Vestibular Perception and Action Employ Qualitatively Different Mechanisms. I. Frequency Response of VOR and Perceptual Responses During Translation and Tilt

Daniel M. Merfeld,1,2 Sukyung Park,1,2 Claire Gianna-Poulin,3 F. Owen Black,3 and Scott Wood3

1Jenks Vestibular Physiology Lab, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts; 2Department of Otology and Laryngology, Harvard Medical School, Boston, Massachusetts; and 3Neurotology Research, Legacy Clinical Research and Technology Center, Portland, Oregon

Submitted 31 August 2004; accepted in final form 13 February 2005

Merfeld, Daniel M., Sukyung Park, Claire Gianna-Poulin, F. Owen Black, and Scott Wood. Vestibular perception and action employ qualitatively different mechanisms. I. Frequency response of VOR and perceptual responses during Translation and Tilt. J Neurophysiol 94: 186–198, 2005. First published February 23, 2005; doi:10.1152/jn.00904.2004. To investigate the neural mechanisms that humans use to process the ambiguous force measured by the otolith organs, we measured vestibuloocular reflexes (VORs) and perceptions of tilt and translation. One primary goal was to determine if the same, or different, mechanisms contribute to vestibular perception and action. We used motion paradigms that provided identical sinusoidal inter-aural otolith cues across a broad frequency range. We accomplished this by sinusoidally tilting (20°, 0.005–0.7 Hz) subjects in roll about an earth-horizontal, head-centered, rotation axis (“Tilt”) or sinusoidally accelerating (3.3 m/s2, 0.005–0.7 Hz) subjects along their inter-aural axis (“Translation”). While identical inter-aural otolith cues were provided by these motion paradigms, the canal cues were substantially different because roll rotations were present during Tilt but not during Translation. We found that perception was dependent on canal cues because the reported perceptions of both roll tilt and inter-aural translation were substantially different during Translation and Tilt. These findings match internal model predictions that rotational cues from the canals influence the neural processing of otolith cues. We also found horizontal translational VORs at frequencies >0.2 Hz during both Translation and Tilt. These responses were dependent on otolith cues and match simple filtering predictions that translational VORs include contributions via simple high-pass filtering of otolith cues. More generally, these findings demonstrate that internal models govern human vestibular “perception” across a broad range of frequencies and that simple high-pass filters contribute to human horizontal translational VORs (“action”) at frequencies above ~0.2 Hz.

INTRODUCTION

The otolith organs are the primary sensory system for transducing both orientation of the head with respect to gravity and head linear acceleration. The otolith receptors contribute to perceptions of both tilt and translation (“perception”) as well as reflexive responses (“action”) like the VOR (e.g., Young 1984). However, as stated by Einstein’s Equivalence Principle, no accelerometer/graviceptor can distinguish gravity (g) from linear acceleration (a); all accelerometers and all graviceptors measure specific gravito-inertial force, which is the vector sum of gravity minus linear acceleration (f = g − a). Therefore additional neural processing is required to elicit behaviorally relevant tilt and translation responses. Two different explanations for how the nervous system might accomplish the necessary processing have been proposed: simple filtering (Mayne 1974; Paige and Tomko 1991) and use of other available information via internal models (e.g., Angelaki et al. 1999; Droulez and Cornilleau-Peres 1993; Droulez and Darlot 1989; Glasauer 1992; Glasauer and Merfeld 1997; Green and Angelaki 2003; Mayne 1974; Merfeld 1995a,b; Merfeld and Zupan 2002; Merfeld et al. 1999a; Young 1984; Zupan et al. 2002).

Internal models describe an explicit set of neuronal mechanisms by which information from disparate sensory systems (e.g., semicircular canals, otolith organs, retina) and nonsensory systems (e.g., efferent copy, cognitive contributions) can be combined and fused by the nervous system. The end products provided by these neural processes of sensory fusion are central estimates of spatial orientation (e.g., central estimates of gravity, central estimates of linear acceleration, central estimates of angular velocity). Even though fusion mechanisms other than internal models may exist, we will use the more specific term, internal models, throughout most of the paper because internal models describe an explicit set of testable hypotheses and because some earlier studies have demonstrated the contributions of internal models to sensory fusion (e.g., Angelaki et al. 1999; Merfeld and Young 1995; Merfeld et al. 1999). There is no question that sensory fusion contributes to the neural processing of perceived tilt. First consider visual influ-

1 A review of the relevant physics as well as a review of experimental data consistent with the internal model hypothesis has been published (Merfeld and Zupan 2003).

2 To provide direct testable hypotheses, we define internal models as neural systems that mimic explicit physical or physiological processes (e.g., physical relationships, sensory dynamics, or motor dynamics) “known” to the nervous system. For example, when a physical process can be described by a mathematical operation (e.g., f = g − a), an internal model of this physical process would signify that a neural process equivalent to this mathematical operation occurs (f = g − a) within the neurons that calculate and encode the neural representations (f, g, or a) of the physical variables (f, g, or a). The model used herein (Merfeld and Zupan 2002) to predict physiological responses provides an implementation of the internal model hypothesis by utilizing internal models as the predominant mechanism for neural processing and sensory fusion.

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
ences; human perception of tilt has long been known to be influenced by visual orientation cues (Asch and Witkin 1948a,b) and visual motion cues (Dichgans et al. 1972, 1974; Howard and Childerson 1994; Zupan and Merfeld 2003) in combination with the otolithic measures of gravity.

It has been hypothesized that internal models explain these well-known visual influences on tilt perception (Young 1984). We have hypothesized that the nervous system also has an internal model of the physical law that specific gravito-inertial force is the sum of gravity minus linear acceleration (e.g., Merfeld 1990, 1995a,b; Merfeld and Zupan 2002; Merfeld et al. 1993a). Putting these two internal models together, we predicted that visual roll rotation cues also contribute to the horizontal translational VOR via internal models (Zupan and Merfeld 2003), which has been confirmed experimentally (Zupan and Merfeld 2003). Visual roll rotation cues may similarly influence translation perception via internal models (Merfeld and Zupan 2003).

The dynamics of human tilt perception as a function of frequency during sinusoidal translation in the dark suggested that human tilt perception is governed by simple low-pass filtering of otolith signals (Seidman and Paige 1996; Seidman et al. 1998). However, published internal model predictions for similar translation paradigms (Merfeld and Zupan 2002) predicted tilt perception dynamics that resemble the output of simple low-pass filters even without the explicit use of low-pass filters. This modeling demonstrates that one cannot determine whether simple filtering contributes to a response by simply measuring the frequency response during Translation. Furthermore, unlike the output of a simple low-pass filter, the phase of perceived tilt has been shown to be relatively constant across a broad frequency range (Glasauer 1995). Also inconsistent with the hypothesis that tilt perception is elicited by simple low-pass filtering are earlier findings that show that roll cues from the semicircular canals contribute to roll tilt perception (Stockwell and Guedry 1970; von Holst and Grisebach 1951). Therefore, our first goal was to determine definitively whether or not simple filtering and/or internal models contribute to perceptions of tilt and/or translation in humans. To achieve this goal, we measured perceptions of both tilt and translation while using both a roll tilt motion paradigm and an inter-aural translation motion paradigm. We performed these motion paradigms using sinusoids across a broad range of frequencies (0.005–0.7 Hz) because sinusoids provide the most direct way to test the predictions of the simple filtering hypothesis.

The second goal was to investigate whether human translational VOR responses are influenced by canal cues or are elicited by simple filtering. We investigated this goal by testing across a broader range of frequencies than earlier studies. Specifically, a number of studies using smaller frequency ranges have demonstrated the influence of canal cues on translation VORs (Angelaki et al. 1999, 2001; Green and Angelaki 2003; Merfeld and Young 1995; Merfeld et al. 1999, 2001; Peterka et al. 2004; Zupan et al. 2000). In one of these studies (Merfeld and Young 1995), squirrel monkeys were quickly tilted in roll such that a high-frequency inter-aural otolith cue was provided by gravity; little or no horizontal VOR was observed. This result was inconsistent with the contributions of simple filtering. This finding was confirmed and substantially extended by a study published by Angelaki et al. (1999) in which rhesus monkeys were roll tilted, translated along their inter-aural axis, or simultaneously roll tilted and translated. The inter-aural shear force during roll tilt alone and translation alone were nearly identical. Only a small horizontal VOR was observed during the roll tilt stimulation, while a substantial horizontal VOR was observed whenever an inter-aural translation was present, showing that the animals used the rotational cue provided via the canals to help resolve the ambiguous gravito-inertial force (GIF) cue. Another set of studies complemented these findings by showing that translational VORs could be elicited via internal models even in the absence of actual translation (Merfeld et al. 1999; Zupan et al. 2000). These findings clearly demonstrate the influence of canal cues on the translation VOR response, but do not rule out the possibility that simple filtering (i.e., simple high-pass filtering) might also contribute to the translational VOR.

The third, and most important, goal of this investigation was to measure perception (perceived roll tilt and perceived inter-aural translation) and action (VOR) simultaneously to determine if the same (or similar) neural mechanisms contribute to perceptual and reflexive responses in humans. It has long been known that there are quantitative differences between eye movement and perceptual responses in humans. For example, the amount of tilt perceived during earth-horizontal inter-aural translation (Glasauer 1995) is not matched by the amount of ocular torsion (Lichtenberg et al. 1982; Merfeld et al. 1996b) for similar stimuli nor is the phase of these two tilt responses the same. A limited number of earlier studies investigated tilt/translation resolution by simultaneously measuring perceptual responses and the VOR (Merfeld et al. 2001; Zupan and Merfeld 2003). Findings from these studies suggest that the translational VOR may share some common mechanisms with perception.

The three goals described in the preceding text led us to design and perform a comprehensive study in which we measured the VOR and perceptions of tilt and translation during sinusoidal translations and sinusoidal tilts across a broad range of frequencies (0.005–0.7 Hz).

**METHODS**

Eye movements and perceptions of tilt and translation were measured in the dark during two different motion paradigms (Fig. 1).
labeled \textit{Translation} and \textit{Tilt}, which are described in detail in the following text. In brief, these motion paradigms provided nearly identical inter-aural (y axis) otolith cues while providing dramatically different roll rotation canal cues. Specifically, there was a substantial roll rotation cue from the canals during the \textit{Tilt} paradigm but not during the \textit{Translation} paradigm.

The subject’s chair and restraints were nearly identical for the two paradigms. Head restraint was provided by a modified motorcycle helmet, which was split down the center to allow custom tightening. Bodies were restrained with a five-point harness and lateral shoulder support. Vacuum-formed “beanbags” and foam pads were added as needed to ensure maximum stability and relative comfort. Knees were secured in place with a strap; feet were secured to a footrest with Velcro straps. The stimuli were always provided in a different randomized order for each subject.

\textit{Tilt paradigm}

Sinusoidal head-centered roll tilts (“\textit{Tilt}” paradigm, Fig. 1A) were accomplished by placing the ear level of the subject near the rotation axis of an earth-horizontal rotator (e.g., Angelaki et al. 1999; Merfeld and Young 1995; Stockwell and Guedry 1970). More specifically, the rotation axis was located such that the linear accelerations (both tangential and centripetal accelerations) at the otolith organs fell below the human acceleration threshold of 0.05 m/s² (Benson et al. 1986; Melvill Jones and Young 1978). For the highest frequency stimuli (0.7 Hz), where the tangential acceleration was largest, this meant that ear placement had to be accurate within \(\pm 1\) cm. This placement accuracy was verified using two redundant measurements of the location of the external auditory meatus with respect to two distinct landmarks previously located with respect to the rotational axis. (Head placement prohibited direct measurement of ear location with respect to rotation axis.)

To avoid discontinuous angular acceleration at motion onset, which becomes substantial at higher frequencies, the angular velocity was linearly increased over an integer number of sinusoidal cycles for all frequencies \(>0.02\) Hz to yield the steady-state sinusoidal stimuli. (While an example of the linear increase for \textit{Tilt} is not shown, Fig. 2, \textit{B} and \textit{C}, shows a similar linear increase for the \textit{Translation} paradigm.) The steady-state sinusoidal roll tilt (0.005 to 0.7 Hz, Table 1) was left/right symmetric about the upright position and had a peak amplitude (\(\Theta_{\text{max}}\)) of 20°. For this paradigm, the steady-state inter-aural (y axis) gravitational force measured by the otolith organs was 
\[ g_\text{i}(t) = -G \sin(\Theta_{\text{max}} \sin(2\pi t)) = -G \sin(\Theta_{\text{max}} \sin(2\pi y/t)), \]

where \(G\) is gravitational force. The inter-aural gravity component had a peak amplitude of 3.3 m/s².

\textit{Translation} paradigm

For the \textit{Translation} paradigms (Fig. 1, \textit{B} and \textit{C}), the subject’s chair was mounted on a platform that could translate horizontally along the rotational arm of a centrifuge (Merfeld et al. 2001). For the highest stimulus frequency (0.7 Hz), the upright subject was simply sinusoidally translated from side to side (Fig. 1B) such that the peak inter-aural linear acceleration was 3.3 m/s². For this condition, the translational linear acceleration, \(a_x = \ddot{x} = -D(2\pi f)^2 \sin(2\pi y/t)\), was the only inter-aural otolith stimulus (\(D\) is peak displacement).

Due to limited track length, for all other frequencies of \textit{Translation} (0.005 to 0.5 Hz), the subject was first seated at the center of rotation of the centrifuge and rotated in yaw about an earth-vertical axis for \(\approx 5\) min, allowing the yaw canal cue to decay, before proceeding with sinusoidal radial translations (Fig. 1C). This is the same technique used in previous studies (e.g., Glasauer 1995; Merfeld et al. 2001; Paige and Seidman 1999; Seidman et al. 1998) to yield low-frequency linear acceleration in the absence of a very long track. The centrifuge yaw angular velocity (\(\omega\)) was trapezoidal, with an angular acceleration of \(25\pi^2\)°/s² to a constant velocity of \(250\pi\)°/s, which was maintained for \(\approx 30–45\) min. During the sinusoidal radial translation, centripetal acceleration (\(\ddot{r} \propto \omega \times \omega\)), radial acceleration (\(\ddot{r}\)), and Coriolis acceleration (\(2\omega \times \dot{r}\)) were present (Fig. 2, Table 1), where \(\dot{r}\) is the displacement vector from the center of rotation and \(\times\) represents the standard vector cross product operator. Centripetal acceleration and radial acceleration were both sinusoidal and naturally synchronized because they were both phase-locked to radial displacement and were both always aligned with the inter-aural axis. We chose the amplitude of the sinusoidal radial displacement (Table 1), such that the centripetal acceleration and radial acceleration summed to yield a peak inter-aural otolith cue of 3.3 m/s², matching the inter-aural gravitational force during a 20° tilt. As for the roll tilt sinusoids, acceleration discontinuities were eliminated by linearly increasing the sinusoidal radial velocity (Fig. 2) over an integer number of cycles for all frequencies \(>0.02\) Hz (Table 1).

The sinusoidal inter-aural otolith cue provided by the precise combination of centripetal and radial accelerations mimics the sinusoidal inter-aural otolith cue present during a sinusoidal roll tilt. Therefore only the Coriolis acceleration is a potential contaminant. As such, it is important to note several facts. First, there was no Coriolis acceleration at the highest frequency (0.7 Hz). Second, the Coriolis acceleration was always smaller than the inter-aural acceleration (Table 1). Third, the Coriolis acceleration was at or below the human acceleration detection threshold (Benson et al. 1986; Melvill Jones and Young 1978) for the two lowest frequencies. Fourth, the Coriolis acceleration was more than an order of magnitude smaller than the

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|}
\hline
\textbf{Frequency, Hz} & \textbf{No. of SS Cycles} & \textbf{No. of Ramp-Up Cycles} & \textbf{Peak angular velocity, °/s} & \textbf{Peak radial displacement, m} & \textbf{Peak centripetal acceleration, m/s²} & \textbf{Peak radial acceleration, m/s²} & \textbf{Peak inter-aural acceleration, m/s²} & \textbf{Peak coriolis acceleration, m/s²} \\
\hline
0.005 & 2 & 0 & 0.6 & 0.175 & 3.3 & 0.0 & 3.3 & 0.05 \\
0.01 & 3 & 0 & 1.3 & 0.175 & 3.3 & 0.0 & 3.3 & 0.10 \\
0.02 & 5 & 0 & 2.5 & 0.175 & 3.3 & 0.0 & 3.3 & 0.19 \\
0.05 & 7 & 1 & 6.3 & 0.174 & 3.3 & 0.0 & 3.3 & 0.48 \\
0.1 & 7 & 1 & 12.6 & 0.172 & 3.3 & 0.1 & 3.3 & 0.94 \\
0.2 & 10 & 1 & 25.1 & 0.162 & 3.1 & 0.3 & 3.3 & 1.77 \\
0.5 & 10 & 2 & 62.8 & 0.115 & 2.2 & 1.1 & 3.3 & 3.16 \\
0.7 & 10 & 3 & 88.0 & 0.173 & 0.0 & 3.3 & 3.3 & 0.00 \\
\hline
\end{tabular}
\caption{Characteristics of the motion stimuli for both \textit{Tilt} and \textit{Translation} paradigms}
\end{table}

The number of steady-state (SS) and ramp-up cycles are shown for each stimulus frequency. For the \textit{Tilt} paradigm, the peak roll tilt amplitude was always 20°, yielding an inter-aural gravitational stimulus of \(-3.3\) m/s² (0.34 g) at each frequency. The peak amplitude of the roll angular velocity during \textit{Tilt} is shown for each stimulus frequency. The peak radial displacement, which was the controlled motion parameter for the \textit{Translation} paradigm, was chosen to yield a peak inter-aural acceleration of 3.3 m/s² at all stimulus frequencies, matching the inter-aural otolith cue during the \textit{Tilt} paradigm. Also shown are the peak radial acceleration and peak centripetal acceleration, which sum to yield the fore-mentioned total peak inter-aural acceleration. For comparison, the theoretical, peak, naso-occipital acceleration, the Coriolis acceleration, is also shown.
Accleration present.

without any rotation (and hence, direction of Coriolis force) on the measured responses, so for this condition, radial acceleration is the only acceleration present.

inter-aural acceleration for half the frequencies. Furthermore, the Coriolis acceleration was always orthogonal to the inter-aural cue in two ways. It was always physically orthogonal because it was always aligned with the naso-occipital axis, and time, because it is 90° out of phase with the inter-aural acceleration. Left: acceleration components present at 0.01 Hz; acceleration is almost exclusively centripetal. Middle: acceleration components for 0.2 Hz. The highest frequency (0.7 Hz) Translation condition was achieved without any rotation (right), so for this condition, radial acceleration is the only acceleration present.

Analog data acquisition

Analog data were acquired at a rate of 60 Hz via a 12-bit analog to digital converter. For the Tilt paradigm, the roll tilt angle was measured via a potentiometer placed on the tilt rotation shaft. For the Translation paradigm, the radial position of the subject’s chair with respect to the centrifuge arm was measured with a potentiometer placed on the linear drive motion shaft.

Subjects and instructions

Eight healthy subjects were prescreened as “normal” via several, standard, clinical, vestibular tests. All subjects signed an informed consent, consistent with institutional procedures prior to participation. The subjects (7 males and 1 female) were between 22 and 60 yr of age. Only data from subjects who successfully completed all trials for a given testing paradigm are included. All eight subjects successfully completed the Tilt paradigm. Seven of eight subjects successfully completed the Translation paradigm; one subject experienced nausea and other symptoms of motion sickness.

Throughout the trials, always performed in darkness, subjects were instructed to open their eyes and look straight ahead without focusing on any point, real or imagined. They were also instructed to remember their perceptions of roll-tilt and inter-aural linear translation so that they could provide verbal reports of their perceived motion after each trial. We used proximal vergence (Jaschinski-Kruza 1990; Schor et al. 1992; Wick and Bedell 1989) to help stabilize and control gaze distance because it is known that human vergence is generally maintained, even in darkness, to be roughly consistent with the distance to surroundings previously viewed in the light. Specifically, we used a dim light to illuminate a poster 1 m in front of the subject for ~15 s before trials. Consistent with the published proximal vergence literature, our measurements showed that vergence was maintained relatively constant for each subject for each test condition.

Eye-movement acquisition and analysis. Binocular eye movements were video-taped with a frame-rate of 60 Hz using an infrared video-oculography (VOG) system (SensoMotoric Instruments). Video cameras were attached to the subject’s head via a facemask and a strap around the back of the head. We further stabilized the goggles by adding a bite bar assembly, providing a stable attachment of the camera to the maxilla (upper jaw), using a custom mold of each subject’s upper jaw made with dental impression compound (3M Express STD). Infrared light-emitting diodes (LEDs), which were invisible to the subject, provided lighting for the cameras. The weight of the goggle assembly was partially supported by a pair of bungee cords attached to the structure moving with the subject. Eye position was calibrated by having subjects sequentially direct their gaze at 13 targets (1 center, 6 horizontal, and 6 vertical). Targets were separated by roughly 5°, with actual angular displacement dependent on exact eye position in space, which was carefully measured prior to each test session, providing a thorough calibration over approximately ±15° in horizontal and vertical dimensions.

Three-dimensional (3D) geometric projection corrections were used to calculate 3D eye position in Fick coordinates (Haslwanter and Moore 1995; Moore et al. 1996). The Fick angles were digitally filtered and differentiated to yield Euler rates, which represent the angular velocity of the eye. These Euler rates were then processed to yield 3D angular velocity of the eye in standard, orthogonal, head-fixed coordinates (Merfeld and Young 1992). Torsional, vertical, and horizontal eye velocities refer to the eye movements about rotation axes aligned with the head x axis (naso-occipital axis), y axis (inter-aural axis), and z axis, respectively. Slow phase velocity (SPV) was calculated using a computer algorithm based on peak acceleration detection to remove the fast phases, with some manual editing by experienced personnel to verify that the automatic algorithm worked properly.

Eye data from five of the seven subjects who successfully completed testing are reported herein. Video data for two of the seven subjects are not reported. Eye images from one of these two subjects showed a small motion artifact on a few trials, indicating that the subject was not properly biting down on the bite bar used to secure the video cameras with respect to the head. The second subject showed continuous variations in pupil diameter (i.e., “hippus”), which made analysis both noisy and inaccurate.

Verbal reports of tilt and translation

The subjects were trained to answer, at the end of each trial, three specific questions focused on head tilt and bridge-of-the-nose trans-
The three specific questions were: what was the maximal, side-to-side, bridge-of-the-nose translation that you experienced? What was the maximal head tilt to the left? And what was the maximal head tilt to the right? Using the right-hand-rule, rightward tilts were defined as positive and leftward tilts as negative. Note that peak-to-peak translation was recorded because translation, unlike tilt, lacks an absolute reference. After answering these 3 questions, subjects were asked to describe any other sensation of motion that they experienced. Not once in 96 trials, did any subject report pitch tilt during the roll tilt condition, and, in only 3 of the 84 translation trials did subjects report a noticeable pitch sensation.

Data analysis

Sinusoidal fits were used to calculate amplitudes and phases of the motion variables and eye movement responses. Steady-state sinusoidal data were fit using a least-mean-square linear regression to the equation, \( x(t) = B + A_c \cos(2\pi f t) + A_s \sin(2\pi f t) \), where \( B \) is a DC bias, \( A_c \) is the amplitude of the cosine component, and \( A_s \) is the amplitude of the sine component. This fit was performed on a cycle-by-cycle basis. The cycle-by-cycle sine and cosine components were averaged to calculate the mean values for that trial, which were then averaged across subjects. Amplitude (A) and phase (\( \phi \)) were then calculated, \( A = \sqrt{A_c^2 + A_s^2} \) and \( \phi = \tan^{-1}(A_s/A_c) \). To calculate variability of amplitude and phase, we calculated the two-dimensional covariance ellipses of the sine and cosine components (e.g., Johnson and Wichern 1982). Error bars shown in all VOR figures represent the SE obtained from the covariance ellipse.

Modeling

Two model structures, “simple filtering” and “internal models,” were investigated to explain the measured responses and to examine physiologic explanations that are consistent both with known neural responses and behavioral data. While our modeling treated these model structures as independent sets of hypotheses, this treatment is not meant to preclude the possibility that both mechanisms may contribute simultaneously to any given response. For example, the canals could contribute to the horizontal translational VOR via internal models while simple filters could simultaneously contribute to the VOR. Simple filters were implemented as first-order filters. Since the characteristics of low- and high-pass filters are well known, we did not perform simple filtering simulations before beginning our experiments. The simulated responses shown (Fig. 3, A and B) were obtained by choosing a cut-off frequency that was consistent with the experimental data and performing the filtering simulations only after the data were in hand. The internal model approach was implemented using a previously published model (Merfeld and Zupan 2002). We performed model predictions using the published model prior to experimentation (solid black lines, Fig. 3, C and D). We also simulated responses after making two parameter changes to the model to fit the experimental data better (dotted gray lines). The right axis is included to show the roll tilt (°) yielded by this inter-aural estimate of gravity. Low-pass model predictions for Translation (empty circle) and Tilt (×) motion paradigms are identical. B: high-pass filter (\( f_{hp} = 0.07 \) Hz) predictions of inter-aural linear acceleration (m/s^2) are shown. High-pass model predictions for Translation (empty circle) and Tilt (×) are identical. C: internal model estimates of inter-aural gravity (m/s^2) predicted prior to experimentation (solid black lines) and simulated results after post hoc parameter changes to better fit the experimental data (dotted gray lines). The right axis is included to show the roll tilt (°) yielded by this inter-aural estimate of gravity. D: internal model estimates of inter-aural linear acceleration (m/s^2) predicted prior to experimentation (solid black lines) and simulated results after post hoc parameter changes (dotted gray lines). Estimates of linear acceleration during Translation exceed the estimates of acceleration during Tilt.

Simple filtering predictions

Low-pass filtering of inter-aural otolith cues has been hypothesized to explain roll tilt responses (e.g., ocular torsion, perceived roll tilt) and is consistent with the fact that roll tilt with respect to gravity introduces inter-aural otolith cues. Complementary to this prediction, high-pass filtering of inter-aural otolith cues has been hypothesized to lead to inter-aural translational responses (e.g., horizontal translational VOR, perception of inter-aural translation, etc.), which is consistent with the fact that inter-aural linear acceleration introduces inter-aural otolith cues. To efficiently investigate these hypotheses, we used models to predict the outputs of fixed-parameter filters. Low-pass (Fig. 3A) and high-pass (Fig. 3B) model simulations are shown for both the Translation (empty circle) and Tilt (×) paradigms. Because the output of each of the simple filters depends only on the otolith input and does not depend on canal cues or on any variable other than the inter-aural force measured by the otolith organs, the simple filter predictions are identical for Translation and Tilt.

Internal model predictions

The internal model predictions (Fig. 3, C and D) demonstrate several key points. First, the predicted tilt response

---

Footnote: 3 Head-fixed references allowed us to focus our study on the influence of the vestibular system and not on somatosensory graviceptors (e.g., Mittelstaedt 1996), which have been shown to influence some low-frequency postural matching tasks but have not been shown to influence eye movements substantially. For example, it has been shown that torso proprioception can elicit reflexive eye movements, but these responses are small or negligible at frequencies above ~0.05 Hz (Mergner et al. 1998), which is the frequency range where we measure substantial VOR responses.
during Translation (empty circle, Fig. 3C) looks qualitatively similar to the low-pass filter predictions (empty circle, Fig. 3A), even though this model does not include an explicit low-pass filter yielding this tilt estimate. Similarly, the predicted translation response during Translation (empty circle, Fig. 3D) is qualitatively similar to the high-pass filter predictions (empty circle, Fig. 3B), even though this model does not include an explicit high-pass filter to yield this translation estimate. However, unlike the predicted responses of simple filters, the predicted tilt (×, Fig. 3C) and translation (×, Fig. 3D) responses for this model during the Tilt condition are substantially different from those predicted for the Translation condition (empty circle, Fig. 3, C and D). Specifically, the predicted tilt responses remain relatively constant during Tilt. Furthermore, the predicted translation responses remain relatively small during Tilt compared with the translation responses during Translation, even though the inter-aural otolith cues are identical. In both cases, the difference between the predicted responses to Translation and Tilt is due to the presence of roll canal cues during Tilt and the absence of a change in roll canal cues during Translation.

Eye movements

Figure 4 shows eye movements during the Translation and Tilt paradigms for one subject. A substantial torsional velocity was always present during Tilt with little or no torsional velocity during Translation. This was true for all subjects. In addition, consistent vertical responses were seldom evident. A substantial horizontal response was observed during both Translation and Tilt with the derivative of horizontal eye position having a substantially larger amplitude during Tilt than during Translation (Fig. 4, 4th row). Application of 3D kinematic corrections yielded the horizontal velocity shown in the bottom row, where the amplitude of the horizontal responses during Translation and Tilt were nearly the same. This demonstrates the importance of the full 3D kinematic corrections (Hess et al. 1992; Merfeld and Young 1992; Tweed et al. 1990) to eye movement analysis because reflexive eye responses have been shown to vary with gaze direction (Seidman et al. 1995). Therefore all eye data discussed from this point onward will only include data that are fully corrected using the 3D kinematic corrections (Merfeld and Young 1992) described in detail in METHODS.

The amplitude of the average horizontal eye response increased with frequency for both Translation and Tilt conditions (Fig. 5A). The eye response amplitudes overlapped at frequencies <0.2 Hz, but the horizontal responses during Tilt became slightly larger than the horizontal responses during Translation at higher frequencies. To help characterize the difference between these Translation and Tilt responses, we calculated the magnitude of the vector difference between the horizontal responses as a function of frequency (Fig. 5A),. Note that the primary difference between these horizontal responses emerges only at higher frequencies.

During Translation, the response demonstrated a 90° phase lead at low frequencies and 0° of phase at higher frequencies, where 0° would be compensatory for the actual linear velocity (Fig. 5B). The phase during Tilt showed similar characteristics, at both the lowest and highest frequencies, but diverged at mid frequencies. The amplitude of the horizontal response was quite small at these middle frequencies. It is important to note that phase does not provide a robust parameterization when responses are small because small variations (e.g., noise associated with fast phase removal) correlated with the motion can lead to large phase variations. Therefore the measured phase differences in the horizontal VOR probably include noise and other artifactual (e.g., “cross-talk,” see DISCUSSION) contributions, especially at frequencies <0.2 Hz.

The average torsional eye velocity increased as a function of frequency during Tilt (Fig. 5C), which was expected because the peak angular velocity of the stimulus increased linearly with frequency (Table 1). In contrast, a tiny torsional velocity response was observed during the Translation condition, which was also expected, because there was no roll angular velocity stimulation. This means that there is a substantial difference in the torsional VOR during Translation and Tilt, which helps explain why the horizontal response during Tilt exceeds that during Translation at frequencies >0.1 Hz. (see DISCUSSION for details regarding this cross-talk.)

Tilt perception

Verbal reports of tilt perception were roughly constant at lower frequencies (<0.05 Hz) and decreased at higher frequen-
cies for both the Translation and Tilt trials but with a very substantial decrease for the Translation trials (Fig. 6A). For the Tilt trials, the magnitude of the tilt response was overestimated at lower frequencies and converged to the actual tilt angle of 20° at higher frequencies while it decreased to near-zero for the Translation trials. Because the inter-aural otolith cues were nearly identical during the Translation and Tilt trials, this difference between the perceptual tilt responses during roll Tilt and during inter-aural Translation demonstrates the influence of roll canal cues on tilt perception.

Translation perception

Verbal reports of translation perception during tilt trials were very small across all frequencies, rising slightly at higher frequencies, consistent with published modeling predictions (Merfeld and Zupan 2002) and those shown in Fig. 3D. In comparison, the subjects reported substantial translation for all Translation trials regardless of the stimulus frequency (Fig. 6B). The translational perception reported during Translation trials showed great variability for all frequencies and was substantially greater than the translation perception during the Tilt trials.

DISCUSSION

Simple filtering contributes to the human translational VOR

VORs compensatory for translational motion occur in many species (Angelaki et al. 2000; Baloh et al. 1988; McCabe 1964; Niven et al. 1966; Paige 1989; Paige and Tomko 1991; Schwarz and Miles 1991; Schwarz et al. 1989). Because VOR magnitude increases with frequency during translation, it has been suggested that high-pass filtering (“frequency segregation”) elicits these responses (Paige and Tomko 1991). However, published “internal model” simulations (Merfeld and Zupan 2002) predict similar frequency response characteristics during translation without explicit high-pass filtering of otolith cues. Therefore unlike any earlier study, we measured VORs across a broad range of frequencies using both Tilt stimuli and Translation stimuli.

We found a substantial human horizontal VOR response at high frequencies (>0.2 Hz) during Tilt stimulation (Fig. 5A) and found that the horizontal VOR behaved qualitatively similarly as a function of frequency during Translation and Tilt. These findings are consistent with the hypothesis that simple filtering (Fig. 3B) yields a translational VOR whenever an inter-aural otolith cue is present. This finding is not consistent with earlier internal model predictions that little (Merfeld and Zupan 2002) or no (Angelaki et al. 1999) horizontal VOR ought to be evident during roll tilts (×, Fig. 3D), when the only inter-aural otolith cue is due to gravity.

Vergence is known to modulate the gain of the horizontal VOR (Paige 1989; Paige et al. 1998), so it is important to note that the internal model hypotheses predict little or no horizontal VOR during Tilt, even if vergence is maintained at a state consistent with a very near target. Therefore the increasingly
large horizontal VOR at frequencies above ~0.2 Hz during Tilt cannot be due to internal models and must, at least in large part, be due to the contributions of simple filters.

This finding that simple filters contribute to the human horizontal translational VOR is consistent with earlier reports suggesting that simple filtering contributes to the human (Wood 2002) and squirrel monkey (Paige and Seidman 1991) translational VORs. It also complements earlier findings that the predominant ocular tilt response, ocular torsion, results from simple filtering (Angelaki 1998; Merfeld et al. 1996b; Paige and Tomko 1991; Telford et al. 1997).

It is interesting to note that the horizontal VOR during Tilt was slightly greater than during Translation. This finding deserves further consideration because this is not consistent with simple filtering predictions that the response should be the same during Translation and Tilt. It appears that this response difference during Translation and Tilt can at least partially be explained by the presence of physiological “cross-talk” that has previously been reported in humans (Tweed et al. 1994). This earlier study showed that a small horizontal response is measured in humans when theoretically a purely torsional response would be expected during roll stimulation about an earth-vertical axis. This noncompensatory horizontal response was ~10% of the torsional response; this physiological cross-talk of unknown origin was present for rotations about an earth-horizontal axis (like our Tilt condition) and for rotations about an earth-vertical axis and was not substantially different for earth-vertical or -horizontal rotations.

At 0.7 Hz, where a 50°/s torsional response was measured in our study, cross-talk of ~10% would be expected to yield about a 5°/s horizontal response component, which is about the same as the difference found between the Translation and Tilt responses (Fig. 5A). At 0.5 Hz, such cross-talk would be expected to yield a horizontal response between 3 and 4°/s, which again is consistent with the observed difference between our Translation and Tilt horizontal responses. At lower frequencies, the torsional response becomes smaller (Fig. 5C), so the cross-talk, while still present, would yield a smaller influence on the horizontal VOR as frequency decreases. All of these expected characteristics of the physiological cross-talk are consistent with our findings (Fig. 5A), consequently, physiological cross-talk probably explains, at least to some extent, why the horizontal eye responses were larger during Tilt than during Translation.

Because previous studies have shown that human translational VORs are larger with near vergence than far vergence (Paige 1989; Paige et al. 1998), it is also possible that vergence could contribute to the measured horizontal VOR difference during Tilt and Translation. Figure 7 shows vergence versus time for one typical subject for both Tilt and Translation at the highest frequency (0.7 Hz), which was chosen because the influence of vergence is known to increase with frequency and also because the difference in horizontal VOR magnitude between the Tilt and Translation paradigms was the largest at 0.7 Hz (Fig. 5A). While small intra-trial variations are evident, the vergence responses were maintained relatively constant even at this maximum frequency of 0.7 Hz, where the VOR responses were largest. All subjects maintained vergence relatively constant during and across trials. As shown in Fig. 7, this subject demonstrated slightly greater vergence during Tilt (5.1°) than during Translation (4.4°). A similar difference in vergence during Tilt and Translation was also observed across all five subjects with the mean vergence maintained nearer for Tilt [7.0 ± 2.2° (SE)] than for Translation [4.1 ± 1.13°] at 0.7 Hz; this difference was not statistically significant (P = 0.27, t-test). Because the poster distance was the same for both Tilt and Translation, this measured difference in vergence is probably because the size of the Translation room was larger than the Tilt room, which probably factored into setting the proximal vergence state.

Thus this vergence difference might also help explain the finding that the horizontal VOR during Tilt is greater than the horizontal VOR during Translation because the increased vergence state would be expected to increase the gain of the horizontal translational VOR during Tilt. We suggest that both the presence of a torsional cross-talk during the Tilt condition and greater vergence during Tilt likely contribute to our finding that the horizontal VOR was greater during Tilt than during Translation.

Given these undesired influences on the horizontal VOR, we cannot rule out the possibility that internal model mechanisms may have also contributed to the horizontal VOR. Hypothetically, with everything else constant, such an influence of internal models would cause the horizontal VOR to be smaller during Tilt than during Translation. As discussed previously, the torsional cross-talk and increased vergence might then act to increase the horizontal response during Tilt such that the total observed horizontal response was greater during Tilt than during Translation.

Although our data unequivocally show that simple filtering contributes to the human horizontal VOR (as evidenced by the increasing response with frequency during Tilt that is not predicted by internal models), we cannot use these data to rule out the possibility that internal models might also contribute to the horizontal VOR, with these contributions masked by cross-talk and vergence variations. Therefore in a companion study, the subject experienced a single test session in a single apparatus. Vergence is maintained constant via this approach, since each subject maintains vergence relatively constant during and across trials. Cross-talk is maintained constant because the roll angular velocity is constant across trials, so the torsional VOR was also the same across trials.

---

Note: The image contains a graph showing vergence versus time with a y-axis labeled "Vergence (deg)" and an x-axis labeled "Time (sec)." The graph includes data points for both Tilt and Translation conditions at 0.7 Hz. The data suggest a difference in vergence between the two conditions, with Tilt showing a slightly greater vergence than Translation. The text provides a detailed analysis of this finding, including comparisons to previous studies and explanations for the observed differences in vergence and horizontal VOR.
we chose to investigate these VOR responses further using a new motion paradigm (Merfeld et al. 2005).

The finding that simple filtering contributes to the human translational VOR appears to conflict with earlier monkey findings (Angelaki et al. 1999, 2001) suggesting that simple filtering does not contribute to monkey translational VORs. More specifically, this finding in humans conflicts with nearly identical measurements made in rhesus monkeys (Angelaki et al. 1999) and similar measurements made in rhesus monkeys (Green and Angelaki 2003) and squirrel monkeys (Merfeld and Young 1995). Similar, and possibly related, differences in how monkey responses (Angelaki and Hess 1994; Merfeld et al. 1993b; Wearne et al. 1999) and human responses (Fetter et al. 1992; Merfeld et al. 1999, 2001) differ during motion paradigms that elicit a sensory fusion of canal and otolith cues have previously been reported. See the companion paper (Merfeld et al. 2004) for a more detailed discussion of these potential species-related influences.

The finding that simple filtering contributes to the human horizontal translational VOR would appear to contradict some earlier studies showing that internal models contribute to the human translational VOR (Merfeld et al. 1999, 2001; Peterka et al. 2004; Zupan and Merfeld 2003; Zupan et al. 2000). These studies were performed using low-frequency stimuli provided via trapezoidal stimulation, and the data showed that internal models contribute to the human horizontal translational VOR; simple filters could not reproduce the observed responses. However, these earlier studies did not refute the possibility that simple filtering might also contribute to the human translational VOR (e.g., at higher frequencies) as demonstrated via the translational VOR responses reported herein.

It is also worth noting that the contributions of the internal models may not be observable in the data presented herein because of the presence of the relatively large simple filter response components and other potential contaminants discussed previously. Therefore we conclude that simple high-pass filtering contributes substantially to the human translational VOR while simultaneously acknowledging that other studies have shown that internal models contribute to the human horizontal translational VOR.

Because both simple filtering and internal models contribute to the human translational VOR, it would be parsimonious to suggest that both response components are also present in monkeys with the internal model response components larger than the simple filtering response components. While somewhat speculative, there is some evidence that simple filtering might make a small contribution to monkey translational VOR responses because a small horizontal VOR appears to be present in monkeys during roll tilt (e.g., Fig. 4 in Angelaki et al. 1999), which could include and/or indicate a very small contribution of simple filtering to the monkey responses.

**Internal models contribute to tilt perception**

Previous studies have investigated the dynamics of roll tilt perception using low-frequency fixed-radius centrifugation and found that perceived tilt gradually aligns with the resultant force in ~1 min (Clark and Graybiel 1963, 1966; Graybiel and Brown 1951). A later study showed that changes in perceived tilt occurred more rapidly (taking just a few seconds) when subjects were tilted in roll (Stockwell and Guedry 1970).

Stockwell and Guedry compared their experimental finding to the earlier findings during centrifugation and concluded that roll canal cues during tilt contribute to rapid shifts in the perception of tilt. However, a recent study showed that the yaw rotational cues experienced during fixed-radius centrifugation delay the alignment of perceived tilt with the resultant force (Merfeld et al. 2001), raising questions about the validity of this comparison and the associated conclusion.

We found that perceived tilt was relatively constant at low frequencies and decreased as the frequency increased above ~0.05 Hz, which is consistent with an earlier report (Glasmu and Brown 1951). A later study showed that changes in perceived tilt during the Translation and Tilt paradigms. Because the difference between these paradigms was the presence (Translation) or absence (Translation) of a dynamic roll canal cue, the measured response difference leads to the conclusion that roll canal cues influence tilt perception and that the canals contribute to the perceptual processes used by the nervous system to resolve the ambiguous otolith cues. While at least one earlier study showed that yaw canal cues influence roll tilt perception (Merfeld and Zupan 2002), to our knowledge, this is the first study showing definitively that roll canal cues influence roll tilt perception, though Stockwell and Guedry (1970) had suggested this to be the case.

Our data show that rotational cues from the semicircular canals influence perceptual tilt responses across a range of frequencies where the canals are sensitive. This finding is inconsistent with predicted contributions of a simple filtering mechanism (Fig. 3A) but is consistent with internal model predictions (e.g., Fig. 3C).

At lower frequencies, subjects overestimated self-tilt (Fig. 6A). This is consistent with the observation that humans often overestimate self-tilt for small tilts during static tilt in the dark (Howard and Templeton 1966). It is worth noting that visual cues normally predominate at low frequencies but were not available during this study because all measurements were made in the dark. Low-frequency visual rotational motion cues (Dichgans et al. 1972; Zupan and Merfeld 2003) and visual perception (Asch and Witkin 1948a; Howard and Childerson 1994) cues are known to have a strong influence on roll tilt perception. The frequency response of the visual (low-frequency) and canal (higher-frequency) components complement one another, analogous to what has previously been found for neurons sensitive to angular velocity in the vestibular nuclei (Henn et al. 1974). It is reasonable to suggest that these (or similar) angular velocity neurons may contribute to the mechanisms that elicit tilt responses.

**Internal models contribute to translation perception**

We also found substantial differences in perceived translation during the Translation and Tilt paradigms. These translation reports (Fig. 6B) are qualitatively similar to internal model predictions (Fig. 3D). This similarity is surprising because the subjects reported peak-to-peak inter-aural translational displacement while the model predicts peak inter-aural linear acceleration. Linear acceleration should be doubly integrated to yield displacement, which would be predicted to yield much larger displacement reports during Translation at low frequen-
cies than at high frequencies. This prediction is inconsistent with our finding that perceived displacement is about the same (within a factor of 3–4) across more than a two-decade range of frequencies, suggesting that perfect double integration of inter-aural acceleration (modeled or real) was not achieved by our subjects. This is consistent with some earlier studies suggesting that the nervous system performs inaccurate integration, sometimes referred to as “leaky” integration (e.g., Becker and Klein 1973; Raphan et al. 1979). However, another study has suggested that linear acceleration cues are doubly integrated to yield displacement predictions (Israel et al. 1993). Our conclusion that perfect double integration is not achieved conflicts with this earlier report, but a number of methodological differences could explain this conflict. Because relatively few data are available to help us understand the dynamics of how the nervous system might process linear acceleration estimates to yield displacement estimates, we decided not to curve-fit our translation displacement data because such a post hoc fit might be misleading, especially because our experiments were not designed to investigate this integration process.

We found no other human translation displacement perception data taken across a broad range of frequencies to compare with our findings. While translation perception has been previously studied, most human studies of translation perception have focused on detection thresholds (e.g., Arrott et al. 1990; Benson et al. 1986; Carpenter-Smith et al. 1995; Gianna et al. 1996; Melvill Jones and Young 1978; Merfeld et al. 1994), manual control of translational motion (e.g., Huang and Young 1987; Merfeld et al. 1994, 1996a), eye tracking of translational displacement (e.g., Berthoz et al. 1987, 1988; Israel and Berthoz 1989; Israel et al. 1993), or translation reproduction (e.g., Berthoz et al. 1995; Israel et al. 1997).

Our finding that perceived translation differs during Translation and Tilt supports the conclusion that canal cues influence translation perception and that the canals contribute to the perceptual processes by which the nervous system resolves the ambiguous otolith cue. This is the first study to show that canal cues influence translation perception, though one earlier report (Merfeld and Zupan 2003) suggested that visual rotational cues might influence translation perception. Our perceptual findings are consistent with internal model predictions (e.g., Fig. 3, C and D) and inconsistent with predicted contributions of a simple filtering mechanism (Fig. 3, A and B).

Summary

The measured eye responses demonstrate that simple filtering contributes to the human horizontal translational VOR. Although these eye-movement data do not show evidence that internal models contribute to the translational VOR, they do not preclude that possibility. Our perceptual data clearly show that internal models contribute to perceptions of both roll tilt and inter-aural translation. Although these perceptual data do not show evidence that simple filtering contributes to perceptions of roll tilt and inter-aural translation, these perceptual data do not preclude the possibility that filtering could make a small contribution.

The finding that internal models govern human perception of both tilt and translation at first appears inconsistent with the finding that simple filtering contributes to human translational VORs. Taken together, these findings lead to the principal conclusion of this study: that qualitatively different mechanisms contribute to human perception and action. This is true even when these various behavioral measures are recorded simultaneously. While human perception and eye movements may sometimes share common neural pathways, the reported differences between eye movements and perception clearly show that one cannot simply measure eye movements and assume that this measure is representative of perception nor simply measure perception and assume that this measure is representative of reflexive responses. Taken together, these behavioral measures (both eye movements and perception) can be used to determine what strategies the nervous system uses to process the incoming sensory information that elicit these behavioral responses and to guide neural recordings that elucidate how these strategies are implemented by the nervous system.

Appendix

Both models included two primary physical effects (Figs. A1 and A2). One of these physical effects is that head rotations (e.g., head “nodding”) influence the relative orientation of gravity with respect to the head. Mathematically, this influence was exactly captured by the integral equation $g(\omega^0) = \int_0^t \omega(t) \times g(t) dt$, where $\omega$ is the angular velocity of the head. The second physical effect was that specific gravitio-inertial force ($f$), which is force per unit mass, as measured by the otolith organs (or any other graviceptor/linear accelerometer), resulted from the combination of gravity ($g$) and linear acceleration ($a$). Mathematically, this was captured by the equation, $f = g - a$.

“Simple Filtering” Simulations. The simple filtering hypothesis was implemented using simple, first-order, fixed-parameter, low-pass and high-pass filters (Fig. A1). The low-pass filter [LPF(s)] used was $[g\hat{g}(s) / f_{\hat{g}}(s)] = 1/(\tau g_{\hat{g}} + 1)$, where $\hat{g}$ is the low-pass filtered estimate (units of m/s²) of inter-aural force, $f_{\hat{g}}$ is the inter-aural force stimuli.

The measured eye responses demonstrate that simple filtering contributes to the human horizontal translational VOR. Although these eye-movement data do not show evidence that internal models contribute to the translational VOR, they do not preclude that possibility. Our perceptual data clearly show that internal models contribute to perceptions of both roll tilt and inter-aural translation. Although these perceptual data do not show evidence that simple filtering contributes to perceptions of roll tilt and inter-aural translation, these perceptual data do not preclude the possibility that filtering could make a small contribution.

Possible causes include differences in the translation perception tasks, different motion profiles, uncontrolled differences in cognition, and/or differences in other motion characteristics (e.g., vibration, acoustics, wind cues, etc.).
 previously been published (e.g., Merfeld 1995b; Merfeld and Zupan 2002; Merfeld et al. 1993a), the description provided here will be brief. The internal model hypothesis suggests that the nervous system somehow “knows” the two physical effects described previously ($f = g - a$ and $g(t') = f^T \omega(t) \times g(t') dt$) and mimics these physical laws via neural mechanisms called internal models. The internal models that mimic these two physical effects are represented mathematically as $\dot{\gamma} = \dot{g} - \dot{a}$, which is the summation synapse shown as a gray circle in the middle of Fig. A2, and $\dot{g}(t) = f^T \dot{\omega}(t) \times g(t) dt$, which is shown as the rotational neural integrator (RNI). The “cap” symbol ($) is used to distinguish the neural quantity (e.g., $\dot{g}$) from the physical quantity (e.g., $g$). The model includes just four free parameters and has been shown to be robust to changes in model parameters (Merfeld and Zupan 2002; Merfeld et al. 1993a). We include model predictions (Fig. 3, C and D, solid lines) performed using the published model (Merfeld and Zupan 2002) prior to experimentation. We also include simulated responses after making two parameter changes to the model (Fig. 3, C and D, dotted lines) to better fit the experimental data.

Specifically, for these latter simulations, the entire structure of the model and the vast majority of parameters were unchanged relative to a recent publication of the model (Merfeld and Zupan 2002). The only change was that two feedback parameters ($k_1$ and $k_2$) were changed from 2.0 (radians/s per radian) to 1.0 to better match the tilt perception frequency characteristics measured during translation (Fig. 6A).

The principal pathway by which canal cues influence the estimation of gravity is shown using thick black lines. In brief, the canal cues are compared with the expected canal cues and the neural signal representing this difference is then scaled and utilized by the rotation neural integrator (RNI) to help keep track of the relative orientation of gravity. These pathways (thick black lines) allow the internal model to use rotational cues from the canals to help estimate the relative orientation of gravity. In other words, this is how the available canal cues during tilt stimulation help accurately estimate tilt with respect to gravity across a broad range of frequencies where the canals are effective.

Furthermore, if the nervous system accurately estimates tilt, then it “knows” that a change in the inter-aural otolith cue during roll tilt is due to a change in orientation with respect to gravity and can determine that little or no linear acceleration is present. These linear acceleration calculations are performed by the neurons and neural pathways represented by thick gray lines. Thus the influence of canal cues on linear acceleration estimation is accomplished via the same pathway discussed above describing the influence of canal cues on gravity estimation plus the additional contributions of the pathway shown using thick gray lines. These additional contributions involve a negative feedback loop ($k_2$ is negative), and a summation synapse where the signal from linear acceleration neurons ($\dot{a}$) subtracts from the signal representing the gravity neurons ($\dot{g}$).

The canal cues are compared to the expected canal cues via the synapse on the far right of Fig. A2. The simulated neural response of this family of cells matches the characteristics of the vestibular-only (VO) cells in the vestibular nuclei that have recently been shown to respond to passive head velocity with little sensitivity to active head rotation (Boyle et al. 1996; McCrea et al. 1996, 1999; Roy and Cullen 2001). While active contributions are not explicitly included in this version of the model, because all motion herein is passive, hypothesized mechanisms for cancellation of active contributions have previously been presented using this same model structure (Merfeld 1995b; Merfeld et al. 1993a).

Cells showing characteristics consistent with these hypothesized tilt neurons ($\dot{g}$) have been found in the vestibular nuclei (Zhou et al. 2000). Neurons showing characteristics similar to the modeled linear acceleration units ($\dot{a}$) included in this paper have also been found in the vestibular nuclei in the brainstem (Dickman and Angelaki 2002; Peterson and Chen-Huang 2002) and the cerebellar fastigial nuclei (Angelaki et al. 2004).

“INTERNAL MODEL” SIMULATIONS. Model predictions were simulated using a previously published model (Merfeld and Zupan 2002). Because comprehensive descriptions of this model (Fig. A2) have
ACKNOWLEDGMENTS

The authors thank T. Bennett and V. Stallings for technical contributions, M. Marsden and P. Cunningham for administrative assistance, and Drs. Lionel Zupan and J. Lewis for commenting on early versions of the manuscript. Experiments were performed at the Legacy Neurotology Research Laboratory.

GRANTS

The authors acknowledge support from National Institute on Deafness and Other Communication Disorders Grants DC-004158 to D. M. Merfeld and S. Park and DC-00205 to C. Gianna-Poulin, S. Wood, and F. Owen Black as well as National Aeronautics and Space Administration Grants NNNJ04HB01G to D. M. Merfeld and NAW9-1254 to F. O. Black.

REFERENCES


Berthoz A, Israel I, Vitte E, and Zee D. Linear displacement can be derived from otolithic information and stored on spatial maps controlling the saccadic system. Adv Otorhinolaryngol 41: 76–81, 1988.


Henn V, Young L, and Finley C. Vestibular nucleus units in alert monkeys are also influenced by moving visual fields. Brain Res 171: 144–149, 1979.


McCrea RA, Gdowski GT, Boyle R, and Belton T. Firing behavior of vestibular neurons during active and passive head movements: vestibulo-