

Can your eyes make you sick?: Investigating the Relationship between the Vestibulo-ocular Reflex and Virtual Reality

Mark H. Draper

29 Apr 1996

This paper discusses the nature of the vestibulo-ocular reflex and its relationship to the tracking response of the eye. Included is a simplistic descriptive model of these mechanisms. Speculations are then made on the potential contributions of the visual and vestibular systems in producing "simulator sickness" and in particular the kinds of artifacts in virtual interfaces that may contribute to this condition.

1.0 Nature of the Vestibulo-ocular Reflex (VOR)

The VOR is a primitive eye-movement reflex that stabilizes visual functions to keep images stabilized on the retina during movement of the head. To facilitate a better understanding of this reflex, a brief overview of the vestibular apparatus will first be presented. This will be followed by a detailed description of the VOR, how it is measured, its characteristics, and its ability to adapt to new situations.

1.1 The Vestibular Apparatus

The vestibular apparatus is a small structure that exists in the bony labyrinth of the inner ear (i.e. there are two vestibular organs, one in each inner ear)(see Howard, 1986a). Its function is to sense and signal movements of the head. This function, although basic, is vitally important because it contributes to the coordination of motor responses, eye movements, and posture. In fact, individuals that have had partial or complete loss of vestibular functioning have found it difficult to perform even the most basic of tasks (e.g. standing, walking, or reading). The vestibular organ consists of two principle sets of structures, the semicircular canals and the otolith organs (Figure 1), which work together to provide optimum information on head movement and positioning. The VIII nerve is the efferent pathway for vestibular signals, transmitting head movement and head positioning data to various centers in the brain and postural control nuclei.

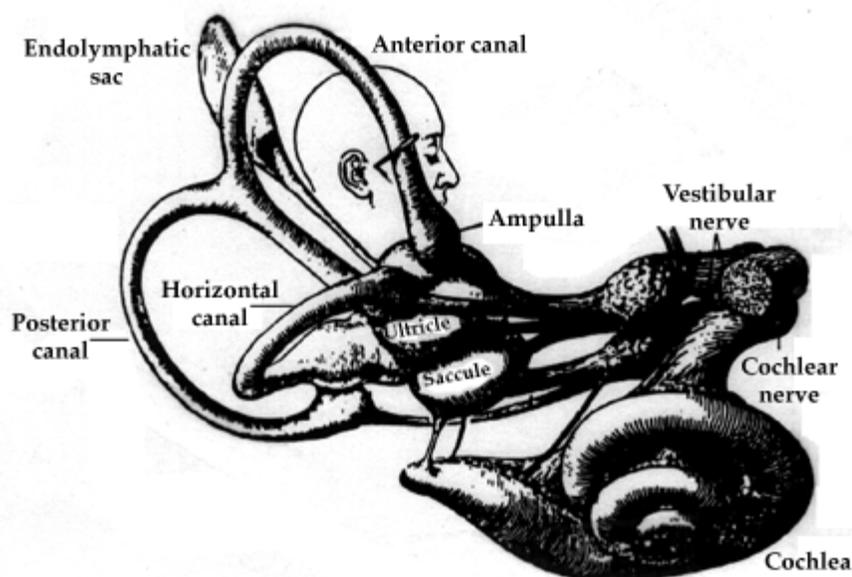


Figure 1: The Vestibular Apparatus (from Howard, 1986a)

There are three semicircular canals (SCC) (termed the anterior, posterior, and horizontal canals) in each vestibular organ whose function is to detect angular accelerations of the head, acting like biological accelerometers. These canals are bi-directionally sensitive and approximately mutually perpendicular so as to detect angular head movement in any direction. Endolymph fluid fills each SCC and is prevented from passing through the ampula (a widened section of each SCC) by the cupula, a thin flap that stretches across the ampula and acts as a barrier to endolymph flow. When the head is rotated, the force exerted by the inertia of the fluid acts against the cupula of those SCCs that are in the plane of motion, causing it to deflect/bend. This deflection causes a displacement of tiny hair cells (located at the base of the cupula in the ampula) which signal this change to the brain via the VIII nerve. For most normal head movements (moderate frequencies), this signal is proportional to head velocity. The receptor system of the SCC can respond to angular accelerations as low as 0.1 deg./sec^2 . However, continued constant angular rotation will decrease the response of the SCC as the fluid decreases its inertial force of the cupula due to lack of acceleration and also due to the elasticity of the cupula striving to return to its initial state. The time constant for the cupula in humans has been found to be approximately 5 to 7 seconds. Therefore, the SCCs encode dynamic changes in head movement only. Also, it is important to note that each SCC transmits a tonic (resting) signal even in absence of motion. This allows the SCC to increase or decrease its response, depending upon whether the head movement is in the direction that the SCC is most sensitive to or in the opposite direction. The brain then integrates information from each SCC pair that occupy the same plane of motion (termed a push-pull pair) to generate an appropriate response for motion in that plane. Between the two vestibular organs, there is a total of three of these push-pull pairs corresponding to three different planes of motion.

In addition to the three SCCs, each vestibular apparatus also contains two otolith organs, the utricle and saccule, that sense dynamic changes in linear acceleration of the head and also provide information on static head position, such as head tilt. These otolith organs are multi-directionally sensitive, as opposed to the bi-directional nature of the SSCs. The receptor portion of these organs contains many hair cells and is called the macula. The macula is covered with a gelatinous substance that contains tiny crystals of calcium carbonate called 'otoliths'. When the head is tilted or undergoes linear acceleration, the otoliths deform the gelatinous mass, which creates a shear force that excite the receptor cells in the macula. This information is then transmitted via the VIII nerve. The utricle's macula is located in the horizontal plane so as to be sensitive primarily to horizontal linear accelerations, while the saccule's macula is positioned vertically to be maximally sensitive to vertically directed linear accelerations, including gravity (Robinson, 1980).

1.2 VOR Description and Characteristics

The greatest potential source of image slip on the retina is due to self-rotation (Robinson, 1980). Therefore, in order to see and move at the same time, the eyes must have an ability to remain stabilized in space as the head rotates. As stated earlier, the VOR is a fundamental eye-movement reflex that functions to keep images stabilized on the retina during movement of the head. Thus it helps to perform a very basic but important function, to allow sight during movement.

When the head begins to move in any direction, the vestibular apparatus senses this movement and sends direction and rate information directly to the oculomotor system. The oculomotor system then responds by moving the eyes (in a conjugate manner) in an equal but opposite direction/rate to compensate for the head movement and keep the visual image stabilized on the retina. This, at a top level, is the VOR. It is a very low latency system in response to head movements, with compensatory eye movements beginning as early as 10-20 msec. after head rotation begins (Viirre, Tweed, Milner & Vilis, 1986).

Although both the SCCs and the otolith organs contribute to the VOR, most researchers agree that the otolith input is minor and transitory while the SCC input dominates the response (Robinson, 1980). Therefore, the rest of this paper will only consider the SCC input to the VOR. However, researchers have shown that under certain conditions (i.e. off-axis rotation, close-in target of fixation) the otolith organs can have a demonstrable affect on the reflex (Viirre et al, 1986; Merfeld, 1995).

A more detailed description of the VOR process is as follows. First an angular acceleration of the head will cause the appropriate SCC cupulas to deflect in the vestibular organs of each ear (creating push-pull pairs as described above).

This deflection will cause the tiny hair receptor cells at the base of the cupula to send head-velocity proportional signals as either excitatory patterns (due to motion in the direction of the directionally sensitive SCC) or inhibitory patterns (due to motion opposite the direction of the directionally sensitive SCC) to the ipsilateral vestibular nucleus (VN). The VN then sends the appropriate eye velocity signal to the oculomotor nuclei (ON), which in turn innervates the three complimentary pairs of muscles that move each eye. This path is often simplistically termed the 'three-arc reflex' (Figure 2). It has also been theorized that the oculomotor neurons require more than just an eye velocity command to drive the eye a certain direction and speed, they also need an eye position command to hold the eye at the new position so the elasticity of the eye plant does not cause the eye to drift back to its original position after the head movement ends. Robinson (1980) and others (Zee, 1980?) have argued there must exist a neural integrator in the system that would integrate the velocity signal from the VN to obtain the required position signal. Goldberg, Eggers, and Gouras (1991) state that this neural integrator, although still undiscovered, requires the cerebellar flocculus, the medial vestibular nucleus, and the nucleus prepositus hypoglossi to operate. Zee (1980?) postulates that this neural integrator may be common to all conjugate eye movement systems.

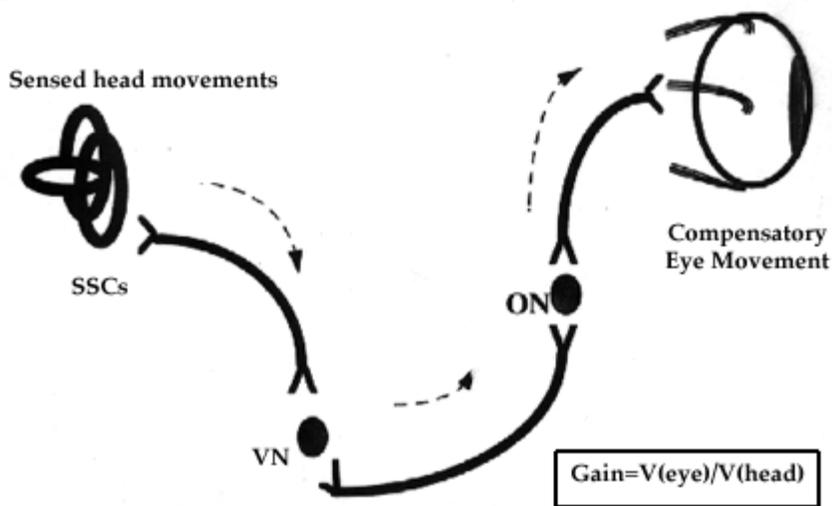


Figure 2: Simplified Schematic of the VOR Three-Arc Reflex

If the head undergoes sustained rotation in any direction, the eyes will exhibit a rhythmic oscillatory pattern called nystagmus. There are many different types of nystagmus, depending on what stimulus is involved. If head rotation occurs in a completely dark environment, the nystagmus is due purely to vestibular input and is called *vestibular nystagmus*. Nystagmus is an oscillatory pattern that is very characteristic of a "saw-tooth". For instance, if the head rotates in the horizontal plane to the right, the VOR will cause the eyes to compensate by moving in an equal but opposite direction to the left. This leftward movement of the eyes will continue until the eye gets near the edge of its orbit, then the eye will rapidly reverse direction, moving back across the center of gaze. This rapid reversal of eye movement in the direction of head rotation is called the *quick phase* of nystagmus. After the quick phase, the eye will again begin to compensate for the rotation by moving to the left. This slower movement in the opposite direction of head movement is called the *slow phase* of nystagmus. This slow phase is often what is measured in VOR research as in VOR response while the quick phase acts more like a correcting saccade.

Vestibular nystagmus does not continue endlessly. The response decays as the SCCs habituate to the constant (zero acceleration) rotation input. The actual time constant of the SCCs is approximately 5-7 sec., but central processes appear to 'extend' this time constant to approximately 25 seconds before the nystagmus ends and the eyes begin to float in space (Robinson, 1980). However, as will be discussed later, nystagmus due to constant rotation in lighted conditions will not decay, due to the added contribution of optokinetic (visual) input to the oculomotor system.

How is the VOR commonly measured? The first general requirement is that measurement take place in a dark room, so that there can be no contributing affects of visually-based image-stabilizing systems (i.e., optokinetic) (one exception to this requirement is a new technique used by Dr. Halmagyi (1995) called a 'head impulse'). A subject is then rotated either sinusoidally or continually in one direction. The movement of the eyes (vestibular nystagmus) is then recorded and the results are compared to the associated movements of the head. A common method of comparison involves the use of bode plots that look at VOR gain (defined as oppositely directed slow-phase eye velocity divided by head velocity) and VOR phase (phase angle between eye movements and head movements) as a function of frequency of head movement. If the VOR was completely compensatory, it would have a gain equal to 1.0 (unity) and no phase deviation.

Previous research has shown that the VOR acts as a band-pass filter. At very low frequency head movements (approximately 0.05 Hz or lower), the VOR compensation is poor, with low gain and a phase lead. At higher head movement frequencies (1 - 7 Hz) the VOR is very responsive with a gain close to 1.0 (unity) and eye movement in phase with head movement. It seems sensible that the VOR would operate well at these frequencies because they are the most likely to be encountered in everyday activities (head movement while walking is approximately 1 Hz, while running it is approximately 4-6 Hz). At still higher frequencies (above approximately 7-8 Hz), the system rolls off with VOR gain decreasing and phase lag increasing. Therefore, VOR responses are often characterized by their gain and phase relationship with the movement of the head at a particular frequency or frequency range.

An important factor that affects the VOR is the mental activity level of the subject. It has been shown (Robinson, 1980) that if the subject is not mentally active, the VOR gain, as tested in a completely darkened environment, is low. If asked to perform mental arithmetic during testing, the human VOR gain increases and averages approximately 0.65. Mental arithmetic involves alertness but no voluntary oculomotor activity. If however, the subject is asked to imagine and fixate on an imaginary spot on the wall in total darkness, the VOR gain increases to approximately 0.95 (this is sometimes called VOR-enhancement). Finally, if the subject is asked to imagine and fixate on a spot on the chair in which he/she is rotating, VOR gain drops to approximately 0.2 (this is often termed VOR-suppression). Apparently being alert is not enough, the subject must also be actively attending to the environment in a way that facilitates VOR functioning. Other factors that may lower the VOR indirectly by decreasing alertness include drugs, pain, and immobilization.

There has also been a differentiation in the literature between the VOR of a subject that is passively moved (Passive VOR) and the VOR of a subject that actively moves his/her head (Active VOR). Collewijn, Martins and Steinman (1983) demonstrated that Active VORs were more robust by 3%-13% than Passive VORs in almost all conditions. The variability of Active VORs was also less than Passive VORs. This could be due to the addition of efference copy (i.e., prediction). A further differentiation in VOR measurement is between Light VOR (sometimes termed Visual VOR) and Dark VOR. Light VOR entails measuring the VOR in a lighted environment and includes the influence of optokinetic stimulation (discussed below) while Dark VOR involves measuring the VOR in darkness and is considered more accurate in isolating the pure VOR response.

An issue that still is ambiguous concerns the effect of predictability of the stimulus on the VOR. It has been shown that the voluntary movement of the head can result in an improved VOR response (Active-VOR, Collewijn, et al., 1983). This is probably due to an efference copy of motor commands affecting the VOR three-arc reflex and is an indication that prediction does affect the VOR response. However, McKinley and Peterson (1985) demonstrated that VOR gain (in relaxed state, enhanced, or suppressed) is independent of whether or not a passive rotation stimulus was predictable or not. This seems to implicate efference copy as the effective method of prediction in the VOR process.

1.3 VOR Adaptation

As described earlier, the VOR is a very low latency reflex that allows the eyes to compensate for self-rotation of the head. An important aspect of the VOR, however, is its ability to change its gain in response to changing conditions that result in a mismatch between the current gain setting and that required to keep an image stabilized on the retina. This mismatch could be internally generated, due to the effects of age, disease, or trauma on the vestibular apparatus, circuitry, and/or eye muscles. Or the mismatch could be externally created by changing the relative movement of the visual scene in response to head movements, such as would occur when putting on prescription eye glasses or an

Collewijn, et. al. also tried to differentially adapt each eye to unequal gain change demands but found that it was not possible. If the unequal demands were close enough in size, the eyes would settle in on an intermediate adaptation level. If the discrepancy was large, however, the dominant eye's adaptation level won out. This also makes sense, as the VOR does function as a conjugate eye movement.

It seems that there can exist simultaneously at least two (and quite possibly more) different VOR gains that can be toggled into use as the appropriate situation arises. For instance, just putting on a scuba mask or a pair of glasses may be enough to toggle the VOR to a gain that is acceptable for that condition. Just the tactile feedback of putting on the goggles may be enough to toggle the adaptation in some cases (Viirre, 1995). However, it has not been completely determined whether these gain values are truly stored or if the adapting mechanism simply speeds up its adaptation to conditions that have occurred often enough in the past.

The above descriptions of the vestibular apparatus, the VOR and its characteristics, and VOR adaptation has hopefully provided insight into the nature of the VOR. However, the VOR does not act in isolation in attempting to keep images stabilized on the retina as the head rotates. There is a visually-based stabilizing mechanism also. This mechanism will be described next, along with its relationship to the VOR.

2.0 Relationship of the VOR to the Tracking Response of the Eye.

The VOR is only one of five major types of eye movements. The other four (optokinetic, saccade, smooth pursuit, and vergence) interact dynamically with the VOR to 1) bring visual targets onto the fovea, and 2) keep them there. Of these eye-movements types, optokinetic and smooth pursuit can be considered tracking eye movements, as they both use input signals for directing the eye that are derived from the retina itself, and their function is essentially one of image stabilization, not image capture. Therefore this discussion will include how both of these tracking mechanisms relate to the VOR.

2.1 Optokinetic Reflex (OKR)

While the VOR compensates for head movement by using input from the vestibular apparatus, the optokinetic reflex (OKR) works to maintain a stable image using visual input. The OKR system actually uses as input visual information coming from the entire retina (not just the fovea) to detect if an image slip is occurring. A slippage is manifest as an optical flow field moving across the retina. A common example of an OKR experience is a person peering out a large window on a moving train, watching the scenery pass by (Furness, 1981). If slippage is occurring, a corrective eye movement is generated to compensate for it by moving the eye with equal gain in the direction of the optical flow. Therefore, while both reflexes serve the same purpose and both can be considered involuntary, the VOR uses vestibular input to generate compensation commands while the OKR uses visual input to do the same. In fact, OKR can produce optokinetic nystagmus that appears similar to vestibular nystagmus discussed earlier, but while head rotation is required for vestibular nystagmus to occur, only a large field-of-view moving image is required to produce optokinetic nystagmus. Why would there need to be two separate systems to perform essentially the same image stabilizing task? It turns out that they both work in synergistic fashion to maximize the eye compensation response to any head movement.

First consider the fact that even under conditions of enhanced VOR, the gain of the resulting VOR averages approximately 0.95, not the required 1.0 to keep the eyes exactly stabilized in space. Yet the same VOR measured in the light does have a gain equal to 1.0 at most natural frequencies. Therefore, visual information provided by the lighted conditions allows for the correction of VOR residual error through some visual tracking mechanism (Peli, 1995). This visual tracking mechanism is in fact the OKR. Below is a description of how the VOR and OKR combine to provide optimum image stabilization performance.

The VOR is a very fast reflex that serves to compensate eye movements effectively for head movements at frequencies in the range of approximately 1-7 Hz, especially if the head movement is voluntary (allowing for efference copy). However, the VOR is less accurate at lower frequencies, especially those lower than 0.1 Hz, where the gain drops significantly and a phase lead appears. The OKR has the opposite performance characteristics. It has longer latency (due to the fact that it uses visual input) but at low frequencies (i.e., less than 0.1 Hz) it has near unity gain and no

phase. From 0.1 Hz to approximately 1 Hz, the OKR begins to lose gain and develop a phase lag due to increasing latency (Peterka, Black, & Schoenhoff, 1987). At higher frequencies it cannot effectively compensate due to its relatively long latency and low gain compared to the VOR. Therefore the combination of the two mechanisms allow for maximal image stabilization all the way from the lowest frequencies (governed mostly by OKR) to the highest frequencies (governed mostly by VOR).

There is another aspect of the VOR/OKR combination that contributes to improved performance over either system alone. This aspect is a timing issue; time of onset and time of offset. Earlier it was mentioned that the VOR has a very short latency (onset time) while the OKR has a longer latency. The VOR then allows for a faster reaction time even at lower frequencies. But it was also mentioned that the VOR will eventually decay during constant, zero-acceleration rotation due to the elasticity of the cupula. Although effectively extended through central processes, the time constant of pure vestibular nystagmus in humans is approximately 25 seconds. The OKR, however, has a long latency but no time constant as it does not decay with repeated stimulation of the retina by an optical flow. Therefore, as the VOR decays, the OKR is building up, creating a continual seamless stabilization of most images on the retina.

<concept of OKN counteracting inappropriate postrotatory vestibular nystagmus; direction of OKAN is opposite to that of perceived self-rotation - Zee 1980, Howard 1986>

<also, loss of peripheral labyrinthine function abolishes OKAN - Zee 1980?>

Finally, it was shown earlier that VOR adaptation requires visual input from the retina regarding image slip. This information is also used by the OKR. Therefore, it may be possible that OKR mechanisms serve to provide this input to the VOR adaptation process, further coupling the two reflexes.

So as discussed, the VOR and OKR have different relative contributions in the areas of latency and frequency range. Alone, neither reflex can account for all possible head movements, but together they provide an excellent means to keep an image stabilized on the retina across the entire spectrum of head movements. A descriptive model of this relationship is shown below (Figure 4).

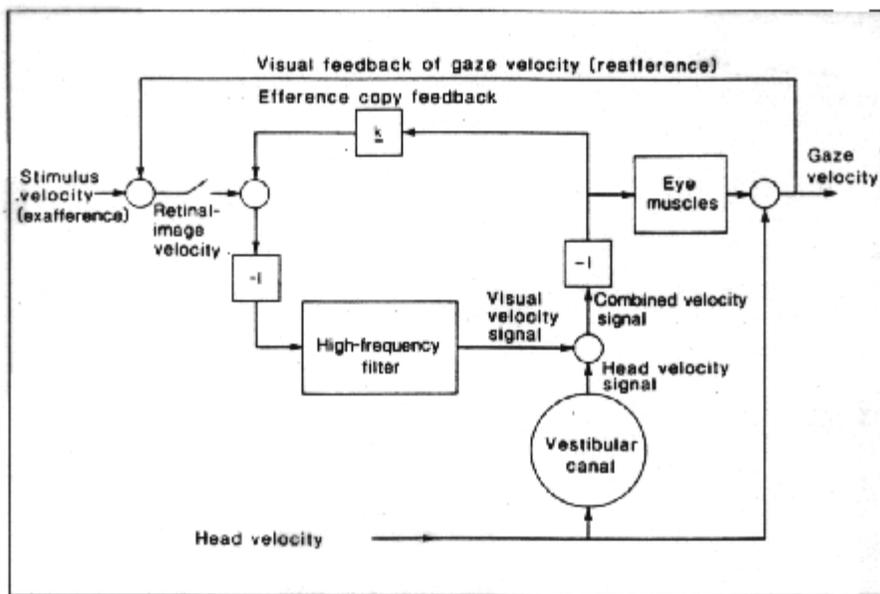


Figure 4: Descriptive Model of OKR/VOR Relationship (from Howard, 1986b)

This model accounts for how visual and vestibular signals cooperate to produce image-stabilizing eye movements (Howard 1986b; derived from an earlier model by Robinson, 1977). Starting at the left of the diagram, stimulus (target) velocity is combined with visual feedback on current eye velocity to generate a retinal image-slip velocity. This slip velocity is adjusted by efference copy (with gain k) which is an inner-feedback loop that provides prediction

information on the current eye movement plans. The resulting visually-mediated eye-command component is phase inverted and run through a high frequency filter so that the high frequency information is rejected (due to the lags/inaccuracies of the optokinetic system at higher frequencies). This final visual-velocity signal is then combined at the vestibular nucleus with the head-velocity signal coming directly from the SCCs, and the resulting combined-velocity signal is phase inverted (due to the compensatory nature of eye movements) and sent to the eye muscles through the oculomotor neurons to stabilize the image. This final combined signal is also the efference copy that is fed back to affect future eye movements. Thus this model has the visual and vestibular signals combining in simple linear-additive fashion at the vestibular nucleus. One addition that 'might' also be included is perhaps a low frequency filter (for frequencies below 0.05 Hz) to negate the poor low frequency response of the vestibular system.

2.2 Smooth Pursuit

Another type of eye-tracking response is accomplished by a smooth pursuit eye movement. Smooth pursuit, like OKR, uses visual input to generate eye movements, but instead of functioning to keep the entire retinal image from slipping, smooth pursuit moves the eye to keep only the image that is on the fovea stabilized (Robinson, 1980). Smooth pursuit uses visual information of the foveal target to calculate its velocity so that the eyes can move accordingly. As an example, following the movement of a rabbit darting in a field would require the rabbit to always remain stabilized on the fovea while the rest of the image (the field) slipped across the peripheral retina. Predictability of target motion will increase smooth pursuit performance and bandwidth (Robinson, 1980; Furness, 1981). Smooth pursuit is also often considered a voluntary eye movement (while OKR is involuntary) although some researchers (as cited by Furness, 1981) have questioned this.

So what would happen if smooth pursuit and VOR occurred at the same time? This could happen, for instance, if you rotated your head from side to side (or up and down) while following the darting rabbit in the field. Each eye movement type has different goals; the VOR strives to keep the eye stabilized in space while the smooth pursuit attempts only to have the moving target image remain on the fovea. Which would win out? Benson and Barnes (1978)(cited from Furness, 1981) showed that VOR suppression by the pursuit commands would occur at up to approximately 1 Hz. Therefore, up to a limit, the pursuit mechanism would cancel the VOR commands (either directly or through other central processes such as parametric changes in the VOR) and allow for the continued pursuit of the target. This makes intuitive sense to anyone who has ever tried to run down a fly ball in a baseball game while his/her head was undergoing rotational vibrations due to the run! Also, remember how subjects could lower the gain of their VOR during testing simply by fixating on an imaginary point on their rotating chair? This is the same VOR suppression, where the pseudo-pursuit task is to follow the point on the chair.

3.0 Speculations on the Visual and Vestibular System Contributions to Simulator Sickness

First a brief overview of the concept and characteristics of simulator sickness will be presented. Second the sensory conflict theory will be offered as a potential link between the visual and vestibular systems and simulator sickness, followed by a discussion of other possible contributions that these systems may offer to understanding the nature of simulator sickness. Lastly an annotated list will be presented of current-technology virtual interface artifacts that may also contribute to simulator sickness, along with the associated rationale.

3.1 Simulator Sickness

Simulator sickness has existed for as almost long as simulators, and although much has been written on the topic, there is still disagreement in the community as to a proper definition and whether or not it is truly different than motion sickness. However, simulator sickness is often used to refer to sickness symptoms that result from an incorrect presentation of a simulation, not sickness caused by a correct simulation of a nauseating experience (Pausch, Crea & Conway, 1992). Motion sickness is considered to include any sickness symptoms that result from the characteristics of actual physical movement. Therefore, motion sickness would require that the vestibular apparatus be stimulated, while there would be no such requirement for simulator sickness (Pausch, et al., 1992; Hettinger & Riccio, 1992). Basically, if the characteristics of a simulator do not match the perceptual requirements of the human, simulator sickness may occur in subjects using that simulator. If all aspects of the simulation are fixed to exactly match the human perceptual

requirements and this simulation is motion based, any further sickness symptoms would be termed motion sickness. However, a review of the literature will reveal that many researchers do not follow the above convention; many either interchange the two terms or simply call everything motion sickness. Therefore care must be taken when reviewing the motion sickness literature to determine what is actually being discussed.

There are many potential symptoms of simulator sickness. They include, but are not limited to, nausea, headaches, sleepiness, sweating, apathy, dizziness, general fatigue, eye strain, and loss of skin color (Ebenholtz, 1992; Pausch, et al., 1992). Another difference between simulator sickness and motion sickness seems to be in the proportion of individuals that develop certain symptoms (DiZio & Lackner, 1992). In general, motion sickness will result in many more instances of nausea/vomiting than simulator sickness. Simulator sickness is more likely to cause fatigue, eye strain, and headaches. Of course there is much overlap and a great deal depends on the specific platforms used in either case.

What causes simulator sickness? This appears to be the million dollar question. What makes it so difficult to decipher is that 1) not everyone that experiences the same simulation will get sick, 2) those that do will often have a variety of different symptoms, 3) many symptoms are internal, non-observable, and subjective, 4) symptoms can arise over a period of minutes to a period of hours, and 5) some individuals may get sick one day and be fine the next (McCauley & Sharkey, 1992; Griffin, 1990). Obviously there are many variables that influence simulator sickness and the resulting variation makes it difficult to acknowledge potential contributing factors without first mentioning a number of caveats. The goal would be to find major contributors to simulator sickness that can result in design guidelines to reduce many of the sickness symptoms that are experienced today.

3.2 Visual-Vestibular Conflicts and the Sensory Conflict Theory

One theory that has arisen to explain motion sickness is called the sensory conflict theory. This theory, developed by Reason and Brand (1975), states that motion sickness will occur if there is a conflict between visual, vestibular, and proprioceptive signals in response to a motion stimulus. This conflict would arise from expectations of signal groupings based upon past experiences. If the mismatch is great, a sensation of oscillopsia (a state in which observed objects appear to oscillate) may occur, along with other motion sickness symptoms.

This sensory conflict may be decomposed into two *categories*: intermodality (between the visual and vestibular systems) or intramodality (between the SCCs and otolith organs within the vestibular apparatus) (Griffin, 1990). Regardless of category, there are three *types* of conflict that could occur: 1) where both signals exist and provide contradictory information, 2) where signal `A' exists but signal `B' is absent, and 3) where signal `B' exists and signal `A' is absent. (`A' and `B' are arbitrary designations for the two signals that are in conflict). The theory claims that all situations that cause motion sickness can be fit into one of the two categories and one of the three conflict types, for a total of six *conditions*.

Although the sensory-conflict theory was developed for motion sickness, it would seem to apply fairly well to simulator sickness with a couple of caveats. First, simulator sickness may include other inaccurate aspects of simulation that do not fall neatly into one of the six categories described above. For instance, a poorly vented helmet-mounted display (HMD) or tracking bodysuit may cause sickness symptoms, although there is no obvious sensory conflict involved. Second, it is doubtful that the second category (SCC-otolith organ conflicts) play much of a role in simulator sickness. So when considering simulator sickness, it may be wise to only consider the visual-vestibular conflicts that arise.

For each of the remaining three conditions involving visual-vestibular conflict, examples can be found from the world of virtual reality and simulation. For Condition 1, where both visual and vestibular signals exist but are in conflict, one example would be an inaccurate or noisy head-position sensor on an HMD. Another example would be if the display optics involved created an image scene magnification or minification from the `real' world, such as would occur if a scene's geometric field-of-view did not match the display's field-of-view. For Condition 2, where a visual signal occurred in the absence of a vestibular signal, a common example would be a flight simulator on a fixed base platform. For Condition 3, where a vestibular signal occurred in the absence of a visual signal, examples would include a very low display update rate, or a head tracking sensor that was either limited to translation or rotational movement only.

This theory also seems viable when one considers the numerous research studies on the nature of VOR adaptation. These studies required mismatches between the visual scene and the VOR at onset to generate the mechanisms that change the VOR gain. In nearly all of these studies, when subjects first encountered the mismatch (or conflict), symptoms of simulator sickness occurred, including headaches, nausea, disorientation, sweating, etc. This would be a classic example of Condition 1, a visual-vestibular mismatch where both signals are present but contradictory.

Regardless of the sensory conflict theory, an important finding is that the existence of a working vestibular apparatus appears to be a necessary requirement for the experience of simulator or motion sickness. Labyrinthine defective individuals fail to exhibit symptoms even under the most extreme conditions and therefore are immune to simulator sickness (Ebenholtz, 1992; McCauley & Sharkey, 1992). However, the mere existence of a working vestibular system is not sufficient for motion sickness to occur, it is simply one necessary requirement.

It is also interesting to note that visual stimulation in the absence of vestibular input can also result in simulator sickness. Hettinger and Riccio (1992) found that individuals that do not experiencevection do not experience simulator sickness. However, not all individuals who experiencevection will get sick. Therefore, it appears that the ability to experiencevection is also a necessary but not sufficient requirement for simulator sickness.

The last two paragraphs implicate both the vestibular and optokinetic systems as necessary components of simulator sickness. These findings, combined with the explanatory ability of the sensory conflict theory in regards to motion sickness situations and the apparent relevance of a modified version of that theory for simulator sickness, definitely point to a strong contribution and interaction of both the vestibular and visual systems in the occurrence of simulator sickness.

Lastly, there is the question of post-effects. The sensory conflict theory implies that if there is a visual-vestibular conflict, motion sickness may result. If there is a conflict of this sort in a virtual environment, the resulting VOR will likely have a mismatched gain. As was discussed earlier, the VOR has a tremendous capability to adapt to relatively small gain conflicts in a short amount of time. During this time, however, there may be definite feelings of sickness, but once it is fully adapted the symptoms may disappear - for awhile. Once the subject exits the virtual environment and returns to the real world, his/her VOR may still be 'tuned' for the virtual world and need time to readapt. In the meantime balance, physical activity, and perception may be degraded (Ebenholtz, 1992). There is also the possibility that some aspect of the environment may at some future point trigger the VOR to momentarily 'switch back' to its VE gain level. These 'VOR flashbacks' could have serious implications for the well-being of the subject. Gauthier and Robinson (1975) actually provide evidence that this can occur. In testing the VOR of under-sea divers in the dark, these researchers found that either of two VOR gains could appear sporadically during the testing. Therefore, studies on simulator sickness need to consider post-effects as well as effects during the simulation.

3.3 Virtual Interface Artifacts that may Contribute to Simulator Sickness

Given that there are many potential factors that affect simulator sickness, there are accordingly many virtual interface artifacts that could contribute to this condition. Therefore, the list was pruned to include only those artifacts that may contribute to simulator sickness by directly affecting the vestibular and vestibular-visual channels. It should be noted that many of these artifacts will contribute to simulator sickness only if they are of sufficient quality to stimulate the reflex responses to visual motion. Very crude displays may not cause simulator sickness but they will not have much intrinsic value either (Griffin, 1990). Below is an annotated list of these major virtual interface artifacts that should be researched for their effects on the VOR and simulator sickness.

3.31 Time Delays/Update Rate

Scene shifting must occur within the bounds determined by the neurophysiology of the human visual system (Viirre, 1995). Earlier it was shown that the VOR has a latency on the order of 20 msec before the eyes begin to compensate for head movement. This creates a definite problem with current virtual interfaces. The time it takes for VR systems to record position information and update the visual scene can often be far slower than the ability of the human's central nervous system to detect position changes. For instance, the Polhemus Fastrak sensor alone has a latency of

approximately 100-250 msec. (Meyer, Applewhite & Biocca, 1992), and additional time must be added to update and render the new scene. This creates a mismatch between vestibular sensations/expectations and what is seen (sensory conflict theory) that can be very disconcerting and cause simulator sickness symptoms to develop. Also, unless the VOR adapts to develop a phase lag commensurate with the lag in the system, image slip may be noticed as the eyes first compensate for head movements, then the display updates itself. It is unclear if and to what extent the phase of the VOR can be modified by visual lags. Of course, much depends on the specific technology used (mechanical systems are very low latency) but the most popular VR systems do levy large time-lag artifacts.

Time lags are often stable/consistent within a specific virtual interface, being based upon hardware specifics or set processing requirements. However, cases may exist where the time lags can fluctuate more drastically, for instance when moving back and forth between a high-complexity visual scene and a low-complexity scene or when varying the distance between a tracking sensor and its transmitter. If the time delays are systematic and fixed, there is a possibility that the human perceptual system could adapt. If, however, the delays are more variable/inconsistent, it is unclear how the human will respond because it has been generally found that unpredictable stimuli are the hardest to adapt to. Both of these cases need to be examined further.

3.32 Tracking (Position) Inaccuracies

Another artifact of current virtual interfaces is the inaccuracy of the tracking sensors used. Some sensors are highly accurate but constrain the user's movements (e.g., mechanical) while others become less accurate the further the sensor gets from its transmitter (e.g., magnetic, ultrasonic, optical), and certain trackers can have their accuracy interfered with by objects/sounds in the environment (e.g., magnetic, ultrasonic, optical) (Meyer, et al. 1992). The result of these inaccuracies is a distorted visual image that can result in image blur during image stabilization attempts by the VOR. This also creates a visual-vestibular mismatch which opens the door to potential simulator sickness symptoms. Again, if the position inaccuracy is consistent, there is a chance that the VOR may adapt to correct for this mismatch. If the inaccuracies are variable (which is more likely the case), it is very unclear whether or not the VOR will adapt, and if so to what value. Also, the virtual interface may actually have different position accuracy's for different types of motion, i.e., rotation or translation. This further complicates the issue.

3.33 Display Field of View (FOV)

Display FOV is generally limited in current VR systems due to its tradeoff with resolution of a display. Most reasonably-priced HMDs on the market today have display FOVs of approximately 40 degrees horizontal by 30 degrees vertical, although some (e.g., Virtual I/O) sacrifice FOV for increased resolution while others (e.g., Division) sacrifice resolution for large FOVs (VR News, 1995). There is a continuing debate as to which aspect should be emphasized, the answer often being task dependent. Simulator sickness does seem to occur more often in large FOV displays, however, as these displays offer a more compelling display of motion (Pausch, et al., 1992).

One potential contributor to this debate that so far has received scant attention is the potential connection between the display FOV and VOR plasticity. Given that the VOR requires optokinetic slip signals to adapt its gain and that these visual signals usually are extracted from image slip over the entire retina due to the optical flow field, there may be a lessening of VOR adaptation at lower FOVs. This lessening may take the form of decreased gain adjustment or increased time to adaptation, or some combination of the two. A study comparing display FOV to VOR adaptation would provide new information to the FOV/resolution debate while potentially providing insight on the causes of simulator sickness.

3.34 Image Quality/Resolution

The corollary to the above argument would serve to consider the impact of image quality/resolution on VOR adaptation. Although there is less face validity for such an experiment, there may be aspects of image quality that affect the optical flow field and recording of image slip on the retina. If image slip detection is affected, OKR and VOR adaptation may also be distorted. It is important to remember, however, that increasing a display's resolution may mean decreasing the associated FOV and/or decreasing the update rate (increasing time lags).

3.35 Geometric FOV (GFOV)

A separate but related concept to display FOV is the geometric FOV (GFOV). The GFOV is the visual angle defined by the viewing frustum used for image *generation* (Danas, 1995). Essentially, when creating an image scene for display, the GFOV defines the viewing volume of the virtual space that will be rendered by computer graphics algorithms. Once it is rendered, it is mapped on to the display FOV, which is the visual angle subtended by physical display screen during image viewing. Therefore, GFOV is the visual angle used during image *creation* to determine which aspects of an image will be seen, and the display FOV is the set visual angle in which the images will be *displayed*. If GFOV matches display FOV, there is a 1:1 correspondence between image generation and image display. However, if GFOV is greater than the display FOV, there will be a perceptual minification of the image relative to its 'true' virtual state, and if GFOV is less than the display FOV the result will be a relative magnification of the scene.

Since software determines the GFOV and HMDs can have widely varying display FOVs, there is a potential for a mismatch between GFOV and display FOV with its resulting perceptual distortions. Danas (1995) discovered that these distortions can affect visual-auditory matching tasks in a virtual environment. Given the resulting perceptual minifications/magnifications that may occur, it seems reasonable that the VOR would need to adapt to these conditions, and the accompanying visual-vestibular conflicts could increase the likelihood of simulator sickness.

3.36 Helmet Optics/Weight

The HMD could affect the VOR and simulator sickness in a couple of ways. First, HMD weight can change the apparent mass of the head, requiring increased effort on the part of the user to make head movements and potentially affecting otolith signaling of tilt (DiZio & Lackner, 1992). This can create a conflict between the vestibular and proprioceptive system according to the sensory conflict theory, creating a potential for sickness symptoms to develop. Second, HMD optics may magnify or minify a visual scene, creating a mismatch between visual and vestibular signals along the lines of VOR adaptation studies. This has definitely been shown to be potentially disconcerting. Distortions and non-linearities in the optics can further degrade the image.

3.37 HMD Accommodation/Vergence cues

The VOR gain varies between 1 and 2 depending on the distance to the target of fixation. HMD close-in optical displays can create artificial vergence cues which could change one's VOR gain. Zee's chapter in Zabel, Howard 1986, Viirre 1986.

3.37 Augmented Displays

Finally, augmented displays offer an interesting situation. These displays allow the presentation of virtual graphics superimposed on a view of the real world. A common example of this type of display is the head-up-display (HUD) found in many aircraft cockpits and a few automobiles. The augmented display types being explored in virtual reality research are head-coupled devices that move with the user, with some computer graphics defined as world-stabilized (fixed in space) while others are head-stabilized (they move with the user's head so as to always remain in the same relative position on the display). There is a question as to which reference frame the VOR would 'tune'. If it remained tuned to the real world image, the VOR would not adapt and this would likely result in image slip for the world-stabilized and head-stabilized computer graphics. If, however, it adjusted its gain to match the world-stabilized computer graphics presentation, it would result in a potential real-world image slip due to system lags/update rates discussed earlier. If the VOR became adapted to the head-stabilized computer graphics, VOR gain would essentially be suppressed and the images of the real world and the world-stabilized computer graphics would slip at slightly different rates (due to time lags involved). It is unclear exactly how the VOR would react to such situations and what parameters (luminance, contrast level, scene complexity, etc.) may affect this reaction.

Although the above virtual interface artifacts highlight potential causes of simulator sickness through the vestibular and

vestibular-visual channels, it is obviously not a complete list. A more detailed review of the literature would serve to identify other areas of concern as well as what efforts have begun to explore the above identified issues.

References

- 1) Collewijn, H., Martins, A.J. & Steinman, R.M. (1983). *Compensatory eye movements during active and passive head movements: fast adaptation to changes in visual magnification*, *J. Physiol* **340**, 259-286.
- 2) Danas, E. (1995). *Mapping auditory space onto visual space*, unpublished masters thesis, University of Washington, Seattle, WA.
- 3) DiZio, P. & Lackner, J.R. (1992). *Spatial orientation, adaptation, and motion sickness in real and virtual environments*, *Presence* **1:3**, 319 - 328.
- 4) Ebenholtz, S. M. (1992). *Motion sickness and oculomotor systems in virtual environments*, *Presence* **1:3**, 302-305.
- 5) Furness, T. A. (1981). *The Effects of Whole-Body Vibration on the Perception of the Helmet-Mounted Display*, doctoral dissertation, University of Southampton.
- 6) Gauthier, G.M. & Robinson, D.A. (1975). *Adaptation of the vestibulo-ocular reflex to magnifying lenses*, *Brain Research* **92**, 331 - 335.
- 7) Goldberg, M.E., Eggers, H.M., & Gouras, P. (1991). *The ocular motor system*, In E.R. Kandel, J.H. Schwartz, & T.M. Jessell (Eds), *Principles of Neuroscience* (Third Edition), Appleton & Lange.
- 8) Griffin, M.J. (1990). *Handbook of Human Vibration*. Academic Press Limited, London.
- 9) Halmagyi, Dr., Presentation on the 3-D measurement of the VOR in normal and labyrinthine deficient subjects using the head impulse stimulus, UW Medical Center, 29 Nov, 1995.
- 10) Hettinger, L.J. & Riccio, G.E. (1992). *Visually induced motion sickness in virtual environments*, *Presence* **1:3**, 306-310.
- 11) Howard, I.P. (1986a). *The vestibular system*, In K.R. Boff, L. Kaufman & J.P. Thomas (Eds), *Handbook of Perception and Human Performance*. New York: John Wiley.
- 12) Howard, I.P. (1986b). *The perception of posture, self motion, and the visual vertical*, In K.R. Boff, L. Kaufman & J.P. Thomas (Eds), *Handbook of Perception and Human Performance*. New York: John Wiley.
- 13) Kelly, J.P. (1991). *The sense of balance*, In E.R. Kandel, J.H. Schwartz, & T.M. Jessell (Eds), *Principles of Neuroscience* (Third Edition), Appleton & Lange.
- 14) McCauley, M.E. & Sharkey, T.J. (1992). *Cybersickness: perception of self-motion in virtual environments*, *Presence* **1:3**, 311 -318.
- 15) McKinley, P.A. & Peterson, B.W. (1985). *Voluntary modulation of the vestibulo-ocular reflex in humans and its relation to smooth pursuit*, *Exp Brain Res.* **60**, 454 - 464.
- 16) Merfeld, D.M. (1995). *Modeling the vestibulo-ocular reflex of the squirrel monkey during eccentric rotation and roll tilt*, *Exp Brain Res.* **106**, 123 - 134.
- 17) Meyer, K., Applewhite, H.L. & Biocca, F.A. (1992). *A survey of position trackers*, *Presence* **1:2**, 173-200.
- 18) Peli, E. (1995). *Real vision and virtual reality*, *Optics and Photonics News*, Jul 95, 28-34.

- 19) Peterka, R.J., Black, F.O. & Schoenhoff, M.B. (1987). *Optokinetic and vestibulo-ocular reflex responses to an unpredictable stimulus*, *Aviat. Space Environ. Med.* **58:9** (Suppl), A180 - A185.
- 20) Robinson, D.A. (1980?). *Control of eye movements*, *The Handbook of Physiology - The Nervous System II*, 1275 - 1320.
- 21) Robinson, D.A. (1976). *Adaptive gain control of vestibulo-ocular reflex by the cerebellum*, *Journal of Neurophysiology* **39:5**, 954 - 969.
- 22) Viirre, E. (1995). *Virtual reality and the vestibular apparatus*. Unpublished documentation.
- 23) Viirre, E., Tweed, D., Milner, K. & Vilis, T. (1986). *A reexamination of the gain of the vestibulo-ocular reflex*, *Journal of Neurophysiology* **56:2**, 439 - 450.
- 24) VR News (1995). *Technology review: Head-mounted displays*, *VR News*, **4:4**, 20 - 27.