR and Perceptual Responses During Combined Tilt & Translation. *J Neurophysiol.* 2005b; 94
- Merfeld DM, Young LR, Oman CM, Shelhamer MJ. A Multidimensional Model of the Effect of Gravity on the Spatial Orientation of the Monkey. *Journal of Vestibular Research.* 1993; 3:141–161. [PubMed: 8275250]
- Mittelstaedt H. The Role of the Otoliths in Perception of the Vertical and in Path Integration. *Annals of the New York Academy of Sciences.* 1999; 871:334–344. [PubMed: 10372083]
- Mittelstaedt ML, Mittelstaedt H. The Influence of Otoliths and Somatic Graviceptors on Angular Velocity Estimation. *J Vestib Res.* 1996; 6:355–366. [PubMed: 8887892]
- Mittelstaedt H. Interaction of eye-, head-, and trunk-bound information in spatial perception and control. *J Vestib Res.* 1997; 7:283–302. [PubMed: 9218243]
- Paige GD, Telford L, Seidman SH, Barnes GR. Human Vestibuloocular Reflex and its Interactions with Vision and Fixation Distance During Linear and Angular Head Movement. *J Neurophysiol.* 1998; 80:2391–2404. [PubMed: 9819251]

- Park S, Gianna-Poulin C, Owen-Black F, Wood S, Merfeld DM. Roll Rotation Cues Influence Roll Tilt Perception Assayed Using a Somatosensory Technique. *J Neurophysiol.* 2006; 96:486–491. [PubMed: 16571732]
- Rader AA, Oman CM, Merfeld DM. Motion perception during variable-radius swing motion in darkness. *J. Neurophys.* 2009; 102:2232–2244.
- Singer H, Purcell AT, Austin M. The effects of structure and degree of tilt on the tilted room illusion. *Perception and Psychophysics.* 1970; 7:250–252.
- Stevens SS. On the psychophysical law. *The psychological review.* 1957; 64:153–181.
- Vingerhoets RA, Vrijer MD, Gisbergen JAMV, Medendorp WP. Fusion of visual and vestibular cues in the perception of visual vertical. *J Neurophys.* 2008
- Wertheim AH, Mesland BS, Bles W. Cognitive suppression of tilt sensations during linear horizontal self-motion in the dark. *Perception.* 2001; 30
- White H. Instrumental Variable Regression with Independent Observations. *Econometrics.* 1982; 50
- Wright WG, Dizio P, Lackner JR. Perceived self-motion in two visual contexts: Dissociable mechanisms underlie perception. *J Vestib Res.* 2006; 16:23–28. [PubMed: 16917165]
- Wright WG, DiZio P, Lackner JR. Vertical linear self-motion perception during visual and inertial motion: More than a weighted summation of sensory inputs. *J Vestib Res.* 2005; 15:185–195. [PubMed: 16286700]
- Wright WG, Schneider E, Glasauer S. Compensatory manual motor responses while object wielding during combined linear visual and physical roll tilt stimulation. *Exp Brain Res.* 2009; 192:683–694. [PubMed: 18830585]
- Yasui S, Young LR. Perceived visual motion as effective stimulus to pursuit eye movement system. *Science.* 1975; 190
- Young, LR. *Handbook of Physiology The Nervous System Sensory Processes.* Bethesda MD: Am. Physiol. Soc.; 1984. Perception of the body in space: mechanisms; p. 1023-1066.
- Zupan L, Merfeld DM. Neural processing of gravito-inertial cues in humans. IV. Influence of visual rotational cues during roll optokinetic stimuli. *J Neurophys.* 2003; 89:390–400.

**Fig. 1.**

(A) Graphical representation of one example combination of actual motion and visual scene combination used in the experiment. All actual motions and visual scenes represent swing motions at a chosen radius (which can agree or differ). (B–E) Alternative perceptions that could result from the stimulus: (B) Perception consistent with swing motion at the actual radius (ignoring the visual cues) (C): Perception consistent with swing motion at the radius shown by the visual scene (ignoring vestibular cues). (D): Perception consistent with swing motion derived from a weighted combination of visual and vestibular cues (our hypothesis). (E): Perception of general motion inconsistent with swing motion.

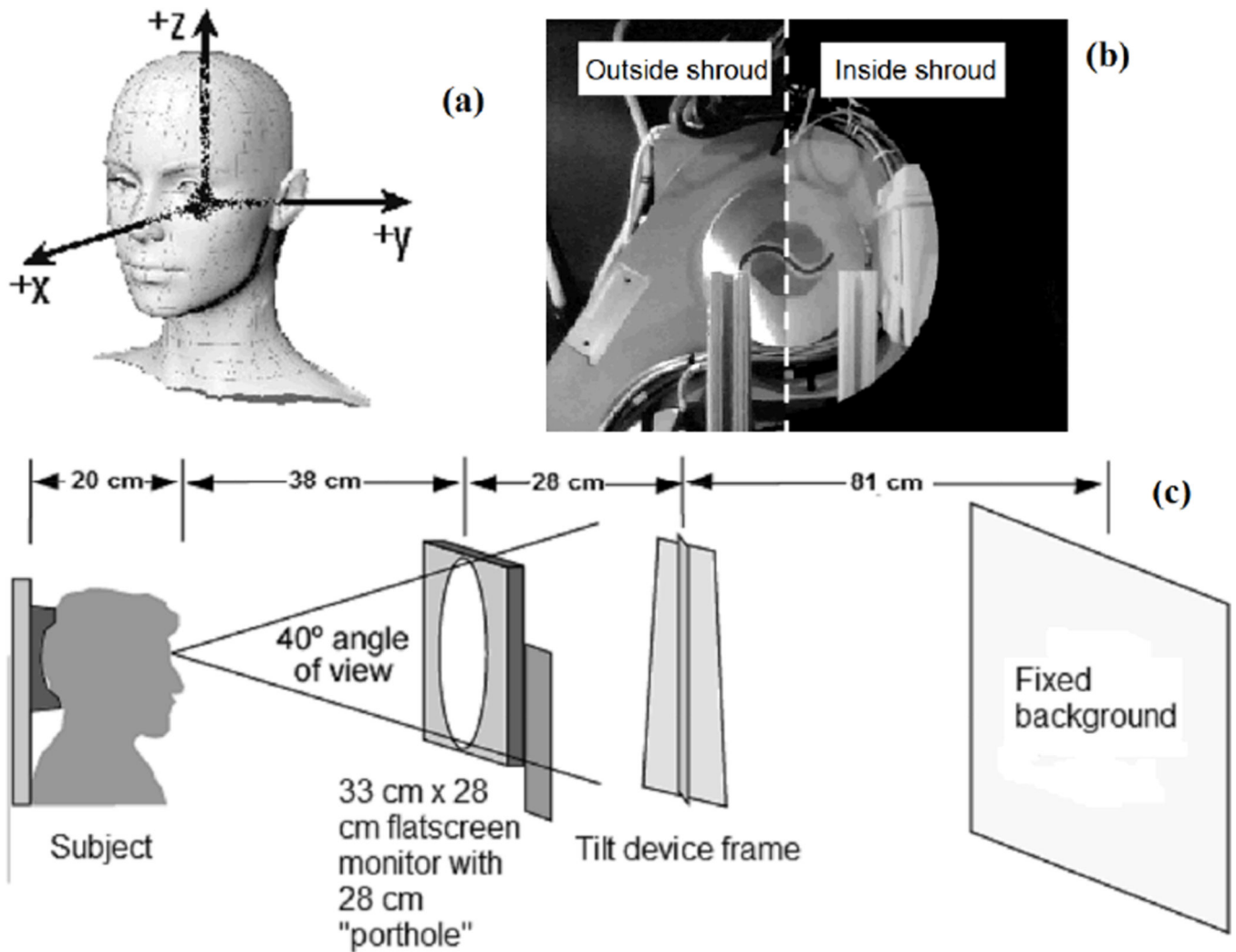


Fig. 2.
 (a) Coordinate frame; (b) Split screen of actual view from subject's position at head center (left) in a lighted room without the flatscreen monitor and shroud, and (right) inside shroud looking at the projected scene; (c) Experimental dimensions.

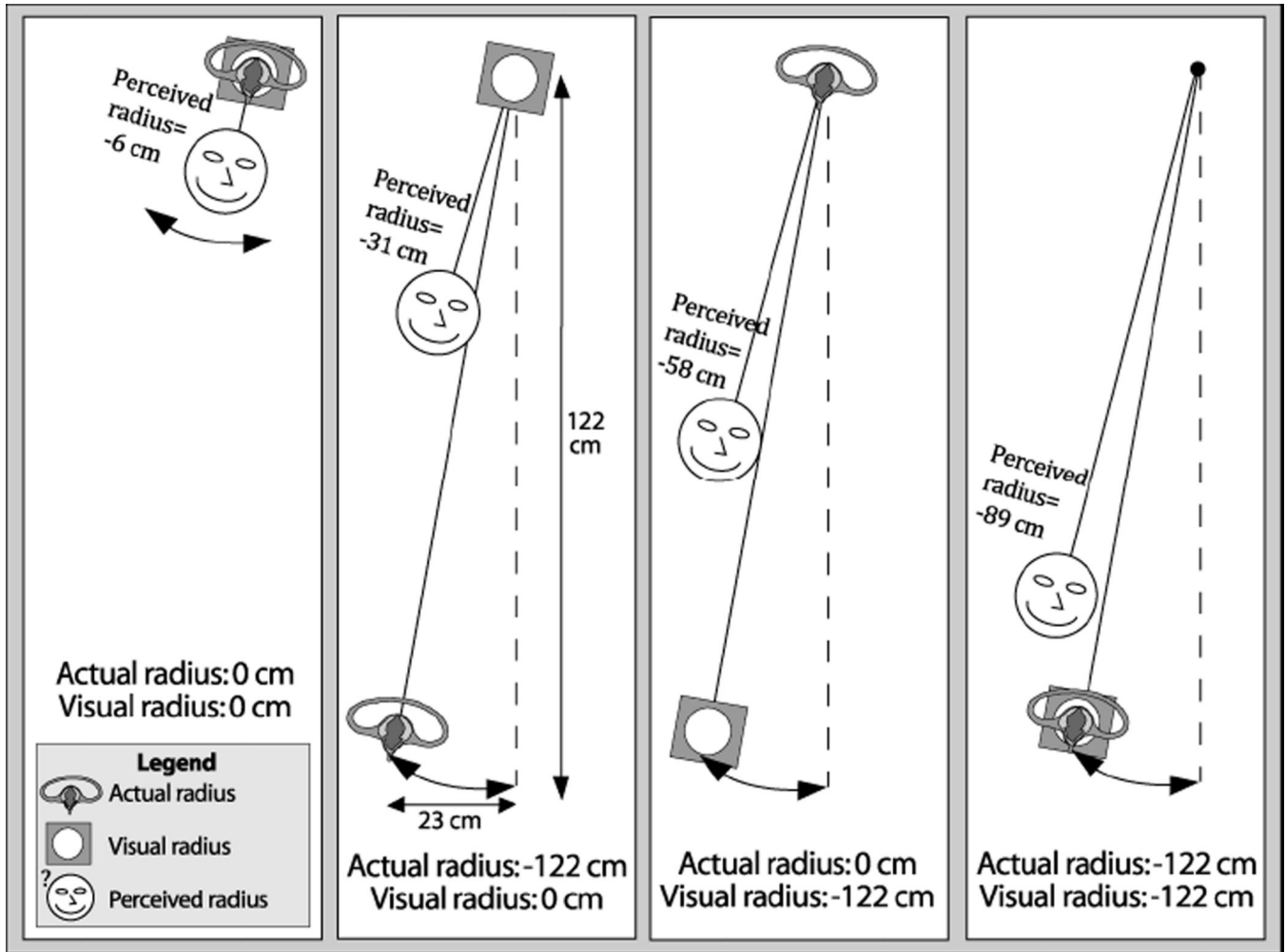


Fig. 3.

Graphical representation of tilt angle and radius perception results: (Left): With the visual and actual radius at 0 cm, the subjects thought they were at 0 cm. (Center-left): With visual radius at 0 cm and the actual radius at -122 cm, the subjects thought they were somewhere between, but closer to the visual radius. (Center-right): With visual radius at -122 cm and the actual radius at 0 cm, the subject thought they were somewhere between, but again closer to the visual radius. (Right): With the visual and actual radius at -122 cm, the subject thought they were close to -122 cm. Note that the tilt angle is overestimated in all cases, yet geometric consistency is preserved.

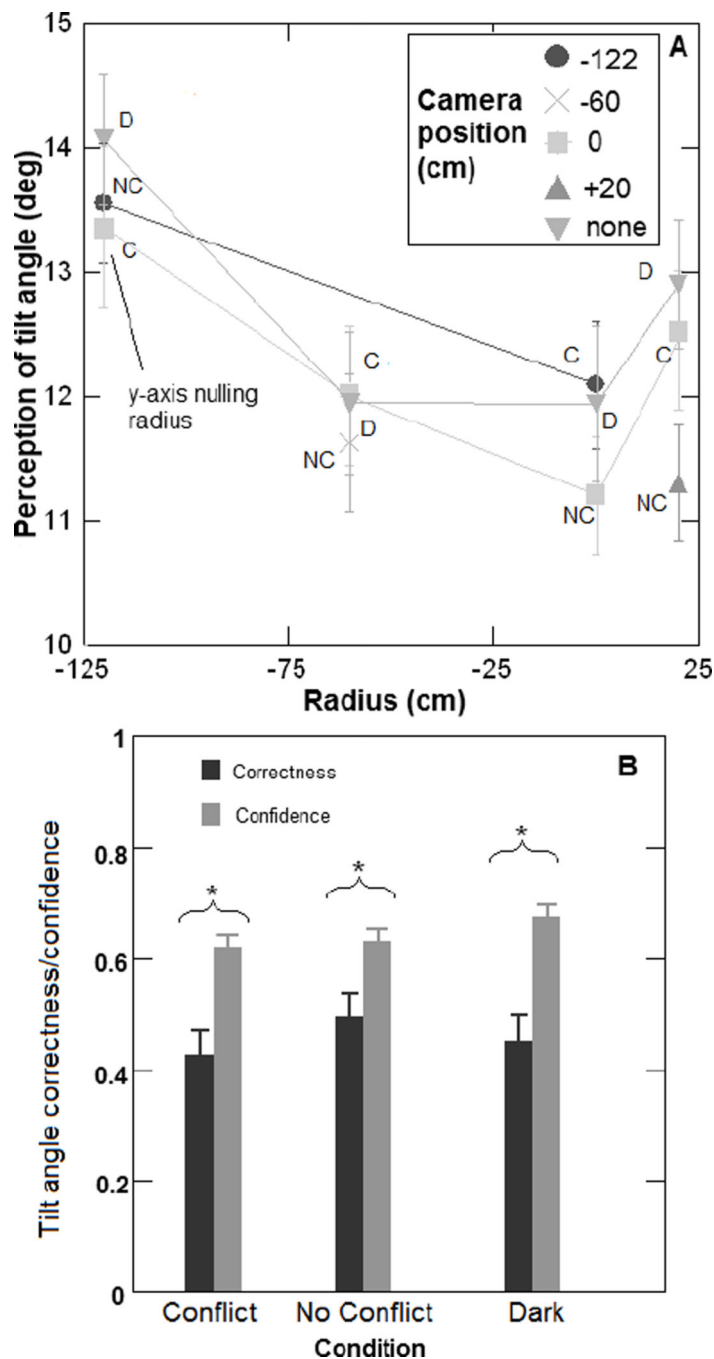


Fig. 4. (A) Subject reported mean tilt perception vs. radius grouped by view condition. Actual tilt angle was 10 degrees in all cases (B) Correctness (fraction of times subjects were within ± 2 degrees) and confidence for “conflict - C” (viewing and actual radius different), “no conflict - NC” (viewing and actual radius the same), and “dark - D” cases. Significant results ($p < 0.05$) are marked *. Correctness and confidence would be equal if subjects were able to perfectly judge their ability to perform the estimation task. Error bars represent standard error.

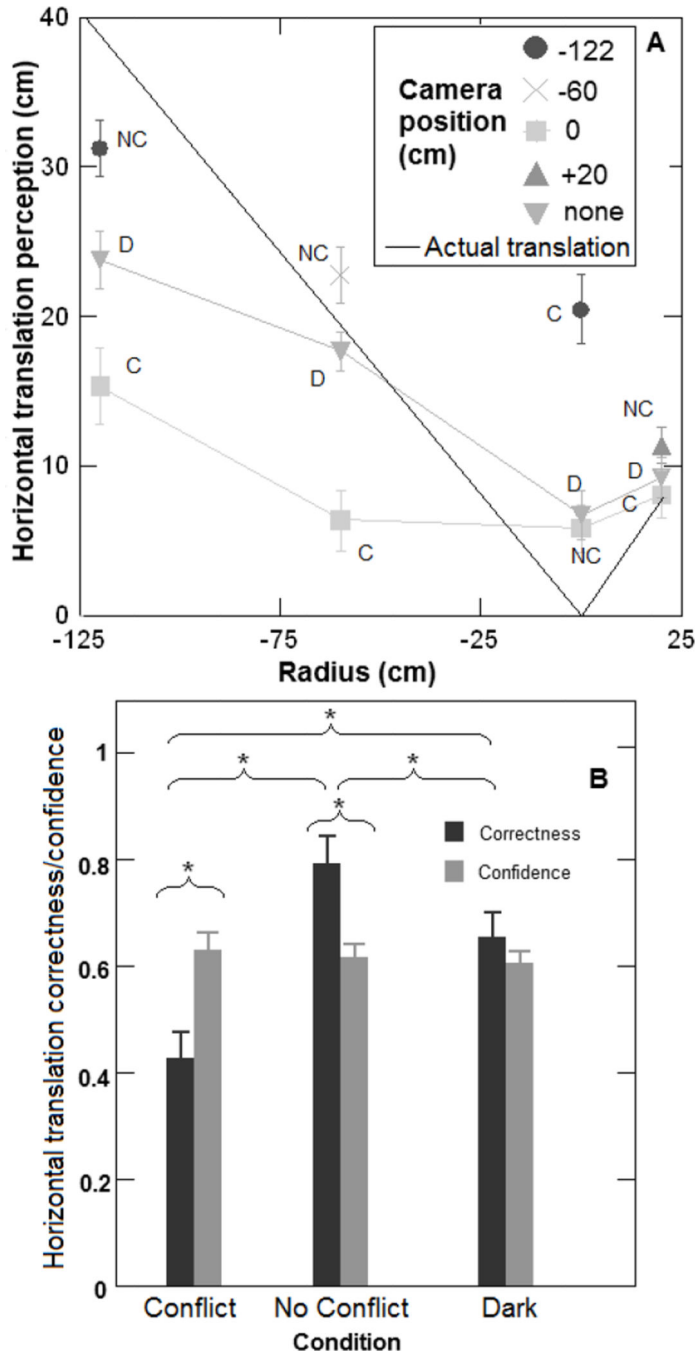


Fig. 5. (A) Subject reported mean horizontal translation perception vs. radius grouped by camera position. (B) Correctness (fraction of times subjects were within +/- 5 inches) and confidence for “conflict - C” (viewing and actual radius different), “no conflict - NC” (viewing and actual radius the same), and “dark - D” cases. Significant results ($p < 0.05$) are marked *. Correctness and confidence would be equal if subjects were able to perfectly judge their ability to perform the estimation task. Error bars represent standard error.

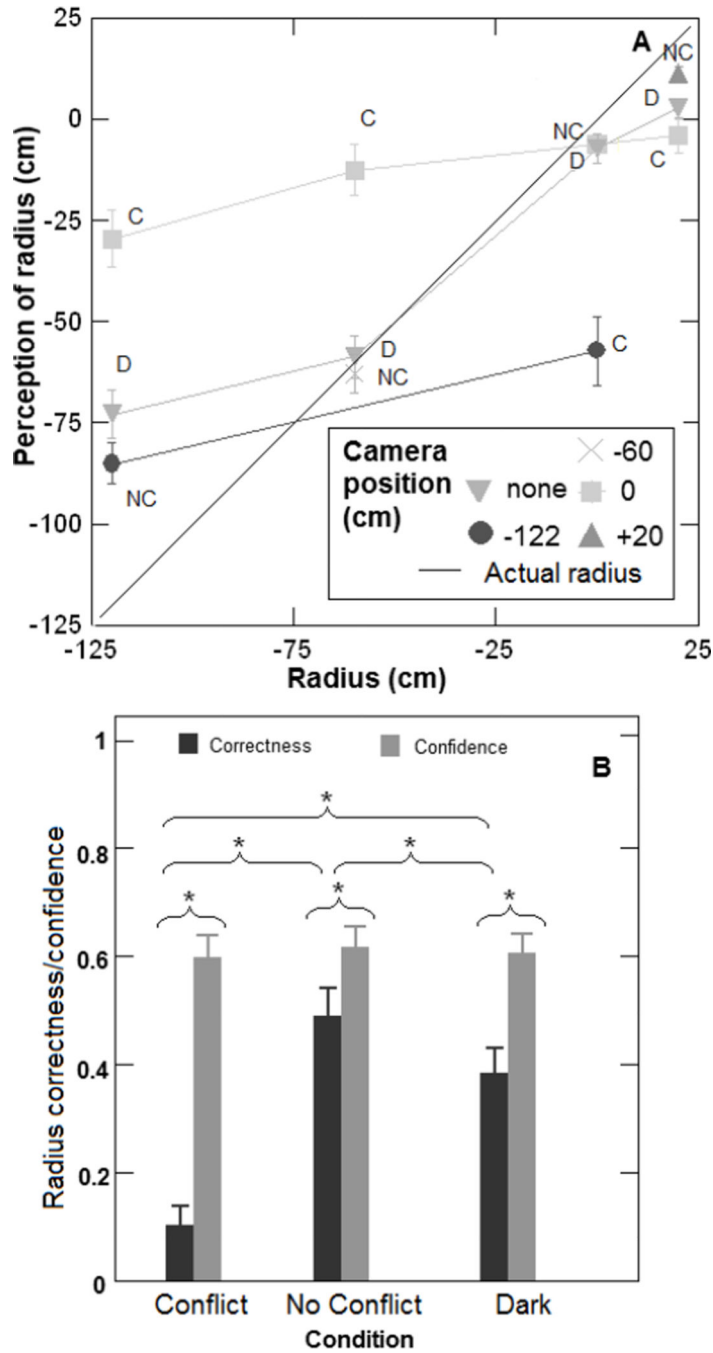


Fig. 6. (A) Subject reported mean radius perception vs. radius grouped by camera position. (B) Correctness (fraction of times subjects were within +/- 5 inches) and confidence for “conflict - C” (viewing and actual radius different), “no conflict - NC” (viewing and actual radius the same), and “dark - D” cases. Significant results ($p < 0.05$) are marked *. Correctness and confidence would be equal if subjects were able to perfectly judge their ability to perform the estimation task. Error bars represent standard error.

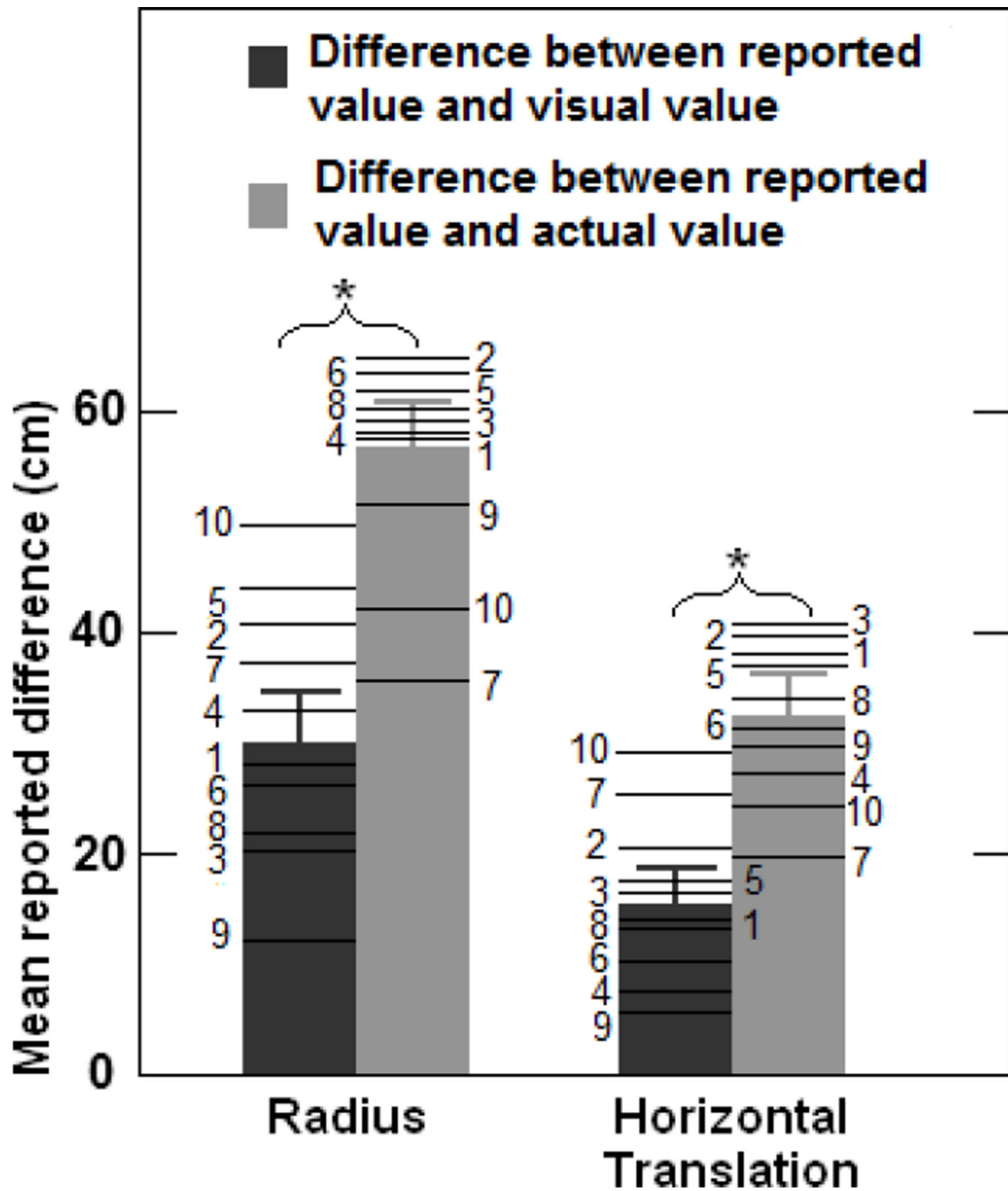


Fig. 7. (A) Overall mean difference from radius and horizontal translation consistent with viewing and actual radius for “conflict” cases. Significant results ($p < 0.05$) are marked *. Error bars represent standard error. Individual subject data are also plotted as a horizontal line with a subject identifier (1–10). Note that only the reports of subjects 7 and 10 were closer to the actual values than the visual values.

Table 1

Training (T) (top 12) and experimental (bottom 12) profiles.

Order	Radius (cm)	Camera (cm)	Conflict?	Tilt amplitude (deg)	Y-axis null	Light?
T1	-122	N	N	30	N	Y
T2*	-122	N	N	20	N	Y
T3	-122	N	N	10	Y	Y
T4*	0	N	N	15	N	Y
T5	0	N	N	10	N	Y
T6*	20	N	N	5	N	Y
T7*	-20	N	N	10	N	N
T8	-60	N	N	10	N	N
T9*	-40	N	N	15	N	N
T10	-122	N	N	20	N	N
T11*	0	N	N	15	N	N
T12	-80	N	N	5	N	N
1/12/7	20	0	Y	10	N	N
2/11/8	-60	N	N	10	Y	N
3/10/9	0	-122	Y	10	N	N
4/9/10	-122	N	N	10	Y	N
5/8/11	-60	0	Y	10	N	N
6/7/12	-122	-122	N	10	Y	N
7/6/1	20	N	N	10	N	N
8/5/2	-60	-60	N	10	Y	N
9/4/3	-122	0	Y	10	Y	N
10/3/4	0	N	N	10	N	N
11/2/5	20	20	N	10	N	N
12/1/6	0	0	N	10	Y	N

* The training sessions marked were omitted for the second and third session. Trial order for each session is marked on the center (for example, 1/12/7 means that the trial was performed first in the first session, last in the second session, and seventh in the third session). Y = Yes; N = No.

Table 2

a Mixed regression effects models for perceived tilt angle.			
Coefficient	Estimate \pm SE	z	p-value
Intercept* (A + Si)	13.2 \pm 0.87°	15.3	<0.01
Actual radius (β_1)	0.025 \pm 0.06°/cm	0.58	0.56
Viewing radius (β_2)	-0.009 \pm 0.025°/cm	-0.26	0.79
Y-axis nulling* (β_3)	2.05 \pm 0.53°	3.9	<0.01

b Mixed regression effects models for perceived horizontal translation.			
Coefficient	Estimate \pm SE	z	p-value
Intercept (A + Si)	5.28 \pm 0.78 cm	6.7	<0.01
Actual radius (β_1)	0.024 \pm 0.02 cm/cm	1.1	0.26
Viewing radius* (β_2)	0.12 \pm 0.012 cm/cm	9.96	<0.01
Y-axis nulling* (β_3)	-3.76 \pm 1.2 cm	-2.78	<0.01

c Mixed regression effects models for perceived radius.			
Coefficient	Estimate \pm SE	z	p-value
Intercept (A + Si)	-0.096 \pm 1.6 cm	-0.059	0.953
Actual radius* (β_1)	0.323 \pm 0.05 cm/cm	-1.410	<0.01
Viewing radius* (β_2)	0.476 \pm 0.03 cm/cm	6.229	<0.01
Y-axis nulling (β_3)	-4.32 \pm 3.11 cm	-1.41	0.159

Significant effects ($p < 0.05$) are marked *. SE = Standard error of estimate.

Significant effects are marked *. SE = Standard error of estimate.