

Exhibition of Stochastic Resonance in Vestibular Tilt Motion Perception

Galvan-Garza R.C.^{a,b}, Clark T.K.^{a,b,c}, Mulavara A.P.^d, Oman C.M.^a

^aMan Vehicle Laboratory, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, ^bJenks Vestibular Physiology Laboratory, Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, MA 02114, ^cSmead Aerospace Engineering Sciences, 429 UCB University of Colorado, Boulder, CO, ^dKBRwyle, 2400 E NASA Pkwy, Houston, TX 77058

Corresponding Author

Raquel C. Galvan-Garza

rcgg@mit.alum.edu

Abstract

Background: Stochastic Resonance (SR) is a phenomenon broadly described as “noise benefit”. The application of subsensory electrical Stochastic Vestibular Stimulation (SVS) via electrodes behind each ear has been used to improve human balance and gait, but its effect on motion perception thresholds has not been examined.

Objective: This study investigated the capability of subsensory SVS to reduce vestibular motion perception thresholds in a manner consistent with a characteristic bell-shaped SR curve.

Methods: We measured upright, head-centered, roll tilt Direction Recognition (DR) thresholds in the dark in 12 human subjects with the application of wideband 0-30 Hz SVS ranging from ± 0 -

700 μ A. To conservatively assess if SR was exhibited, we compared the proportions of both subjective and statistical SR exhibition in our experimental data to proportions of SR exhibition in multiple simulation cases with varying underlying SR behavior. Analysis included individual and group statistics.

Results: As there is not an established mathematical definition, three humans subjectively judged that SR was exhibited in 78% of subjects. “Statistically significant SR exhibition”, which additionally required that a subject’s DR threshold with SVS be significantly lower than baseline (no SVS), was present in 50% of subjects. Both percentages were higher than simulations suggested could occur simply by chance. For SR exhibitors, defined by subjective or statistically significant criteria, the mean DR threshold improved by -30% and -39%, respectively. The largest individual improvement was -47%.

Conclusion: At least half of the subjects were better able to perceive passive body motion with the application of subsensory SVS. This study presents the first conclusive demonstration of SR in vestibular motion perception.

Introduction

Stochastic resonance is a phenomenon in which the response of a non-linear system to an input signal is benefited by the presence of a particular non-zero level of noise (for reviews see [1–5]). The signature of SR is a pseudo-bell shape performance curve with a peak in performance at some optimal noise level associated with optimal system output [3]. In its original conception, SR was defined as pertaining to bi-stable dynamical systems [6]. Later the concept of threshold SR or non-dynamical SR began to appear in the literature. We will not distinguish dynamical

from non-dynamical SR but will instead use SR to broadly describe the exhibition of “noise benefit” as suggested and discussed by McDonnell and Abbott [4].

SR has been demonstrated in a variety of human sensory systems, typically through psychophysical experiments. The application of stochastic noise to sensory input has been shown to improve visual contrast sensitivity and detection [7], letter recognition [8], perception of ambiguous figures [9], and visual depth perception [10]. SR has also been observed in human hearing [11, 12], and has been identified as an important component in cochlear coding strategy [13]. SR in tactile sensation has been demonstrated in response to weak mechanical stimuli in healthy individuals [14–17] and stroke survivors [17]. The application of subsensory mechanical noise to the feet has been shown to reduce sway in young and elderly subjects [18–20] and patients with diabetes and stroke [21], and gait variability in elderly fallers [22]. Similarly, mechanical noise applied to ankle muscles was shown to improve balance in patients with functional ankle instability [23]. Electrical noise applied to the tibial nerve improved tactile perception in the soles of the feet in older adults [24], and electrical noise applied to the back of the knee has also been shown to improve balance [25]. Further, the application of whole body vibration over many days was shown to be an effective therapy to improve postural stability in Parkinsonian patients [26].

Galvanic vestibular stimulation (GVS) is the application of electricity to the vestibular system. GVS typically refers to the use of *suprasensory* stimuli levels applied in order to disrupt normal body responses. The effects of constant level, pseudo-random or stochastic GVS have largely been examined through posture, balance, and gait [27–30]. This study uses *subsensory* frequency

band limited stochastic vestibular stimulation (henceforth referred to as SVS), with the goal of improving threshold vestibular responses via the hypothesized SR effect. SR exhibition due to the application of SVS has been investigated previously using action responses as performance metrics. SVS typically improves responses in only a fraction of subjects tested. For example, in healthy subjects, SVS was shown to improve balance stability in 52-67% of subjects, depending on the frequency content [31] and 49-71% of subjects, depending on the direction of interest [32]. Another study showed that locomotor stability was improved with SVS in 77% of subjects [33]. These SVS studies stated that SR occurred in an individual when performance measurements taken on an individual in a single session improved during nonzero SVS relative to when no SVS was applied. These studies and others also chose to characterize SVS improvement in terms group mean responses. For example, SVS improved group mean postural balance performance on an unstable compliant surface in Parkinsonian patients [34]. However, not every subject, considered individually, necessarily demonstrated the effect. In contrast, a study of patients with primary autonomic failure did present individual statistical analysis, showing that SVS *significantly* improved autonomic and motor responsiveness within 55% and 75% of subjects, respectively [35]. The last two studies mentioned were careful not to claim that SR was definitively exhibited, but that SR could possibly explain the improvements measured when SVS was applied.

Although SVS studies are consistent in that not every subject shows performance improvement with SVS, in the SVS literature there has not been clear consensus on what criteria should be used to determine whether SR was demonstrated or not. All human performance metrics have inherent variability. Factors such as differences in anatomy, electrode placement, and internal

sensory noise level may introduce variability in the size of the anticipated SR effect, with some tests possibly exhibiting no clear performance improvement at all.

Thus, our goals were to (1) determine for the first time if SVS can improve vestibular motion *perception*, which is believed to employ qualitatively different mechanisms than “action” (e.g. vestibular ocular reflex, vestibular spinal reflex, etc.) [36,37], and (2) investigate criteria for determining the presence of SR in experimental data sets, including within an individual subject. To do this we measured roll tilt vestibular direction recognition (DR) thresholds, an assessment of the smallest tilt motion direction that can be reliably perceived. We defined and then applied two criteria for SR identification – a relatively liberal subjective one and a more conservative statistical criterion. We also employed simulations to quantify the likelihood of false positives using these SR criteria. To our knowledge, this is the first study to demonstrate that vestibular perception is systematically and significantly enhanced due to the application of SVS and the phenomenon of SR, and to explore the implications of the criteria used to determine the presence of SR in individual subject responses.

Materials and Methods

Subjects

Fourteen participants were recruited for this experiment. All were healthy, with no known history of vestibular problems. Subjects signed informed consent forms and filled out a health screen questionnaire through the REDCap tool [38] hosted at the Massachusetts Eye and Ear Infirmary (MEEI). Two subjects were excluded from the study based on results of a

suprathreshold Galvanic Vestibular Stimulation (GVS) task, not discussed in this paper. One was hypersensitive to the GVS, reporting uncomfortable tingling at the electrode site, while the other reported unilateral motion sensation, inconsistent with the stimulus. The remaining 12 subjects (6 female) were 25.8 ± 2.7 (SD) years old. All testing took place in the Jenks Vestibular Physiology Laboratory (JVPL) at the MEEI.

Stimulation Application

The stimulation device and data acquisition system used in this research was obtained on loan from the Neuroscience Laboratory at NASA Johnson Space Center and has been previously described [31]. Bilateral bipolar electrical stimulation was generated using the controlled current stimulator with subject electrical isolation and delivered via leads connected to 5.1×10.2 cm UltraStim Electrodes (Axelgaard Manufacturing Co., LTD.) placed on the mastoid processes. Before electrodes were applied, the skin was prepared with Nuprep skin prep gel and cleaned with alcohol wipes. Although the electrodes included a layer of Multistick gel, an additional, generous layer of Signagel electrode gel (Parker Labs) was added for improved conductivity. Stimulation was applied only after impedance was measured to be less than $1 \text{ k}\Omega$.

The SVS signals, 0-30 Hz white noise, used in this experiment were identical to those used previously [31], except the initial and final ramps in current intensity were removed. The remaining 20.5 s stimulus period of the profile (zero mean ($\pm 1\%$) and RMS [$(30 \mu\text{A RMS}/100 \mu\text{A} (\pm 5\%))$]) was repeated continuously to achieve SVS for the duration of testing. The SVS level refers to the (\pm) peak of the signal over the profile duration (Table 1).

Direction Recognition Threshold Data Collection

Direction recognition (DR) thresholds were measured using an approach similar to that previously described [39–41]. Subjects were secured in a chair mounted on a Moog 6DOF2000E motion platform by a five-point harness and an adjustable head restraint. To minimize non-vestibular sensory cues, testing was done in total darkness and subjects wore headphones that played white noise during motions. Microphones enabled communication between the subject and operator. Each trial consisted of a 0.2 Hz (5 s duration) head-centered roll tilt motion made of a single cycle acceleration to either the left (left ear down) or right (right ear down). The direction was randomized on each trial [41,42]. Perception of this motion involves integration of cues from the semicircular canals and otoliths of the vestibular systems [43].

After each motion, subjects indicated the direction of their perceived motion by tapping either a left or right button. Following a pause of at least three seconds, the next trial began. The peak velocity of the motions varied using an adaptive staircase procedure, starting at 2 deg/s (5 deg tilt). Initially, the motion magnitudes were determined using a 2-down, 1-up staircase (i.e., two motion directions had to be correctly identified in a row in order for the next trial's magnitude to decrease). After the first incorrect response, the staircase switched to a 3-down, 1-up staircase with a 90% correct motion direction target [41,44]. While intuitive, each subject completed several practice trials to ensure familiarity with the task. Vestibular perceptual thresholds measured in the dark do not appear to improve with training [45]. 150 trials were completed for each individual DR threshold assessment, during which a level of SVS (± 0 , ± 200 , ± 300 , ± 500 , or ± 700 μA) was applied. Subjects experienced all five SVS levels in randomized order. The SVS was delivered continuously during testing including during the motion, subject response, and

delay prior to the next motion. None of our subjects reported sensations of motion, discomfort, or awareness of the electrical stimulation during the application of SVS at any level. Subjects were given short breaks within each test (i.e., SVS level) and were removed from the motion chair between tests. Total testing time, including setup and breaks, was approximately 3.5 hours with each 150 trial assessment taking about 35 min.

Direction Recognition Threshold Determination

Each DR threshold dataset was analyzed by fitting with a bias-reduced generalized linear model (MATLAB function `brglmfit.m` [41]). In this method, a cumulative Gaussian psychometric curve was fit, such that the mean and standard deviation correspond to the vestibular bias and “one-sigma” DR threshold, respectively [39]. The threshold, in this case, corresponds to our performance metric, and is the magnitude of roll tilt which can be correctly distinguished at a rate of 79.4% (i.e., lower thresholds correspond to better perceptual performance) [39]. Before determining the DR threshold, a lapse identification algorithm was applied to identify and remove trials in which subject reports were inconsistent with the other trials using a delta-deviance method (Clark and Merfeld, 2017). Conclusions were not different with or without lapse identification applied (See Supplementary Materials).

The standard error for each DR threshold estimate was calculated using a jackknife leave-one-out method [46]. To determine if one estimated DR threshold was statistically significantly lower than another, we conducted one tailed t-tests. The number of degrees of freedom for the t-test was appropriately estimated using a Satterthwaite approximation [47] as previously described [48]. Significance was set at $\alpha=0.05$.

Stochastic Resonance Curve Fitting

In an effort to better visually determine if SR was present in our experimental data, the plot of threshold vs. SVS level for each subject was fit with an equation previously developed to describe the general SR phenomenon [49,50]. This equation is discussed in [51] and in Supplementary Materials. Successful fits using this equation could not be accomplished for every subject's dataset when the data did not have the appropriate shape. These curve fits were used only as a visualization aid for subjective judgments of SR exhibition, and not as a formal mathematical representation of vestibular perceptual SR.

Definitions of SR Exhibition

Since there is not a formal mathematical definition of what sufficiently constitutes SR behavior as a function of SVS level, two alternative criteria were defined and used to determine whether SR was present in our SVS data. The first was a "subjective SR" criterion, requiring experienced human judges to evaluate plots of each individual's DR threshold data. The judges identified whether an SR effect was present, by looking for a characteristic bell-shaped SR trend in the data. DR threshold data plots showed mean and standard error and were overlaid with a theoretical SR curve, as shown in Figure 1 and described below. However, the judges were instructed to report their overall impression of SR, and not to necessarily require that the data strictly match the plotted theoretical curve, or that the mean and standard error necessarily indicate a statistically significant decrease in threshold due to the SVS. Our second criterion, which we call "statistical SR", utilized human judges in the same way, but was intended to be particularly conservative: it not only required that an individual's data be subjectively consistent

with the characteristic SR bell shape, but also required that the lowest threshold be significantly lower than the no SVS DR threshold of the individual subject.

Numerical simulations for SR criterion sensitivity assessment

We aimed to determine the sensitivity (% correct detection) of the two SR exhibition criteria used [52], and how sensitivity changed with the magnitude of the underlying SR. To do this, we performed Monte Carlo simulations of a modeled subject's responses to the threshold test protocol [41,44]. A more detailed description of the simulation process can be found in Supplementary Materials. In all simulations the underlying vestibular bias was set to zero for simplicity. As a baseline, we used an underlying DR threshold of 0.34 deg/s (0.85 deg tilt), the arithmetic mean DR threshold for the 10 SR curve fit subjects excluding S9 who had a relatively high threshold compared to the group. From this baseline, the underlying threshold was then varied as follows. We used three types of simulation cases each with different underlying SR magnitudes based on actual subject results shown in Figure 1. The cases were: 1) Strong underlying SR exhibition, 2) Weak underlying SR exhibition, and 3) No underlying SR exhibition. The SR curve representing strong SR was defined by the fit to S1's data and was characterized by a decrease in DR threshold from baseline to optimal of -46%. The Weak SR case was defined using the SR curve fit to S5's data with a decrease of -14.9%. In both SR cases, the optimal DR threshold was a 300 μ A. Case 3 was defined by a constant DR threshold, thus quantifying the inherent measurement variability in *estimated* DR thresholds. Case 3 demonstrated that multiple simulations (or repeated empirical tests) of the same *underlying* DR threshold yield different *estimated* DR thresholds due to measurement and perceptual variability. The simulated subject data were plotted in the same manner as the experimental data, including

SR curve fit (whether it matched a typical SR shape or not) and 95% confidence interval box. These data plots, along with the experimental data plots were de-identified, randomly grouped, and given to three judges (two were authors) who were asked to indicate if SR was exhibited in each plot (i.e., yes or no). The judges were familiar with the characteristic SR curve and were given a non-labeled, theoretical SR curve figure for reference. In total, 360 data plots were presented to the judges consisting of the 12 experimental subjects and 116 simulations of each of three SR cases: none, weak, and strong. This blinded judging process allowed us to control for the false positive rate at which judges might identify SR when in fact no underlying SR existed.

SR Repeatability Test

The repeatability of SR exhibition within an individual is important for practical SVS implementation in the future. Although not a focus of this paper, we re-tested subjects on a subsequent day (requiring reapplication of electrodes), measuring their DR thresholds again with 1) no SVS and 2) at their individual optimal SVS level from the first test normalized by their GVS threshold [51].

Results

Direction Recognition Thresholds with SVS

Vestibular perceptual SR exhibition was subjectively identified by the blinded judges in the majority of subjects. In these subjects, as SVS increased, the DR thresholds first decreased, reached a minimum value, and then increased to a value generally below the threshold at 0 μ A SVS level (i.e., the “baseline” threshold). 10/12 subject data sets had successful SR curve fits (overlaid in Figure 1). Of these subjects, 6 *also* had at least one SVS level with a statistically

significantly lower threshold compared to their baseline threshold (t-test, $p < 0.05$). Including all 10 SR fit subjects, the decrease in threshold due to the “optimal” level of SVS ranged from 0.02 deg/s (S6) to 0.41 deg/s (S9), with a percent change from baseline ranging from -5.4% (S6) to -47% (S11), and an average decrease of -30%. For the 6 subjects that also had statistically significantly lower thresholds with SVS, the smallest decrease was 0.10 deg/s (S2), the smallest percent change from baseline was -26% (S10), and the average percent change was -39%.

The group average improvements in DR thresholds with SVS application were also analyzed (Figure 2). With all 12 subjects included, and when the data was normalized by each individual’s baseline threshold, the two lowest average motion thresholds were achieved at the 500 and 300 μA SVS levels, with a normalized change from baseline of -16.7% and -16.0%, respectively (Figure 2A, solid grey line). There was a statistically significant decrease of 0.086 deg/s from the 0 to 300 μA SVS group means (paired $t(11)=2.530$, $p=0.014$), and of 0.082 deg/s between the 0 and 500 μA SVS means (paired $t(11)=2.377$, $p=0.019$). When we allowed the data sets to be shifted such that each individual’s minimum threshold occurred at the same normalized SVS level, the change was -25.6% between the group average baseline and minimum threshold (Figure 2B, solid grey circle). This corresponded to a statistically significant mean decrease of 0.118 deg/s (paired $t(11)=3.706$, $p=0.002$).

With the two non-SR fitted subjects (S7 and S8) removed from the dataset, the group optimal SVS level was 300 μA with a normalized change of -22.4% between baseline and optimal SVS (Figure 2A, black solid line), yielding a statistically significant mean decrease of 0.109 deg/s (paired $t(9)=3.005$, $p=0.008$). When allowing for varying optimal SVS levels by subject, the

normalized change between the average group baseline and optimal was -29.5% (Figure 2B, black filled circles) for a statistically significant decrease of 0.136 deg/s (paired $t(9)=3.856$, $p=0.002$).

Comparison of Judges Responses to Simulated and Experimental Data

When split by group (3 simulation groups and experimental group), as expected, all three judges identified subjective SR in a higher percentage of subjects in the Strong underlying SR case (mean=95.7%) than in the Weak underlying SR case (mean=54.3%) and the No underlying SR case (mean=31.3%) (Figure 3A). We note that even in simulations in which there was no underlying SR, our judges reported subjective SR (false positives) in 15.5 - 40.5% of cases. For the experimental data set, two of the three judges identified subjective SR in 83.3% of subjects (the same 10 subjects in both cases) while Judge 2 was more conservative and identified only 66.7% (8/12) of experimental subjects as exhibiting subjective SR. Although 10/12 subjects had “successful” SR curve fits, we will refer to the more conservative average judge vote of 78% from this point forward as the percentage of subjective SR exhibitors in our experimental group. This proportion of subjective SR exhibitors in the experimental group was statistically significantly higher than that of the simulation group with No underlying SR ($z=3.20$, $p<0.001$), demonstrating that the proportion of SR exhibition found experimentally was not likely to have occurred by chance. The proportion was higher than the Weak underlying SR simulation group but was not significantly different ($z=1.56$, $p=0.059$), and was statistically significantly lower than that in the Strong underlying SR group ($z=-2.49$, $p=0.006$).

When the criterion was made even more conservative by requiring “statistical SR”, again as expected, the percentage of simulated subjects identified as SR exhibitors was higher for the

Strong underlying SR case (mean=93.4%) than for the Weak underlying SR case (mean=36.2%) than for the No underlying SR case (mean=16.1%) (Figure 3B). Again, we highlight that even with no underlying SR, 16% of simulated subjects were identified as statistical SR exhibitors, meaning they exhibited subjective SR and also had a significantly lower than baseline threshold within their data set. For the experimental data set, statistical SR (subjective criteria + significant decrease) was found in 50% of subjects (the same 6/12 for all three judges). Using this unanimous judge vote, the experimental group proportion of statistical SR exhibitors (50%) was statistically significantly higher than that of the simulation group with No underlying SR ($z=2.83$, $p=0.002$), not different from the Weak underlying SR ($z=0.940$, $p=0.174$), and lower than the proportion in the Strong underlying SR group ($z=-4.72$, $p<0.001$). Thus the statistical SR seen in our experimental data was statistically unlikely to have occurred by chance.

Repeatability of SR exhibition

As a secondary aim, SR repeatability was evaluated using an abbreviated protocol on a separate test day. On reassessment, we found 4/10 subjects demonstrated lower thresholds with the individualized optimal level of SVS than with no SVS (statistically significantly lower for 3/10). Of the remaining six subjects deemed to not exhibit SR repeatability, five had similarly low DR thresholds at optimal SVS level upon reassessment (as expected), but had an unexpected decrease in their baseline DR threshold (0 μ A SVS).

Discussion

We demonstrated the exhibition of stochastic resonance in passive vestibular motion perception with the external application of subsensory electrical noise in both individual subject and

grouped responses. Subjective SR exhibition (as determined by three blinded judges) was shown in 78% of subjects, while statistical SR exhibition (defined by both subjective SR exhibition and a significantly lower than baseline minimum DR threshold) was found in 50% of subjects. SVS elicited an improvement in upright roll tilt DR thresholds with a maximum percent change of -47% within an individual and a mean of -30% and -39% for the subjective and statistical SR groups, respectively. In fact, even when including the two “non-exhibitors”, and making no assumptions about which SVS level was optimal for each individual subject, compared to baseline we found significant improvements in thresholds at 500 uA and 300 uA of -16.7% and -16.0%, respectively. These results indicate that with subsensory SVS, at least half of the subjects were better able to correctly perceive passive self-motion in the absence of visual cues. Further, the characteristic shape of the relationship between DR threshold and SVS level was consistent with SR theory (Figure 1).

Even for the numerical simulations Case 3 where there was absolutely no underlying SR effect, our judges (incorrectly) identified an average of 31% of subjects as subjective SR exhibitors, which resulted in 16% meeting the stricter statistical SR criteria. This latter percentage of false positives is noticeably higher than 5% (expected rate with $\alpha = 0.05$) because there were four threshold measurements (at nonzero SVS levels) that could each by chance be statistically lower than the baseline threshold. This false positive rate is also less than 18.5%, (expected rate accounting for four comparisons) because of the additional conservative requirement of subjectively matching the shape of a SR curve. This suggests that our statistical SR identification approach better controlled for false positives expected from the combination of multiple tests and subjective SR shape judgments. By comparing the blinded subjective judgments of experimental

vs. simulated datasets, we were able to rigorously demonstrate that the proportion of subjects exhibiting SR in our experiment (subjective and statistical) was significantly higher than that expected to occur by chance. Without comparing to simulations in which there is no SR, subjective judgments are likely to identify SR at a relatively high rate, even when it does not exist. We suggest that one or both of our SR identification approaches, including requirements on SR shape and within subject significant differences, be considered in future SR studies.

Although we tested a range of SVS levels to allow for possible inter-subject differences in the optimal level of SVS, it is intriguing that subjects did not vary a great deal in the level of SVS that was optimal in improving their DR threshold. Half of our subjects achieved their minimum measured DR threshold at 300 μ A SVS (82 μ A RMS) and 8/12 subjects had lower DR thresholds than baseline with 300 μ A SVS applied. We note this could be an artifact of the SVS levels that we assessed (i.e., an individual whose optimal SVS level was 900 μ A or 400 μ A could not be identified as such since these levels were not tested). Nonetheless, concerning practical use of SVS in the future, it is promising that a relatively small range of SVS (i.e, 300 to 500 μ A (82 to 136 μ A RMS)) was beneficial for most subjects, such that the SVS level may not need to be tuned specifically for each individual in order to see improvement in perceptual performance.

While we did not find SR to be consistently repeatable (3/10 subjects), with $\alpha = 0.05$ and a sample of 10 subjects, by chance we would expect only roughly half of a subject to reach statistical significance, and yet we observed three. The observed changes in baseline DR thresholds on the re-test day were unexpected considering previous work from the JVPL that found high repeatability of DR threshold measurements across days, weeks, and months

(unpublished data). We believe that potential effects of repeated SVS applications deserve further targeted investigation.

This experiment showed that SR is common in the context of upright roll tilt DR thresholds at 0.2 Hz. This motion type is known to involve integration of information from both the semicircular canal and otolith vestibular sensors. Due to the non-specific nature of externally applied SVS, it remains an open question as to whether vestibular perceptual SR is dependent on the involvement of the semicircular canals, otoliths, or both. Vestibular motion perception was measured using a relatively simple two-alternative forced choice procedure, however we emphasize that this process is not only a sensory task, but also involves a higher-level cognitive component of decision making. Where along the perception pathway SR is occurring remains unknown.

We found no correlations between SR exhibition or optimal SVS level and age, GVS threshold, or baseline DR threshold. However, among SR exhibitors, we found a statistically significant correlation between an individual's baseline motion threshold and the maximum improvement in threshold due to SVS ($R=0.892$, $p=0.001$). This suggests that subjects who were relatively poor performers in our sensory recognition task benefited relatively more from SVS. Roll tilt DR thresholds (at 0.2 Hz, identical to those in this study) have been found to be increased in people above the age of 40, and further, increased roll tilt thresholds correspond to an increased likelihood of failing the modified Romberg balance test, even accounting for age [53]. Failures on this balance test also correlate with increased likelihood of falls [54]. While speculative, our

results together with the findings of Bermudez Rey et al. [53] suggest that an SVS device may be particularly helpful for older people at risk of falling.

Acknowledgments

This study was supported by a NASA Space Technology Research Fellowship Grant # NNZ13AM68H and Universities Space Research Association (R.C. Galvan-Garza), and the National Space Biomedical Research Institute (NSBRI) Grant NCC9-58 (T.K. Clark). We thank Daniel Merfeld for his input and for allowing this study to take place in his laboratory, Faisal Karmali for his helpful comments, Jacob Bloomberg and Leia Stirling for their guidance, and our subjects for their participation. This study was part of a doctoral thesis [51].

References

- [1] Gammaitoni L, Marchesoni F, Menichella-Saetta E, Santucci S. Stochastic Resonance in Bistable Systems. *Phys Rev Lett* 1989;62:349–52.
- [2] Collins JJ, Priplata A a, Gravelle DC, Niemi J, Harry J, Lipsitz L a. Noise-enhanced human sensorimotor function. *IEEE Eng Med Biol Mag* 2003;22:76–83.
- [3] Moss F. Stochastic resonance and sensory information processing: a tutorial and review of application. *Clin Neurophysiol* 2004;115:267–81. doi:10.1016/j.clinph.2003.09.014.
- [4] McDonnell MD, Abbott D. What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology. *PLoS Comput Biol* 2009;5:e1000348. doi:10.1371/journal.pcbi.1000348.

- [5] Aihara T, Kitajo K, Nozaki D, Yamamoto Y. How does stochastic resonance work within the human brain? – Psychophysics of internal and external noise. *Chem Phys* 2010;375:616–24. doi:10.1016/j.chemphys.2010.04.027.
- [6] Nicolis C, Nicolis G. Stochastic aspects of climatic transitions- Additive fluctuations. *Tellus* 1981;33:225–34.
- [7] Simonotto E, Riani M, Seife C, Roberts M. Visual perception of stochastic resonance. *Phys Rev Lett* 1997;256:6–9.
- [8] Piana M, Canfora M, Riani M. Role of noise in image processing by the human perceptive system. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 2000;62:1104–9.
- [9] Riani M, Simonotto E. Stochastic Resonance in the Perceptual Interpretation of Ambiguous Figures: A Neural Network Model. *Phys Rev Lett* 1994;72:3120–3. doi:10.1097/01.NURSE.0000425876.00812.7d.
- [10] Ditzinger T, Stadler M, Strüber D, Kelso J a. Noise improves three-dimensional perception: stochastic resonance and other impacts of noise to the perception of autostereograms. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 2000;62:2566–75.
- [11] Zeng FG, Fu QJ, Morse R. Human hearing enhanced by noise. *Brain Res* 2000;869:251–5.
- [12] Jaramillo F, Wiesenfeld K. Mechano-electrical transduction assisted by Brownian motion: a role for noise in the auditory system. *Nat Neurosci* 1998;1:384–8. doi:10.1038/1597.
- [13] Morse R, Evans E. Enhancement of vowel coding for cochlear implants by addition of noise. *Nat Med* 1996;2:928–32.
- [14] Collins J, Imhoff T, Grigg P. Noise-enhanced tactile sensation. *Nature; Nat* 1996;383:770.
- [15] Collins J, Imhoff T, Grigg P. Noise-mediated enhancements and decrements in human

- tactile sensation. *Phys Rev E* 1997;56:923–6. doi:10.1103/PhysRevE.56.923.
- [16] Richardson K a., Imhoff TT, Grigg P, Collins JJ. Using electrical noise to enhance the ability of humans to detect subthreshold mechanical cutaneous stimuli. *Chaos* 1998;8:599–603. doi:10.1063/1.166341.
- [17] Enders LR, Hur P, Johnson MJ, Seo NJ. Remote vibrotactile noise improves light touch sensation in stroke survivors’ fingertips via stochastic resonance. *J Neuroeng Rehabil* 2013;10:105. doi:10.1186/1743-0003-10-105.
- [18] Priplata A, Niemi J, Salen M, Harry J, Lipsitz L, Collins J. Noise-Enhanced Human Balance Control. *Phys Rev Lett* 2002;89:238101. doi:10.1103/PhysRevLett.89.238101.
- [19] Priplata A a, Niemi JB, Harry JD, Lipsitz L a, Collins JJ. Vibrating insoles and balance control in elderly people. *Lancet* 2003;362:1123–4. doi:10.1016/S0140-6736(03)14470-4.
- [20] Dettmer M, Pourmoghaddam A, Lee B-C, Layne CS. Effects of aging and tactile stochastic resonance on postural performance and postural control in a sensory conflict task. *Somat Mot Res* 2015:1–8. doi:10.3109/08990220.2015.1004045.
- [21] Priplata A a, Patritti BL, Niemi JB, Hughes R, Gravelle DC, Lipsitz L a, et al. Noise-enhanced balance control in patients with diabetes and patients with stroke. *Ann Neurol* 2006;59:4–12. doi:10.1002/ana.20670.
- [22] Galica AM, Kang HG, Priplata AA, D’Andrea SE, Starobinets O V, Sorond FA, et al. Subsensory Vibrations to the Feet Reduce Gait Variability in Elderly Fallers. *Gait Posture* 2009;30:383–7. doi:10.1016/j.gaitpost.2009.07.005.Subsensory.
- [23] Ross SE, Linens SW, Wright CJ, Arnold BL. Customized noise-stimulation intensity for bipedal stability and unipedal balance deficits associated with functional ankle instability. *J Athl Train* 2013;48:463–70. doi:10.4085/1062-6050-48.3.12.

- [24] Breen PP, Serrador JM, O'Tuathail C, Quinlan LR, McIntosh C, Laighin G. Peripheral tactile sensory perception of older adults improved using subsensory electrical noise stimulation. *Med Eng Phys* 2016;38:822–5. doi:10.1016/j.medengphy.2016.05.015.
- [25] Gravelle DC, Laughton C a, Dhruv NT, Katdare KD, Niemi JB, Lipsitz L a, et al. Noise-enhanced balance control in older adults. *Neuroreport* 2002;13:1853–6.
- [26] Kaut O, Brenig D, Marek M, Allert N, Wüllner U. Postural Stability in Parkinson ' s Disease Patients Is Improved after Stochastic Resonance Therapy. *Hindawi* 2016;2016. doi:10.1155/2016/7948721.
- [27] Pavlik A, Inglis J, Lauk M, Oddsson L, Collins J. The Effects of Stochastic Galvanic Vestibular Stimulation on Human Postural Sway. *Exp Brain Res* 1999;124:273–80.
- [28] Scinicariello AP, Inglis JT, Collins JJ. The effects of stochastic monopolar galvanic vestibular stimulation on human postural sway. *J Vestib Res* 2003;12:77–85.
- [29] MacDougall HG, Moore ST, Curthoys IS, Black FO. Modeling postural instability with Galvanic vestibular stimulation. *Exp Brain Res* 2006;172:208–20. doi:10.1007/s00221-005-0329-y.
- [30] Moore ST, MacDougall HG, Peters BT, Bloomberg JJ, Curthoys IS, Cohen HS. Modeling locomotor dysfunction following spaceflight with Galvanic vestibular stimulation. *Exp Brain Res* 2006;174:647–59. doi:10.1007/s00221-006-0528-1.
- [31] Mulavara AP, Fiedler MJ, Kofman IS, Wood SJ, Serrador JM, Peters B, et al. Improving balance function using vestibular stochastic resonance: optimizing stimulus characteristics. *Exp Brain Res* 2011;210:303–12. doi:10.1007/s00221-011-2633-z.
- [32] Goel R, Kofman I, Jeevarajan J, De Dios Y, Cohen HS, Bloomberg JJ, et al. Using low levels of stochastic vestibular stimulation to improve balance function. *PLoS One*

- 2015;10:1–24. doi:10.1371/journal.pone.0136335.
- [33] Mulavara AP, Kaufman IS, De Dios YE, Miller C, T. PB, Goel R, et al. Using low levels of stochastic vestibular stimulation to improve locomotor stability. *Front Syst Neurosci* 2015;9:1–14. doi:10.1371/journal.pone.0136335.
- [34] Samoudi G, Jivegard M, Mulavara AP, Bergquist F. Effects of stochastic vestibular galvanic stimulation and LDOPA on balance and motor symptoms in patients with Parkinson's disease. *Brain Stimul* 2015;8:474–80. doi:10.1016/j.brs.2014.11.019.
- [35] Yamamoto Y, Struzik ZR, Soma R, Ohashi K, Kwak S. Noisy vestibular stimulation improves autonomic and motor responsiveness in central neurodegenerative disorders. *Ann Neurol* 2005;58:175–81.
- [36] Merfeld D, Park S. Vestibular Perception and Action Employ Qualitatively Different Mechanisms. I. Frequency Response of VOR and Perceptual Responses During Translation and Tilt. *J Neurophysiol* 2005;93:186–98. doi:10.1152/jn.00904.2004.
- [37] Merfeld DM, Park S, Gianna-Poulin C, Black FO, Wood S. Vestibular perception and action employ qualitatively different mechanisms. II. VOR and perceptual responses during combined Tilt&Translation. *J Neurophysiol* 2005;94:199–205. doi:10.1152/jn.00905.2004.
- [38] Harris P a., Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap) - A metadata driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81. doi:10.1016/j.jbi.2008.08.010.Research.
- [39] Merfeld DM. Signal detection theory and vestibular thresholds: I. Basic theory and practical considerations. *Exp Brain Res* 2011;210:389–405. doi:10.1007/s00221-011-

2557-7.

- [40] Lim K, Merfeld DM. Signal detection theory and vestibular perception: II. Fitting perceptual thresholds as a function of frequency. *Exp Brain Res* 2012;222:303–20. doi:10.1007/s00221-012-3217-2.
- [41] Chaudhuri SE, Merfeld DM. Signal detection theory and vestibular perception: III. Estimating unbiased fit parameters for psychometric functions. *Exp Brain Res* 2013;225:133–46. doi:10.1007/s00221-012-3354-7.
- [42] Grabherr L, Nicoucar K, Mast FW, Merfeld DM. Vestibular thresholds for yaw rotation about an earth-vertical axis as a function of frequency. *Exp Brain Res* 2008;186:677–81. doi:10.1007/s00221-008-1350-8.
- [43] Lim K, Karmali F, Keyvan N, Merfeld DM. Perceptual precision of passive body tilt is consistent with statistically optimal cue integration. *J Neurophysiol* 2017. doi:10.1152/jn.00073.2016.
- [44] Karmali F, Chaudhuri SE, Yi Y, Merfeld DM. Determining thresholds using adaptive procedures and psychometric fits: evaluating efficiency using theory, simulations, and human experiments. *Exp Brain Res* 2015;234:773–89. doi:10.1007/s00221-015-4501-8.
- [45] Hartmann M, Furrer S, Herzog MH, Merfeld DM, Mast FW. Self-motion perception training: thresholds improve in the light but not in the dark. *Exp Brain Res* 2013;226:231–40. doi:10.1007/s00221-013-3428-1.
- [46] Tukey JW. Bias and Confidence in Not-quite Large Samples. *Ann. Math. Stat.*, vol. 29, 1958, p. 614–23.
- [47] Satterthwaite FE. Synthesis of variance. *Psychometrika* 1941;6:309–16. doi:10.1007/BF02288586.

- [48] Johnson EG, Rust KF. Effective Degrees of Freedom for Variance Estimates from a Complex Sample Survey. *Am. Stat. Assoc. Annu. Meet.*, 1993.
- [49] Rouvas-Nicolis C, Nicolis G. Stochastic resonance. *Scholarpedia* 2007;2:1474.
doi:10.4249/scholarpedia.1474.
- [50] Gammaitoni L, Hänggi P, Jung P, Marchesoni F. Stochastic resonance. *Rev Mod Phys* 1998;70:223–87.
- [51] Galvan-Garza R. Enhancement of Perception with the Application of Stochastic Vestibular Stimulation. Massachusetts Institute of Technology, 2016.
- [52] Green DM, Swets JA. Signal detection theory and psychophysics. 1966.
doi:10.1901/jeab.1969.12-475.
- [53] Bermúdez Rey MC, Clark TK, Wang W, Leeder T, Bian Y, Merfeld DM. Vestibular Perceptual Thresholds Increase above the Age of 40. *Front Neurol* 2016;7:1–17.
doi:10.3389/fneur.2016.00162.
- [54] Agrawal Y, Carey JP, Della Santina CC, Schubert MC, Minor LB. Disorders of balance and vestibular function in US adults. *Arch Intern Med* 2009;169:938–44.
doi:10.1001/archinternmed.2009.66.

Table 1. Stochastic Vestibular Stimulation Details. All values are in μA .

\pm SVS level	Min	Max	Mean	Calculated RMS
200	-200	178	1.56	54.4
300	-258	300	-1.14	81.6
500	-500	481	4.12	136
700	-655	700	-4.41	190.4

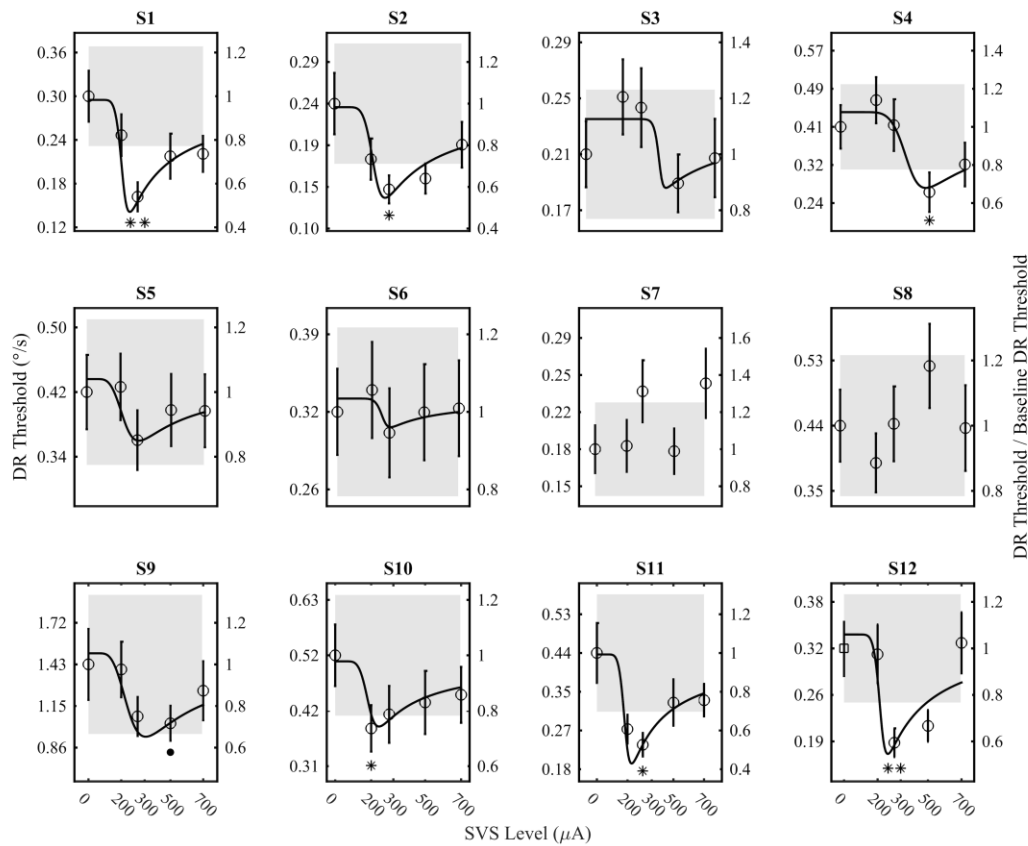


Figure 1. Upright roll tilt DR thresholds vs. SVS level. Left axis indicates the DR threshold in deg/s . Right axis indicates the threshold normalized by the baseline DR threshold (with 0 μA SVS). Error bars indicate standard error of the threshold estimate. Shaded areas indicate the 95% confidence interval of the baseline DR threshold (extended across the plot in a band to allow for comparison to the DR threshold at each SVS level). Overlaid line is the fit SR curve for each individual subject, when a successful fit was possible. S12's baseline DR threshold was measured on a separate test day and is therefore represented by a square data point. p-values refer to a statistically significantly lower DR threshold

compared to the baseline DR threshold. ** $p < 0.005$, * $p < 0.05$, • $p = 0.08$. The 348 simulation data sets also shown to the judges were presented in the same format.

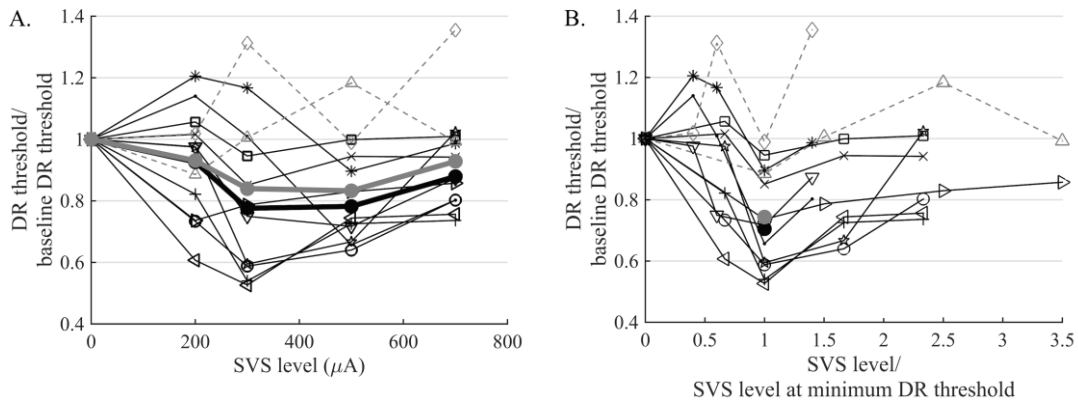


Figure 2. Group upright roll tilt DR thresholds. A: DR thresholds (open shapes connected by a thin line) for each individual subject normalized by their baseline threshold vs. SVS level delivered. B: DR thresholds normalized by baseline vs. SVS level normalized by each subject's personal optimal SVS level (i.e., that which resulted in the lowest threshold). In both panels A and B, grey dotted lines represent the two non SR curve fit subjects, S7 and S8. Thick black filled circles represent the average threshold in only the 10/12 SR curve fit subjects. Thick grey filled circles represent the average thresholds including all subjects.

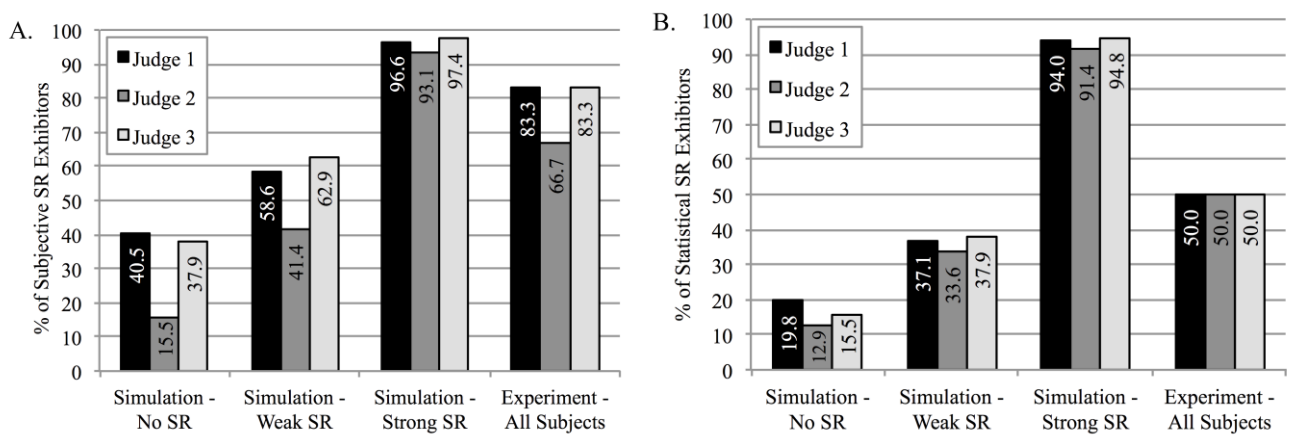
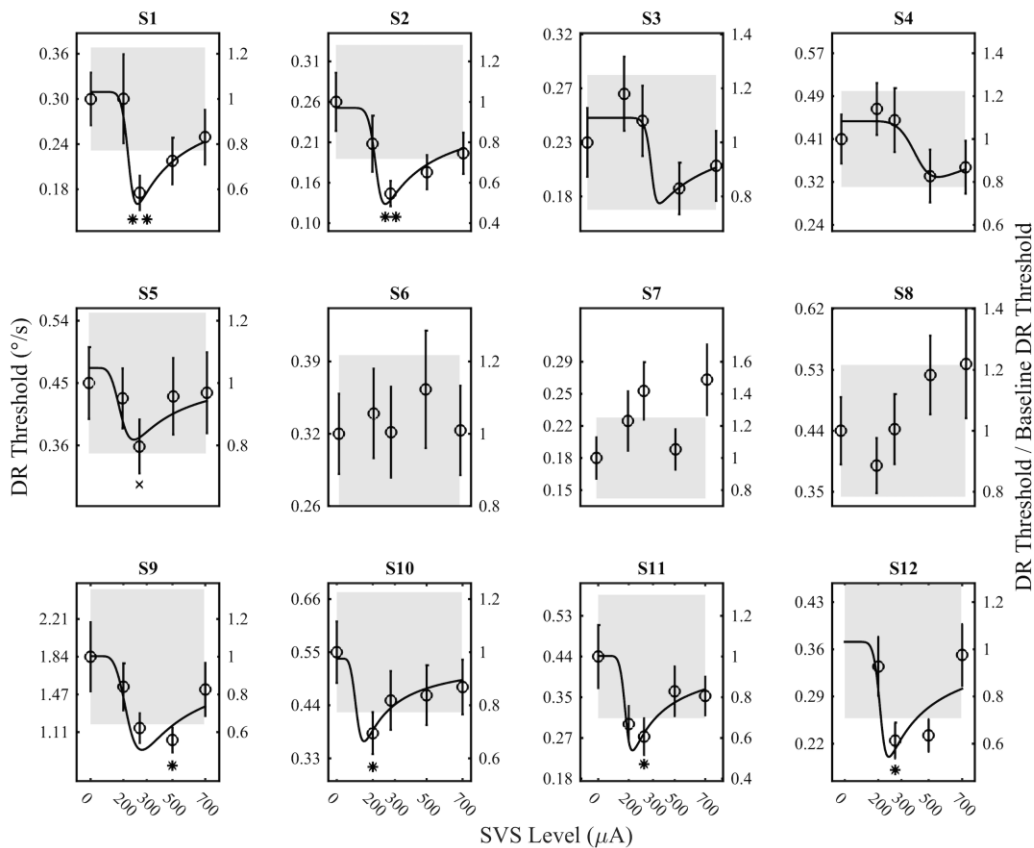


Figure 3. Percentage of subjects judged to be SR exhibitors for each of the three simulated cases and experimental data. A: Subjective SR. B: Statistical SR. N=116 in simulation groups and N=12 in experimental group.

Supplementary Material

DR thresholds estimated without lapse detection:

The figure below matches Figure 1 in the paper with the exception that the DR thresholds plotted were estimated without lapse detection applied. The resultant subplots differ slightly from those created when lapse detection was applied, but the overall conclusion is similar. S2 goes from significant at $p < 0.05$ (with lapse detection) to significant at $p < 0.005$ (without lapse detection), S4 loses statistical SR, S5 loses statistical SR ($p = 0.08$), S6 loses an SR fit, S9 goes from $p = 0.08$ to $p < 0.05$, S12 goes from $p < 0.005$ to $p < 0.05$. Overall, without lapse detection the ratio of subjects that exhibit subjective SR is 9/12 (lower than with lapse detection, 10/12) and the ratio that exhibit statistical SR is 6/12 (same as with lapse detection).



Supplemental Figure 1. Upright roll tilt DR thresholds without lapse detection applied vs. SVS level. Left axis indicates the DR threshold in deg/s. Right axis indicates the threshold normalized by the baseline DR threshold (with 0 µA SVS). Error bars indicate standard error of the threshold estimate. Shaded areas indicate the 95%

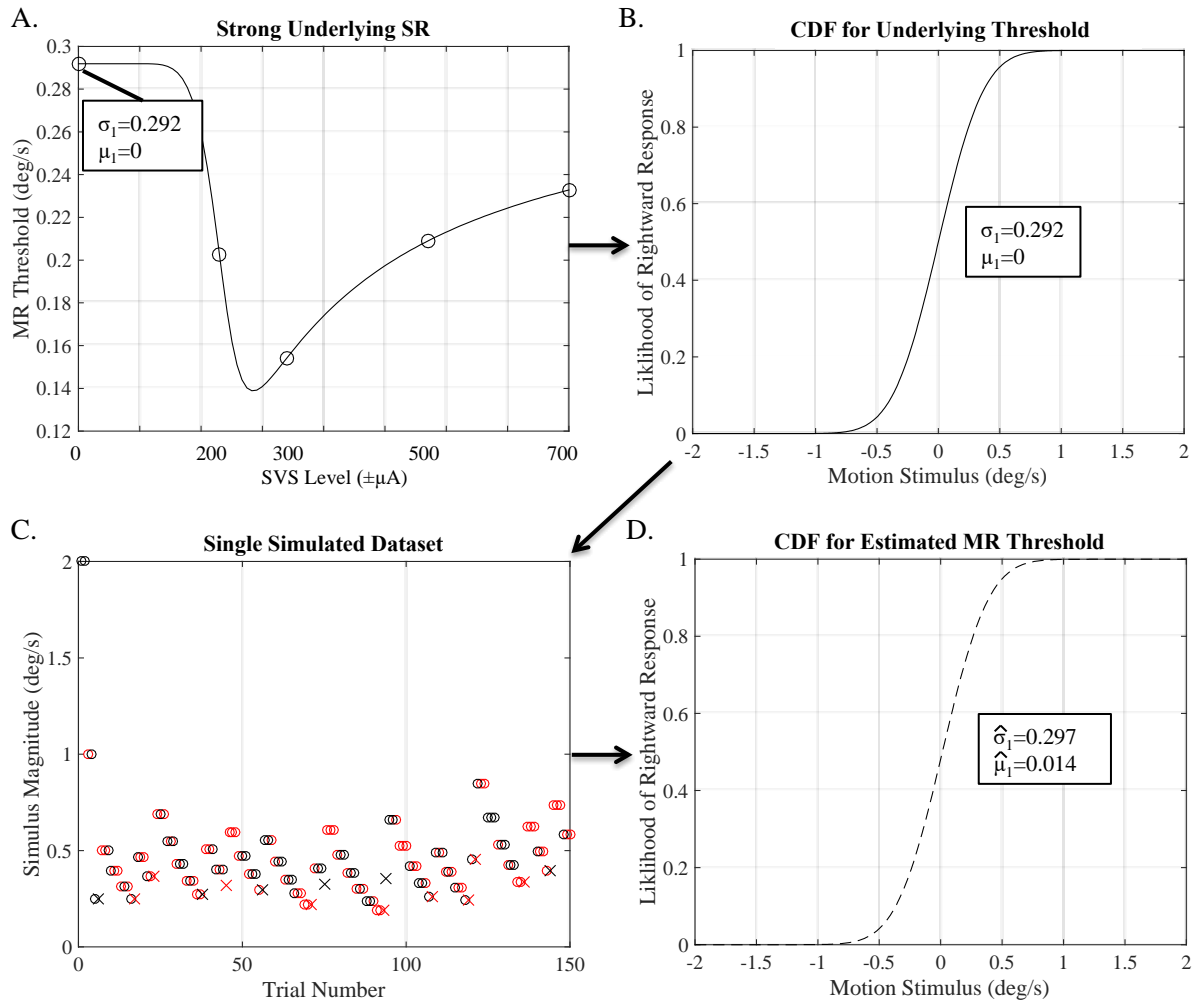
confidence interval of the baseline DR threshold (extended across the plot in a band to allow for comparison to the DR threshold at each SVS level). Overlaid line is the fit SR curve for each individual subject, when a successful fit was possible. p-values refer to a statistically significantly lower DR threshold compared to the baseline DR threshold. ** $p < 0.005$, * $p < 0.05$, ^x $p = 0.08$.

Stochastic Resonance equations fit to DR threshold data:

The equation fit to the data describes the amplitude of the periodic response of a Brownian particle in a bi-stable potential well system in the presence of noise and weak periodic forcing, $A = A_0 \frac{\lambda}{q^2} \frac{r(q^2)}{(4r(q^2) + \omega_0^2)^{1/2}}$, where $r(q^2)$ represents the Kramer's rate in the presence of Gaussian white noise with zero mean and strength equal to q^2 and is described by $r(q^2) = \frac{1}{\sqrt{2\pi}} \lambda \exp\left(-\frac{\lambda^2}{2q^2}\right)$. The additional system parameters include the amplitude of weak periodic forcing, A_0 , the frequency of weak periodic forcing, ω_0 , and the quartic potential parameter, λ (related to depth and spread of the potential wells). These three parameters were included as free parameters in the fit and an additional shifting parameter was included that allowed the baseline (0 uA SVS), or starting point, of the curve to vary across subjects. The equation was also made negative, corresponding to an optimal dip, rather than peak. The fit was done using MATLAB's `fminsearchbnd` function restricting the fit parameters to only positive fit values. We would like to point out that this equation does not act as a formal mathematical representation of our study task. The equation has 4 free parameters per subject dataset, which only includes 5 data points, suggesting a possible over fitting of the data. Additionally, we are not yet able to confidently interpret the fit parameter values and any mapping from theory to our study task remains speculative. For these reasons, we did not use a goodness of fit to this theoretical SR equation as a primary SR exhibition criterion in our study. This curve was fit and overlaid on the data as a visualization aid only. We acknowledge that a future derivation of a formal mathematical definition of SR exhibition applicable to human perception and performance could be extremely useful.

Direction Recognition Threshold Simulations:

A single simulated subject data set was constructed as follows. The SR curve from one of the three simulation cases was used to define the underlying motion threshold at each of the 5 SVS levels (set to be the same as those tested: $\pm 0, 200, 300, 500,$ and $700 \mu\text{A}$) (Supplemental Figure 2A). With the bias (μ) set to zero, each individual motion threshold (σ) was input to the experiment simulation. A cumulative distribution function (CDF('norm',X, μ, σ)) with X equal to the range of motion stimuli was then fully defined (Supplemental Figure 2B). One trial at a time, the simulation then calculates a subject response to a single motion stimulus (initial motion amplitude was set to 2 deg/s , matching the experiment). The subject response (left or right) and whether the response was correct or incorrect, was determined by calculating the probability of a rightward response for the specific motion stimuli (using the Gauss error function to solve the CDF) and then assessing whether a random number selection, from 0 to 1, falls above or below the calculated probability (Supplemental Figure 2C). This random element corresponds to human perceptual processing, modeled using signal detection theory (Green and Swets 1966) in that subjects will not always choose the same response even if given the identical motion stimuli many times. For example, although the probability for responding correctly may be high (e.g. 74%) for a relatively large motion stimulus, there is still a chance (e.g. 26%) that the subject's perception of the motion will result in an incorrect response. It is from this random component of the simulation that the output, estimated motion threshold (and bias) varies even when the same underlying motion threshold (and bias) is used as an input. In the simulation once the subject response was determined, the next motion stimulus was defined using the same adaptive 3-down, 1-up 90% target with an initial 2-down, 1-up phase method, as was used experimentally. This process was repeated until all 150 trials had been simulated, resulting in a set of 150 motion stimuli and their associated simulated subject responses for each underlying motion threshold.



Supplemental Figure 2. Example of the simulation steps for calculating an estimated DR threshold with strong underlying SR. This example is for the first threshold at 0 SVS. A: Strong underlying SR curve with the underlying thresholds for each SVS level (normalized by GVS threshold) shown. B: Cumulative Distribution Function (CDF) defined by the first underlying threshold with the bias set to zero. C: Simulated data set for the underlying threshold. Red and black circles represent correctly identified rightward motions and leftward motions, respectively. X's represent incorrectly identified motions. D: CDF fit to the simulated data defined by the estimated DR threshold $\hat{\sigma}_1$ and bias $\hat{\mu}_1$.

The data set was then fit using a bias-reduced glm fit (Chaudhuri and Merfeld 2013), again matching the experimental fitting approach, defining a new CDF for the estimated DR threshold (Supplemental Figure 2D).

Lapse detection and standard error calculations were done in the same manner as well and the result was an estimated direction recognition threshold at each SVS level.

Chaudhuri, S. E., & Merfeld, D. M. (2013). Signal detection theory and vestibular perception: III. Estimating unbiased fit parameters for psychometric functions. *Experimental Brain Research*. 225(1), 133–46.
doi:10.1007/s00221-012-3354-7

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley & Sons, Inc. New York.
doi:10.1901/jeab.1969.12-475