Postural responses exhibit multisensory dependencies with discordant visual and support surface motion

Emily A. Keshner\textsuperscript{a,b,∗}, Robert V. Kenyon\textsuperscript{c} and Jessica Langston\textsuperscript{a}

\textsuperscript{a}Sensory Motor Performance Program, Rehabilitation Institute of Chicago, Chicago IL, USA
\textsuperscript{b}Department of Physical Medicine and Rehabilitation, Feinberg School of Medicine, Northwestern University, Chicago IL, USA
\textsuperscript{c}Department of Computer Science, University of Illinois at Chicago, Chicago IL, USA

Keywords: Virtual reality, vestibular, postural control, motion analysis, sensory re-weighting

Received 6 June 2003
Accepted 10 December 2003

Abstract. The purpose of this study was to identify how the postural system weights coincident yet discordant disturbances of the visual and proprioceptive/vestibular systems. Eleven healthy subjects (25–38 yrs) received either fore-aft translations of an immersive, wide field-of-view visual environment (0.1 Hz, ± 3.7 m/sec), or anterior-posterior translations of the support surface (0.25 Hz, ± 15 cm/sec), or both concurrently. Kinematics of the head, trunk, and shank were collected with an Optotrak system and angular motion of each segment plotted across time. With only support surface translation, segmental responses were small (1°–2°) and mostly opposed the direction of sled translation. When only the visual scene was moving, segmental responses increased as the trial progressed. When the inputs were presented coincidentally, response amplitudes were large even at the onset of the trial. Mean RMS values across subjects were significantly greater with combined stimuli than for either stimulus presented alone and areas under the power curve across subjects were significantly increased at the frequency of the visual input when both inputs were presented. Thus, intra-modality dependencies were observed, such that responses to the visual inputs significantly increased and responses to the somatosensory signals reflected the stimulus amplitude only when the two inputs were combined. We believe it unlikely that the role of any single pathway contributing to postural control can be accurately characterized in a static environment if the function of that pathway is context dependent.

1. Introduction

There is significant evidence that dynamic visual inputs induce postural instability through measures of increased sway during quiet stance [26,46,47]. Studies with visual field motion have demonstrated large center of pressure changes [9,10,31], with the most robust postural changes in the roll and pitch planes [14,46] and at frequencies below 0.2 Hz [4,21,31] Motion of the visual field affects more than center of pressure values, however. Velocity and frequency of visual field stimuli have been correlated with individual segmental velocities [10,13,24,27,28,35], muscle electromyographic amplitudes [12], and the direction of gaze [17] in quietly standing subjects.

We might hypothesize that when all sensory systems are functioning and subjects are faced with coincident disturbances in the visual field and at the base of support, they are able to disregard the information from the environment (visual inputs) and attend only to their in-
ternal signals (proprioceptive and vestibular) to maintain postural orientation. But our previous study of healthy subjects walking within a room with a virtual environment projected at a constant velocity in roll that was uncorrelated with the parameters of their locomotion revealed that the subjects were forced to either alter the organization of their locomotion pattern or lose their balance while walking [24].

The purpose of this study was to identify how the postural system weights coincident yet discordant disturbances of the visual and proprioceptive/vestibular systems. Most studies examining the effect of vision on postural reactions to a disturbance of the support surface have investigated the effect of the visual inputs only with earth-fixed visual environments, or with eyes closed [1,5,23,51], or by diminishing the reliability of somatosensory inputs with a sway-referenced support surface [39,44]. These studies have also placed the visual axis of rotation at the ankle and assumed that the multisegmental body functions as an inverted pendulum. In this study we exposed our subjects to simultaneous motion of an immersive, wide field-of-view stereo image of a virtual environment and perturbations of the support surface. The frequency, velocity, and amplitude of the visual scene and the support surface perturbation were large enough to elicit a physical response [5], but different enough from each other to present stimuli with no common parameters. We measured segmental kinematics and calculated power of the segmental response at each input frequency when presented individually and concurrently to examine how segmental components of the postural system principally responded to a single modality and multimodal parameters were incorporated into segmental behaviors.

2. Methods

2.1. Subjects

Eleven healthy young adult subjects (aged 25–38 yrs) gave informed consent according to the guidelines of the Institutional Review Board of Northwestern University Medical School to participate in this study. Subjects had no history of central or peripheral neurological disorders or problems related to movements of the spinal column (e.g., significant arthritis or musculoskeletal abnormalities) and a minimum of 20/40 corrected vision. All subjects were naive to the virtual environment (VE).

2.2. Apparatus

A linear accelerator (sled) that could be translated in the anterior-posterior direction was situated within a light-tight room. Motion of the sled was controlled by D/A outputs from an on-line PC. The sled was placed 40 cm in front of a screen on which a virtual image was projected via a stereo-capable projector (Electrohome Marquis 8500) mounted behind the back-projection screen (Fig. 1).

The projection wall in our system consisted of back projection material measuring 1.2 m × 1.6 m. An Electrohome Marquis 8500 projector throws a full-color stereo workstation field (1024x768 stereo) at 96 Hz onto the screen. An SGI desk side Onyx II with a Reality Engine created the imagery projected onto the wall. The field sequential stereo images generated by the Onyx II were separated into right and left eye images using liquid crystal stereo shutter glasses worn by the subject (Crystal Eyes, StereoGraphics Inc.). These glasses limited the subject’s horizontal FOV to 100° of binocular vision and 55° for the vertical direction. The correct perspective and stereo projections for the scene were computed using values for the current orientation of the head supplied by a position sensor (Flock of Birds, Ascension Inc.) attached to the stereo shutter glasses (head). Consequently, virtual objects retained their true perspective and position in space regardless of the subjects’ movement. The total display system latency measured from the time a subject moved to the time the resulting new stereo image was displayed in the environment was 50–75 ms. The stereo update rate of the scene (how quickly a new image is generated by the graphics computer in the frame buffer) was 14–24 stereo frames/sec. Flock of birds data was sampled at 48 Hz.

2.2.1. Scene characteristics

The scene consisted of a room containing round columns with patterned rugs and painted ceiling (Fig. 1). The columns arranged in a regular pattern were 6.1 m apart and rose 6.1 m off the floor to the ceiling. The rug patterns were texture mapped on the floor and consisted of 10 different patterns. The interior of the room measured 30.5 m wide by 6.1 m high by 30.5 m deep. The subject was placed in the center of the room between two rows of columns. The subjects stood 3 m from the nearest pillar on either side of them. Since the sled was 64.8 cm above the laboratory floor the image of the virtual room was adjusted so that its height matched the sled height (i.e., the virtual floor
and the top of the sled were coincident). Beyond the virtual room was a landscape consisting of mountains, meadows, sky and clouds. The resolution of the image was 7.4 min of arc per pixel when the subject was 40 cm from the screen. When the scene moved in fore-aft, objects could come into view and then go out of view depending on their position in the scene.

2.3. Procedures

Subjects stood on the sled with their hands crossed over their chest and their feet together. The first trial for all 11 subjects was to stand quietly in the dark for 80 sec. Seven of the subjects then received a $\pm 15$ cm/sec, 0.25 Hz sinusoidal translation of the sled moving $\pm 10$ cm in the anterior-posterior direction. The next trial was a $\pm 3.7$ m/sec, $\pm 3$ m, 0.1 Hz sinusoidal fore-aft motion of the computer generated stereo image. In the last trial they received the sled and scene motion concurrently. Four additional subjects received the concurrent sled and scene motion first, followed by the scene motion alone and then the sled motion alone. The sled frequency was selected in order to produce a postural disturbance having minimized intersegmental joint motion [5] but enough vestibular information to be reliable [7]. The visual frequency was selected because Lestienne et al. [31] have demonstrated that the destabilization induced by visual cues conflicting with vestibular and proprioceptive cues is roughly equivalent to the absence of visual cues at frequencies be-
low 0.2 Hz and postural sway is increased. Each trial lasted a total of 280 sec. In all trials, 20 sec of data was collected before scene or sled motion began (pre-perturbation period). In the trial where both scene and sled moved, the scene moved alone for 20 sec and then the sled was also translated for the next 220 sec. In the last 20 sec only the visual scene moved. When only the sled was translated, the visual scene was visible but stationary, thus providing appropriate visual feedback equivalent to a stationary environment.

2.4. Data collection and analysis

Three-dimensional kinematic data from the head, trunk, and lower limb were collected at 100 Hz using video motion analysis (Optotrak, Northern Digital Inc., Ontario, Canada) with a resolution of 0.1 mm. Infrared markers placed near the lower border of the left eye socket and the external auditory meatus of the ear (corresponding to the relative axis of rotation between the head and the upper part of the cervical spine) were used to define the Frankfort plane and to calculate head angular position relative to the earth vertical [45]. Other markers were placed on the back of the neck at the level of C7, the left greater trocanter, the left lateral femoral condyle, the left lateral malleolus, and on the translated surface. Markers placed at C7 and the greater trocanter were used to calculate trunk angular position relative to earth vertical. Knee angular position was the angle from the greater trocanter to the lateral femoral condyle relative to earth vertical, and shank angular position was the angle between the lateral femoral condyle and the lateral malleolus relative to earth vertical. For trials where the sled moved, sled motion was subtracted from the linear motion of each segment prior to calculating segmental motion. Segmental angles were then calculated with respect to the initial position of the segment in each trial (pre-perturbation). The angular position of each segment was plotted relative to time.

Root mean square (RMS) values were calculated at 40 sec intervals for the segmental angles to examine changes in response magnitude across trials. Power of the segmental response at each stimulus frequency was calculated using a 40 sec sliding window following a Fast Fourier transform analysis, and area under the power curve was calculated. Differences in mean segmental RMS, peak excursion differences, and mean areas under the power curve were tested for significance with repeated measures ANOVA and paired t-tests. Bonferroni post-hoc comparisons were performed on the dependent measures at the $p < 0.05$ level.

3. Results

3.1. Postural sway with sled motion and earth-fixed vision

When only the sled was translated the visual scene behaved as any earth-fixed environment would (e.g., the laboratory). Consequently subjects perceived that they were standing within the computer generated room and that this room was fixed to earth. Subjects did not perceive the room as moving with the platform, but as fixed to the earth and they were being translated within the computer generated room. Therefore there was no conflict between the movement generated feedback from the sled and the visual feedback. Responses of the head, trunk, and shank were in the same direction in about half of the subjects (see subjects 3–5, 7–9 in Fig. 2), thereby acting as an inverted pendulum. In the other subjects (subjects 1, 2, 6, 10, 11 in Fig. 2), the head and trunk moved opposite the direction of the sled but the shank moved in the direction of the sled implying that subjects were bending at the hip. Amplitudes of the trunk stayed constant over time (usually about 1°–2° peak to peak), although some subjects reduced the amplitudes of the shank response over time (see 4th and last subject in Fig. 2). Amplitudes of the head were greater than the trunk (see Fig. 6), and head saccades were observed in 3 of the subjects (subjects 5, 8, and 11 in Fig. 2). The power spectra for all subjects (Fig. 3) indicated that the primary response frequency was the same as the frequency of the stimulus (0.25 Hz) and that the power was about the same in all segments although the head tended to exhibit considerable low frequency noise.

3.2. Postural sway with visual-only motion

When only the visual scene was moving, the performer would experience a conflict between the visual perception of motion and the vestibular and somatosensory systems signaling an absence of physical motion. A decision would need to be made about whether the visual motion signal was due to self motion or motion of the environment. Response amplitudes were less with visual scene motion than with sled motion (Fig. 6). In some subjects the initial small response increased from about 1°–2° to 5°–6° (e.g., subjects 1 and 7 in Fig. 4), but this was not a significant trend (Fig. 6). In the beginning of the trial, some subjects did not respond to the scene motion, but in the latter part of the trial their segmental responses were matched with
Fig. 2. Angular motion during the sled motion only trial is plotted for all subjects. Each row is the response of the three body segments (head, trunk, and shank) of one subject for an early (left plot) and middle (right plot) 40 sec period of the trial. Motion of the sled is plotted in the bottom traces. The thin vertical line in each column has been added to help trace the response to an upward peak of the sled. Upward peaks of the data represent anterior motion relative to the room. Downward peaks of the data represent posterior motion relative to the room.

the sinusoidal motion of the scene (see head subjects 1, 6, 7, and 8 in Fig. 4). The peak of the segmental responses was delayed in most subjects. Responses took at least 20 sec to occur, and some responses took as long as 150–200 sec (see 1st and 7th subjects in Fig. 4). Two response strategies were observed across subjects. Either the subject responded as an inverted pendulum with the head, trunk, and shank moving in the direction of the visual scene motion at the same time (subjects 3, 7–10 in Fig. 4). Or, the trunk and shank moved together, often with smaller amplitudes of motion than the head, and the head exhibited large saccades (subjects 1, 2, 4–6, and 11 in Fig. 4). Examination of the power spectra (Fig. 3) reveals that the stimulus frequency (0.1 Hz) was the primary response frequency in most subjects but had the least amount of power in the shank except for one subject (subject 10 in Fig. 4).

3.3. Postural sway with both visual and sled motions

When both the sled and visual scene were moving, the vestibular and somatosensory systems were signaling real motion of the body. But feedback from the visual system conflicted because of the spatial and temporal discordance between the sled and the visual scene. Segmental response amplitudes increased dramatically in this condition and, unlike the trial in which only the scene was moving, segmental response amplitudes were large from the beginning of the dual stimulus trial (Fig. 6). The segmental responses were more complex with combined stimuli, often exhibiting multiple frequency components in the response (e.g., subjects 2 and 6 in Figs 5 and 7). Only four subjects synchronized the motion of all their segments thus acting as an inverted pendulum (subjects 3, 6, 7, and 9 in Fig. 5). All of the other subjects kept the head and trunk moving in the
same direction at the same time, but the responses of the shank were very different. Motion of the head and trunk peaked with the direction of scene motion while the ankle tended to follow the peaks of sled motion (Fig. 7). Examination of the segmental power spectra for combined sled and scene motion (Fig. 3) indicated that the increased response activation occurred predominantly at the frequency at the visual stimulus (0.1 Hz). Power at the frequency of sled motion did, however, demonstrate a small increase in the shank with respect to the response to sled motion alone. Plotting the power spectra for the three conditions across time (Fig. 8A) confirmed that the visual scene frequency became the dominant frequency in the response to combined sled and scene motion. With sled motion alone, the power of the segmental responses at the frequency of the input remained relatively constant across the trial. With scene motion alone, there was usually either little power at the visual scene frequency, or a dramatic increase in the power of the segmental response late in the trial. With both sled and visual scene motion, subjects demonstrated either no increase (see 1st subject in Fig. 8A) or a small increase in the power at the frequency of the sled.

A significant increase in power of the segmental responses was elicited at the frequency of the visual stimulus, and this was true across all subjects (Fig. 8B). Area under the mean power curves was significantly greater for the head ($F(3,30) = 11.91$, $p < 0.0001$) and trunk ($F(3,30) = 11.35$, $p < 0.0001$) across subjects at 0.1 Hz during combined sled and visual scene motion than to either frequency when presented alone or to the sled frequency when the stimuli were combined. Area under the mean power curve of the shank was also greater to the combined stimuli at the frequency of the scene ($F(3,30) = 6.50$, $p < 0.002$) than to either the sled or scene presented alone. This increase in power at the visual frequency was significantly greater than a simple summing of the area under the mean power curves of the trials with single frequency inputs ($t(32) = 2.35$, $p < 0.01$).

There was a greater overall effect of the combined stimuli on the trunk and shank than in the head. RMS values of the head with combined inputs was significantly greater than either stimulus presented alone only
in the latter portion of the trial ($F(6, 60) = 3.41, p < 0.006$), and the response to the combined inputs was not significantly greater than the response of the head in the dark. RMS values of the trunk with combined inputs were significantly greater than those when standing quietly in the dark and to either the sled or scene motion alone ($F(6, 60) = 21.15, p < 0.0001$). The same was true of the shank ($F(6, 60) = 12.47, p < 0.0001$), and in addition, RMS values of the shank were significantly smaller when standing quietly in the dark than when the sled alone was translated ($p < 0.05$).

4. Discussion

The results presented here argue that the response to visual information was strongly potentiated by the presence of physical motion (e.g., Fig. 5). Either stimulus alone produced marginal responses in most subjects, but when combined, the response to visual stimulation was dramatically enhanced even though the visual inputs were incongruent with those of the physical motion.

Preliminary studies in our laboratory demonstrated that subjects responded much less to fore-aft motion than to roll stimuli of the same magnitude. For that reason we designed a large visual stimulus motion so that we could insure a response to the visual input. Nonetheless, the size of the visual stimulus motion doesn’t negate the finding that when motion of the sled was combined with motion of the visual scene, the response to the visual input was much larger than with a stationary sled and a moving visual scene although in both cases subjects were receiving discordant feedback from visual, vestibular, and somatosensory systems. It is possible, however, that our results may not generalize to other amplitudes and frequencies of visual motion.
Fig. 5. Angular motion during the combined sled and scene motion trial is plotted for all subjects. Each row is the response of the three body segments (head, trunk, and shank) of one subject for an early (left plot) and middle (right plot) 40 sec period of the trial. Motion of the scene (bold line) and the sled (thin line) is plotted in the bottom traces. The thin vertical line in each column has been added to help trace the response to an upward peak of the sled and scene. Upward peaks of the data represent anterior motion relative to the room. Downward peaks of the data represent posterior motion relative to the room.

Although we employed the same visual frequency as Peterka and Benolken [44], saturation of visually induced sway was not observed in either our current or previous [24] data. The most probable explanation for this difference was that their subjects were physically restricted so that they could only respond at the ankle joint. Our subjects were free to move at all joints, and indeed the greatest increase in motion with combined inputs was observed mostly at the head and trunk signifying that bending at the hip rather than an inverted pendulum was the strategy of choice when the visual world conflicted with support surface motion. When only the platform moved and the scene was earth fixed, subjects’ responses were evenly divided between an inverted pendulum and hip strategy. This strategy selection could have been a random occurrence, but we believe they chose the strategy they preferred because the response selection was stable within each subject.

Another consideration is that many investigators [15, 37,39,43] have chosen to sway reference the support platform in order to diminish feedback from the lower limb, thereby assuming that subjects would be more reliant on vestibular inputs when the eyes were closed. We have chosen to keep all inputs activated and have observed that rather than shifting from a reliance upon one pathway to another, subjects were attempting to incorporate characteristics of all relevant pathways into their response. For example, with combined inputs, subjects not only increased their response to the visual signal but also increased their responses to the sled motion. As observed by McCollum [36], sensory reafference depends very much upon the specific movement performed, but the movement response that is chosen depends equally as much upon the sensory inputs that are available to the performer.
There is ample evidence that postural response characteristics are task dependent [19,20,23,38,40] and this finding has been used to suggest the presence of supraspinal control of postural responses. Motor evoked potentials from the soleus and tibialis anterior muscles during transcranial magnetic stimulation [30,50] and galvanic vestibular stimulation [15] have demonstrated that the short and medium latency responses of these muscles were larger when standing on an unstable as compared to a stable support surface. Under quiet sway conditions (e.g., when there was no motion of the support surface), inputs from the vestibular and somatosensory systems could be relied upon to reduce unexpected or discordant inputs from the visual system via inhibitory reciprocal pathways [2,11]. But increased response amplitudes with combined inputs imply that the absence of any static cues or the presence of unpredictable cues interfere with the subject’s ability to differentiate between self-motion and environmental motion [8,18,34]. Thus, when the world is moving as well, exteroceptive feedback (e.g., vision) becomes more important because we have to shape our responses to match the changing external constraints [16, 49]. Continuous modulation of the postural responses during periods of unpredictable instability would require an adaptive gain controller, most probably at the cortical level [53].

Supraspinal control of these postural responses may also explain the variability we observed across subjects. We speculate that the postural behaviors elicted in this paradigm were not the behavior of sub-cortical reflex pathways but emerged from a complex perceptual process defining vertical orientation. For example, a study of subjects rotating in a linearly moving environment concluded that subjects used conflicting visual and non-visual information differently according their individual perceptual styles [29], i.e., whether they weighted the visual or proprioceptive information more heavily. Other studies have demonstrated that perceptual weighting of visual and vestibular inputs was a good predictor of the subject’s reliance on visual reafference to stabilize posture [22,32], and recent PET and MRI studies have demonstrated that simultaneous visual and vestibular stimulation activated the medial parieto-occipital visual area and parieto-insular vestibular cortex, potentially to resolve intersensorial conflict [2,11].

Oie et al. [41] predicted that the gain to a particular sensory input would increase when that input becomes more reliable. We found this to be true when both the physical and visual stimuli were moving. Even though they were disparate in frequency content, the two sensory inputs reinforced each other. A similar phenomenon has been observed in flight simulation studies using transient rather than sinusoidal physical motion [54]. The pilot becomes more responsive to the visual scene when it is combined with physical motion. As in our paradigm, the physical excursion in the simulator is a small percentage of the visual excursion but the pilots respond appropriately to the
Fig. 7. Angular motion of three body segments (head, trunk, and shank) is plotted for an early (left plots) and middle (right plots) 40 sec period of the trial with both sled and scene motion for two subjects (subjects 6 and 10 in Fig. 5). The data has been repeated to better compare with each stimulus. Upward peaks of the data represent anterior motion relative to the room. Downward peaks of the data represent posterior motion relative to the room.

The responses observed here were much greater than one would expect from a simple summation of inputs, and multiple input frequencies were included in the response. Thus a simplified linear model, such as that proposed by Peterka [42], could not be applied to these results. The non-linear increases observed in these responses are instead indicative of a time dependent, non-linear postural system [6,32] attempting to resolve the mismatch between the visual and vestibular inputs, perhaps through sensory re-weighting [41,43,52]. Similar to the findings of Oie et al. [41], intra-modality dependencies were observed, such that responses to the visual inputs significantly increased and responses to the somatosensory inputs quantitatively reflected the somatosensory amplitude only when the two inputs were combined.

We infer from our findings of multiple response frequencies and increased response magnitudes with combined inputs that, in our protocol, the postural system does not respond as a redundant system capable of ignoring one input pathway in favor of another. Instead, each input is accommodated by a continuous monitoring of the environmental signals in order to appropri-
ately modulate the frequency and magnitude characteristics of the response [41]. It is important to note that our findings of visual enhancement would not have been observed if we had relied only upon the sway response as measured through the center of pressure. As reported in previous studies [24,25], the enhanced response to the visual component was only observed in the head and trunk; the shank continued to be primarily responsive to input from the base of support. These data support previous assertions that the shank is differentially controlled by base of support inputs whereas the head and trunk were more responsive to the visual inputs.

The current findings have significant impact on studies of motor control and on rehabilitation interventions. In the past, postural responses have been examined...
through isolating individual control pathways in order to determine their specific contribution. However, if these pathways are responsive to functionally relevant contexts, then their response may well be different when the system is receiving simultaneous inputs from multiple pathways especially when those systems exhibit non-linear behaviors. Furthermore, we believe it unlikely that the role of any single pathway contributing to postural control can be accurately characterized in a static environment if the function of that pathway is context dependent.

Acknowledgements

This work was supported by National Institute of Health grants DC01125 and DC05235 from the NIDCD and AG16359 from the NIA. We gratefully acknowledge the permission to use the CA VElib and TrackD software to generate and control the virtual scene from VRCO, Virginia Beach, VA.

References


