



Testing the ‘differences in virtual and physical head pose’ and ‘subjective vertical conflict’ accounts of cybersickness

Stephen Palmisano¹ · Lance Stephenson¹ · Rodney G. Davies¹ · Juno Kim² · Robert S. Allison^{3,4}

Received: 17 March 2023 / Accepted: 7 October 2023 / Published online: 17 January 2024
© The Author(s) 2024

Abstract

When we move our head while in virtual reality, display lag will generate differences in our virtual and physical head pose (known as DVP). While DVP are a major trigger for cybersickness, theories differ as to exactly how they constitute a provocative sensory conflict. Here, we test two competing theories: the *subjective vertical conflict theory* and the *DVP hypothesis*. Thirty-two HMD users made continuous, oscillatory head rotations in either pitch or yaw while viewing a large virtual room. Additional display lag was applied selectively to the simulation about the same, or an orthogonal, axis to the instructed head rotation (generating *Yaw-Lag + Yaw-Move*, *Yaw-Lag + Pitch-Move*, *Pitch-Lag + Yaw-Move*, and *Pitch-Lag + Pitch-Move* conditions). At the end of each trial: (1) participants rated their sickness severity and scene instability; and (2) their head tracking data were used to estimate DVP throughout the trial. Consistent with our *DVP hypothesis*, but contrary to *subjective vertical conflict theory*, *Yaw-Lag + Yaw-Move* conditions induced significant cybersickness, which was similar in magnitude to that in the *Pitch-Lag + Pitch-Move* conditions. When extra lag was added along the same axis as the instructed head movement, DVP was found to predict 73–76% of the variance in sickness severity (with measures of the spatial magnitude and the temporal dynamics of the DVP both contributing significantly). Ratings of scene instability were also found to predict sickness severity. Taken together, these findings suggest that: (1) cybersickness can be predicted from objective estimates of the DVP; and (2) provocative stimuli for this sickness can be identified from subjective reports of scene instability.

Keywords Head-mounted display · Motion sickness · Cybersickness · Virtual reality · Motion-to-photon latency · Sensory conflict

1 Introduction

Despite the enormous potential of head-mounted display (HMD)-based virtual reality (VR), cybersickness continues to limit its use (Rebenitsch and Owen 2016; Yildirim 2019). Unfortunately, this sickness appears to be worse than that produced by other types of virtual reality or simulation (Howarth and Costello 1997; Kim et al. 2014; Munafò et al.

2017; Sharples et al. 2008). For example, recent studies have shown that first-person videogames cause more severe sickness when played using HMDs (compared to desktop monitors or televisions—see Dennison et al. 2016; Martirosyan et al. 2021; Yildirim 2020). While none of the participants in these studies dropped out playing desktop versions of these games, many could not last even 10 min when playing in HMD VR. Indeed, it has now been shown that cybersickness can occur quite early during active HMD VR exposures (e.g. Clifton and Palmisano 2019; Palmisano and Constable 2022; Risi and Palmisano 2019; Teixeira and Palmisano 2021; Teixeira et al. 2022; Yildirim 2020). Increased sweating, nausea, burping, and stomach awareness are all quite common experiences when wearing HMDs, in addition to user disorientation and increased oculomotor discomfort (Gavgani et al. 2017; LaViola 2000; McCauley and Sharkey 1992; Stanney et al. 1998a, b; Palmisano and Constable 2022; Teixeira et al. 2022). Unfortunately, these unpleasant symptoms often persist long after the VR simulation

✉ Stephen Palmisano
stephenp@uow.edu.au

¹ School of Psychology, University of Wollongong, Wollongong, NSW 2522, Australia
² School of Optometry and Vision Science, University of New South Wales, Sydney, NSW, Australia
³ Centre for Vision Research, York University, Toronto, ON, Canada
⁴ Department of Electrical Engineering and Computer Science, York University, Toronto, ON, Canada

has stopped and the HMD has been removed (Kennedy and Lilienthal 1994; Kennedy et al. 1994; Merhi et al. 2007; Teixeira et al. 2022). It is therefore imperative to find effective ways to mitigate, or even better avoid, this cybersickness if we are going to unleash the full potential of HMD VR.

Display lag (also referred to as motion-to-photon latency) is known to play a major role in the cybersickness experienced with HMDs (Golding 2016; Howarth and Finch 1999; Kinsella et al. 2016). This lag is the time taken for the tracked user's head movements to produce compensatory visual changes on their HMD. Many studies have examined the effects of display lag on cybersickness during HMD VR (see Palmisano et al. 2020 for a recent review of this literature). They typically injected additional constant or time-varying display lags into their VR simulations, while their HMD users were performing some sort of physical activity (e.g. making repetitive head movements, or carrying out a virtual visual search task). Researchers varied the magnitude (and sometimes other characteristics) of this added display lag from trial to trial. Most studies found that the extra display lag increased the likelihood and severity of cybersickness (e.g. Caserman et al. 2019; DiZio and Lackner 1997; Feng et al. 2019; Jennings et al. 2000, 2004; Kim et al. 2020; Kinsella et al. 2016; Palmisano et al. 2019, 2020, 2023; St. Pierre et al. 2015; Stauffert et al. 2018). However, a few of these studies failed to find any significant display lag effects on cybersickness (Draper et al. 2001; Moss and Muth 2011; Moss et al. 2011). The latter null findings suggest that we cannot understand cybersickness based simply on display lag alone. Recently, we have argued that information about the HMD user's head movements is also important (i.e. in addition to the system's display lag).

1.1 The DVP hypothesis for cybersickness

While display lag is not noticeable when HMD users are stationary, it often results in obvious conflicts between their visual and non-visual information during physical movement. During head movements, the user's visually displayed head orientation in the HMD will sometimes trail, and other times lead, their head's actual physical orientation. We have previously argued that these *differences in virtual and physical head pose* (or *DVP* for short) are key to understanding cybersickness during active HMD VR (see Kim et al. 2020; Palmisano et al. 2020). This is because the DVP (at any instant) depends on the user's head velocity as well as the magnitude of the VR system's display lag. Our *DVP hypothesis* proposes that cybersickness is caused by large amplitude, time-varying patterns of DVP. Thus, according to this hypothesis, cybersickness should increase when: (1) additional (constant or time-varying) lag is injected into the system; and (2) the user's head-velocity increases (as both manipulations would be expected to generate more provocative patterns of DVP). These indirect

predictions of the theory have recently been supported by the findings of Feng et al. (2019), Kim et al. (2020) and Palmisano et al. (2019). Other studies have attempted to test the *DVP hypothesis* directly by estimating the DVP produced by their different HMD VR conditions (Kim et al. 2020; Palmisano et al. 2020, 2023). In these studies, we had participants make continuous oscillatory head movements (at 0.5 or 1.0 Hz and in either yaw, pitch or roll) under different levels of imposed constant lag (ranging from 0 to 222 ms on top of the system's ~4 ms baseline lag). We then estimated the DVP experienced on each trial based on our knowledge of its display lag and the participant's head movements. In the first of these studies, Kim et al. (2020) found significant positive relationships between their participants' sickness severity ratings and the mean of their DVP during active HMD VR trials. Subsequently, Palmisano et al. (2020, 2023) found that sickness severity was also predicted by the peak and the standard deviation of the DVP. These 3 DVP summary measures were able to reliably predict cybersickness in HMD VR during all types (yaw, pitch and roll) and speeds (0.5 Hz or 1.0 Hz) of user head movement tested.

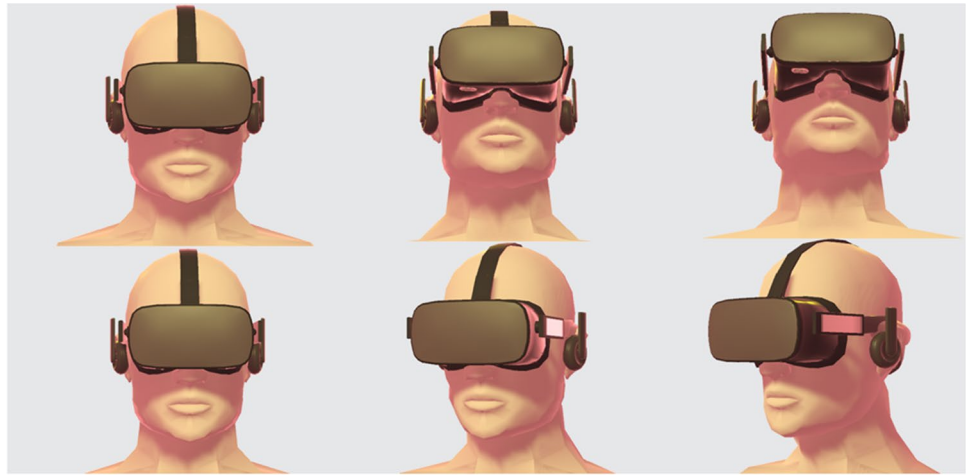
When HMD users move their heads during laggy VR, they often consciously perceive the DVP generated as *scene instability* [i.e. their virtual world will appear to swim and oscillate around them; effects sometimes referred to as 'oscillopsia'—see Allison et al. (2001)]. Like their DVP, these perceptions of scene instability tend to increase with head movement speed and the amount of imposed display lag (e.g. Allison et al. 2001; Kim et al. 2020, 2021; Palmisano et al. 2019). Thus, it has also been proposed that consciously perceived scene instability might provide confirmatory evidence of provocative patterns of DVP (Kim et al. 2020). Although other findings suggest that time-varying DVP might still be problematic when these differences fail to reach the threshold for conscious detection (e.g. Stauffert et al. 2018).

The above findings appear to be consistent with the predictions of our *DVP hypothesis*. However, this hypothesis is focussed only on *sensory input conflicts* (i.e. discrepancies in the visual and non-visual information about head orientation arriving at the HMD user's sense organs). It does not attempt to model any of the subsequent sensory or neural processing that is generated by the DVP. As it is based only on directly observable factors, this makes the *DVP hypothesis* straightforward to test. However, it is possible that we may not be able to fully understand cybersickness in HMD VR based on DVP alone.

1.2 Neural mismatch and subjective vertical conflict theories of cybersickness

Some theorists argue that we actually need to model the neural activity generated by the person's multisensory

Fig. 1 An HMD user making head movements in pitch (top row) and yaw (bottom row). On the trials in this study, participants made continuous oscillatory (up-down or left-right) head movements for 35 s



stimulation in order to understand their experience of motion sickness (e.g. Reason 1978; Oman 1982, 1990; Bles et al. 1998). According to Reason (1978), motion sickness only occurs when the sensory input conflict causes a specific type of *neural mismatch*, referred to as an *expectancy violation*. These neural mismatches can arise whenever our current sensory information differs from what we were expecting to receive based on our past experience and planned movement. Reason's *neural mismatch model* therefore explains why cybersickness often decreases with repeated exposure to the same (initially provocative) HMD VR simulation [because our expectations change during each subsequent exposure, the same DVP tends to result in less neural mismatch each time—see Palmisano and Constable (2022)]. The model has, however, also faced substantial criticism (e.g. Ebenholtz et al. 1994; Keshavarz et al. 2014; Lackner and DiZio 2020; Stoffregen and Riccio 1991; Stoffregen et al. 2017). According to Stoffregen et al. (2017), one of the theory's major problems is that scientists “cannot know ... the history of an individual's interactions with the environment” and thus, they “cannot know ... what patterns of intersensory stimulation might be expected” (p. 4). As a result, they argue that the *neural mismatch model* cannot be used to make quantitative predictions about future experiences of motion sickness (because it does not have access to the necessary information about our past exposures to similar motion scenarios). Others have argued that it is also difficult for the model to make qualitative predictions (e.g. Ebenholtz et al. 1994; Keshavarz et al. 2014). Because the range of potentially provocative conflicts is so wide according to the model, they argue that it is difficult to devise experiments that would falsify it (see Keshavarz et al. 2014, p. 654).

The *subjective vertical conflict theory* of motion sickness was created to address some of these criticisms of Reason's *neural mismatch model* (Bles et al. 1998, 2000; Bos and Bles 1998, 2002; Bos et al. 2008; de Graaf et al. 1998). While the theory still relies on a hypothetical internal model of sensory

and neural processing, it does provide clear, testable criteria for distinguishing provocative (from non-provocative) conflict situations. According to the theory, “all situations which provoke motion sickness are characterised by a condition in which the *sensed vertical* ... is at variance with the *subjective vertical* as predicted on the basis of previous experience” (Bles et al. 1998, pp. 481–482). This predicts that cybersickness due to DVP should be less likely and severe when upright HMD users make yaw, compared to pitch (or roll), head movements (see Fig. 1). This is because pitch (or roll) head movements: (1) tilt the user's head away from alignment with gravity (whereas yaw head movements do not); and (2) generate patterns of DVP that are more likely to result in misperceptions of the direction of gravity. When some lag is present, yaw, pitch and roll head rotations should all generate DVP in the user—with the largest amplitude DVP in each case occurring about the primary axis of the movement. However, during yaw head movements, most of the DVP will be orthogonal to the direction of gravity. Thus, according to *subjective vertical conflict theory*, this yaw DVP should not be a problem. Only roll and pitch DVP should induce significant cybersickness.

Support for the *subjective vertical conflict theory* of motion sickness has thus far been mixed. Consistent with the theory, motion sickness does appear to be more common when participants move their heads away from alignment with gravity (e.g. Lackner and DiZio 2006; Thornton and Bonato 2013; Chen et al. 2016). Also, consistent with the theory, studies examining stationary upright observers have found that motion sickness increases when the axis of their visually simulated self-motion is tilted away from alignment with gravity (e.g. Bubka and Bonato 2003). However, contrary to its predictions, physically upright observers still appear to experience motion sickness during visually induced illusory self-rotations in yaw [when there should be no subjective vertical conflict—see Bonato et al. (2005), Nooij et al. (2017)]. Also, contrary to its predictions, display

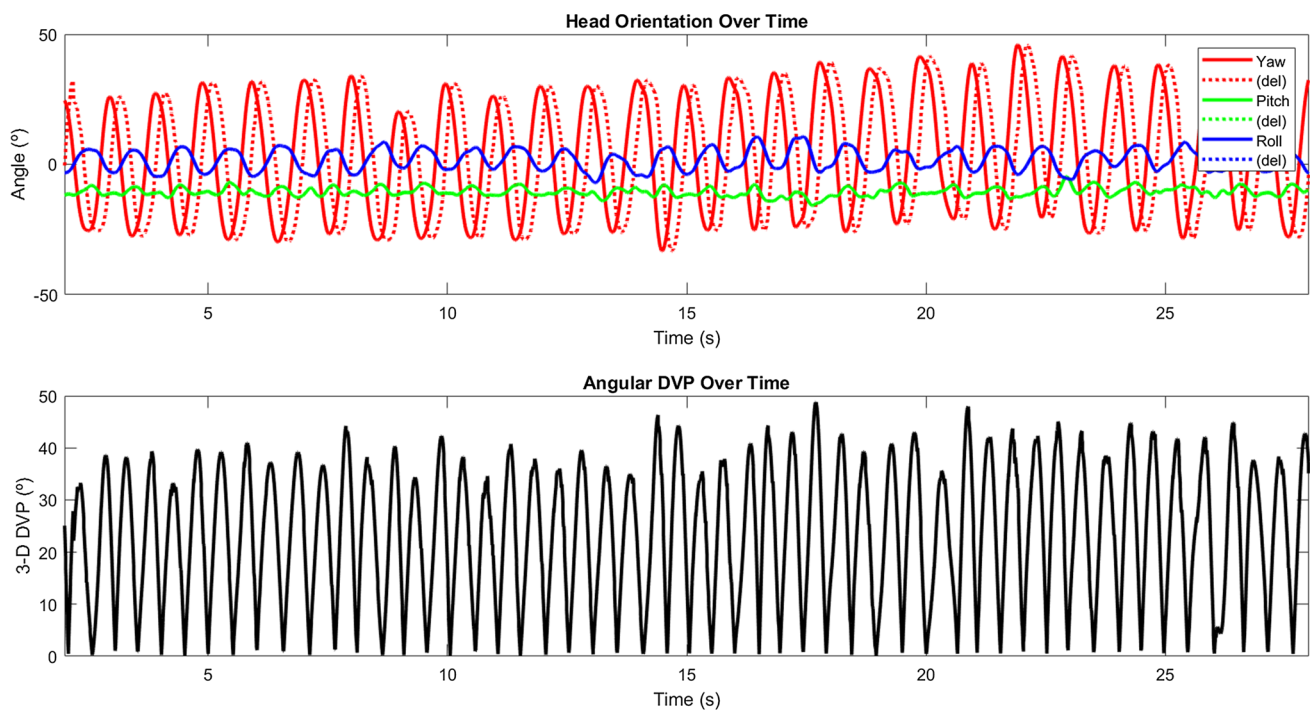


Fig. 2 (Top) This plot shows one participant's physical yaw, pitch and roll head orientations (in degrees) during a *Yaw-Lag + Yaw-Move* trial in the current study (see the red, green and blue solid lines, respectively). P2 was instructed to move their head in yaw at 1.0 Hz and the visual consequences of their yaw (but not pitch and roll) head movements were delayed by an extra 222 ms in the HMD. P2's estimated virtual head orientations during the trial are shown as dotted lines in

that same plot. Only the red dotted line is visible—since only yaw head movements were delayed. (Bottom) This shows the unsigned differences between P2's physical and virtual head orientations over time (also in degrees). At each instant this DVP was calculated as the absolute angle of the smallest rotation from the quaternion specifying their actual head pose to the quaternion specifying their virtual head pose (color figure online)

lag effects on cybersickness do not appear to be reduced during self-generated yaw, compared to pitch or roll, head movements in HMD VR (Feng et al. 2019; Kim et al. 2020; Palmisano et al. 2019, 2020, 2023). However, the designs of these more recent HMD VR studies were not ideal for testing the *subjective vertical conflict theory*. In the next section, we discuss some potential difficulties making conclusions about this theory based on data from these HMD VR studies.

1.3 Limitations of past cybersickness studies

In our recent studies, we instructed upright HMD users to make continuous oscillatory head movements—in either yaw, pitch or roll—while extra display lag was injected into the system (Feng et al. 2019; Kim et al. 2020; Palmisano et al. 2019, 2023). Our three earlier studies examined cybersickness during either self-generated yaw, or self-generated pitch, head movements—not both (Feng et al. 2019; Kim et al. 2020; Palmisano et al. 2019). Cross-experimental comparisons suggested that display lag effects on cybersickness were similar during the yaw and pitch head movement instruction conditions of these studies. However, this was not a strong test of the *subjective vertical conflict theory* because these studies

each tested different groups of participants in different virtual environments. In our most recent study (Palmisano et al. 2023), we directly compared the effects of display lag and head speed on cybersickness during yaw, pitch and roll head movements. Importantly, this study presented all three head movement instruction conditions to the same group of participants in the same virtual environment. In apparent contradiction to the predictions of *subjective vertical conflict theory*, we again found that cybersickness: (1) still occurred during yaw head movement conditions; (2) increased in severity with display lag magnitude during all three head movement conditions; and (3) was roughly similar (on average) across the yaw, pitch and roll head movement conditions. While this Palmisano et al. (2023) study was a clear improvement on our previous HMD VR studies, it was still not ideal for testing the predictions of *subjective vertical conflict theory* (e.g. that participants should not experience sickness during *pure* yaw rotations). Participants in that study were instructed to rotate their heads exclusively around a single axis (either pitch, or roll, or yaw). However, inspection of their head tracking data revealed that they also made other unintended head movements as well. Since lag was introduced by buffering the six degrees of freedom (6-DOF) head pose data used to

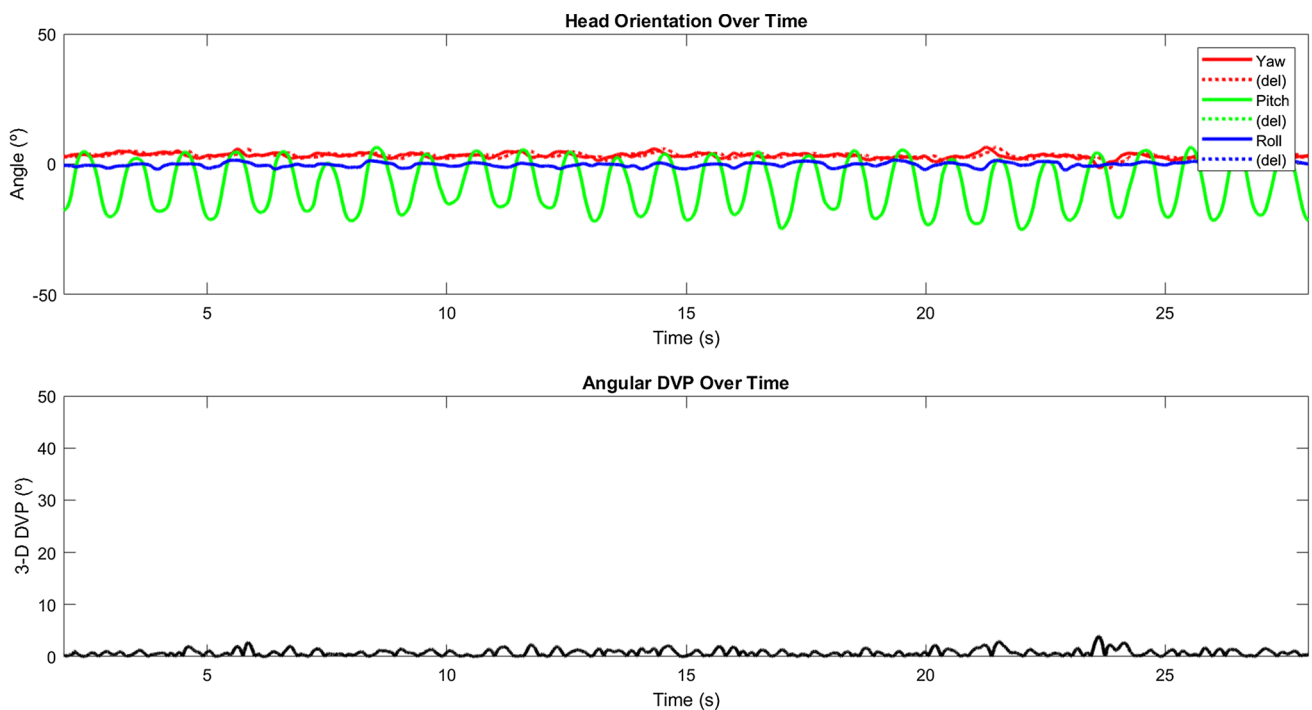


Fig. 3 (Top) This plot shows P2’s physical yaw, pitch and roll head orientations (in degrees) during a *Yaw-Lag + Pitch-Move* trial in the current study. P2’s estimated virtual head orientations during the trial are also shown as dotted lines in that plot. As in the previous example, the visual consequences of P2’s yaw axis head movements were delayed by an extra 222 ms. However, on this trial, P2 was instructed to move their head in pitch (not yaw) at 1.0 Hz. Thus, the added lag only had an impact on P2’s virtual head orientation due to their

unintended yaw movements. Only a red dotted line is visible in this plot—since only yaw head movements were delayed. (Bottom) This plot shows the unsigned differences between P2’s physical and virtual head orientations at each instant. As their yaw head movements were reduced in this trial (compared to the trial shown in Fig. 2), P2’s DVP is also considerably reduced (even though the same magnitude of yaw axis display lag was applied)

render the scene, this meant that the visual consequences of both their voluntary and involuntary head movements were delayed. As a result, *all* head movements—not just the voluntary ones about the requested axis—generated DVP during the different experimental conditions (based on our use of *6-DOF display lag*). Thus, it was not possible to study the effects of *pure yaw* (or *pure pitch* or *pure roll*) DVP in our past studies. When participants in these studies were asked to self-generate yaw head movements, they did produce mainly yaw DVP as instructed, but they also produced some pitch and roll DVP as well. In this specific case, most (but not all) of their DVP would have been orthogonal to the direction of gravity. Thus, proponents of the *subjective vertical conflict theory* could still argue that our participants became sick during laggy yaw head movement instruction conditions, because of the DVP generated by their involuntary pitch and roll head movements. That is, it was possible that the smaller pitch and roll head movements in our yaw rotation instruction trials were still capable of generating significant *subjective vertical conflicts* in the form of provocative pitch and roll DVP. Thus, instead of lagging the visual consequences of all HMD user head movements, it would have been better to selectively lag only those

associated with head-movements along a single axis (e.g. we could have selectively delayed display updates based on the participant’s yaw head movements, but not those based on their pitch and roll head movements). This is precisely what we planned to do in the current study, which examined the effects of selectively applying *1-DOF display lag* to the same axis as, or an orthogonal axis to, the participant’s instructed head movement (see Figs. 2, 3).¹

1.4 The current study

In this study, we tested the predictions of two different theories of cybersickness during active HMD VR: (1) the

¹ There are other ways one might test the predictions of the *subjective vertical conflict theory* during HMD VR. The HMD users in the current study were seated upright when they made their pitch (nodding “yes”) and yaw (“no” gesture) head-movements. Under these specific conditions, the *subjective vertical conflict theory* would predict that lag-induced sickness should be greater when they made “yes” (as opposed to “no”) head-movements. However, if they were made to lie on their sides (say in a left lateral recumbent posture), then the theory might predict that they would experience more lag-

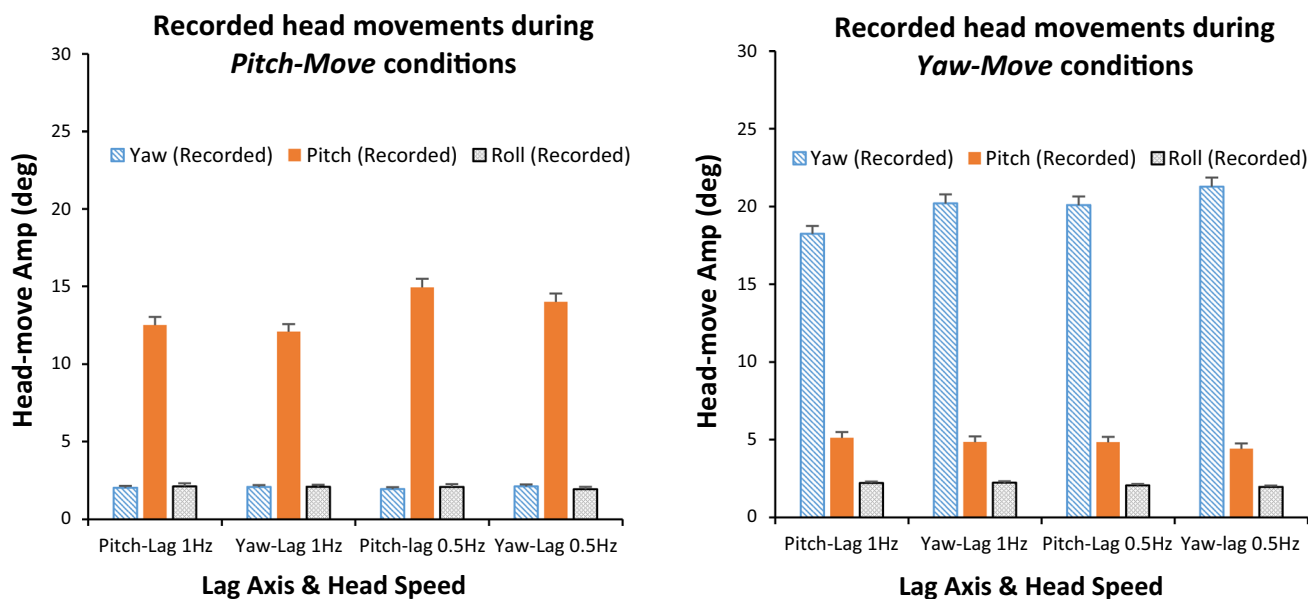


Fig. 4 Mean recorded head movement amplitudes (in degrees) across participants along all 3 axes for *Pitch-Move* (Left) and *Yaw-Move* (Right) instruction conditions. In each plot, data are shown separately

for each display lag axis and requested head speed condition. Error bars represent standard errors of the mean

subjective vertical conflict theory; and (2) the *DVP hypothesis*. According to the *DVP hypothesis*, sickness should increase with both the display lag and the HMD user’s head velocity (as both types of change are proposed to generate more provocative patterns of DVP). However, the *DVP hypothesis* does not make specific predictions about the severity of this sickness based on the axis of the user’s head movement—yaw head movements could potentially be as provocative as pitch head movements if the other characteristics of their DVP are similar. By contrast, the *subjective vertical conflict theory* predicts that cybersickness due to display lag should be less likely and severe when upright HMD users make yaw, compared to pitch, head movements. According to this theory, yaw DVP should not cause sickness—only pitch and roll DVP. So, if HMD users could make *pure* yaw head movements, then they should not become sick at all (even when presented with simulations that have very long display lags). Unfortunately, it is not possible to self-generate *pure* yaw (or *pure* pitch) head movements (please see the head movement data in Palmisano et al. 2023 and Fig. 4 of the current study). This means that when HMD users are asked to move their heads in yaw during simulations with 6-DOF display lag, some pitch and roll DVP will inevitably be generated (in addition to the

desired yaw DVP) by their involuntary head motions. Thus, in the current study, we used 1-DOF (as opposed to 6-DOF) display lag in order to strongly test the *subjective vertical conflict theory*. On different trials, participants were asked to make either yaw or pitch head movements (i.e. *Yaw-Move* and *Pitch-Move* conditions) at one of two rates (0.5 or 1.0 Hz). On each trial, lags of either 0, 89 or 222 ms were selectively added to display updates made along one axis (producing either *Yaw-Lag* or *Pitch-Lag* conditions). Lags were at approximately baseline levels for all updates made along the other two orthogonal axes. This meant that if sickness was experienced during a *Yaw-Lag* condition, then only DVP along the yaw axis could have been responsible for causing it.

According to the *subjective vertical conflict theory*, some pitch lag and some pitch movement would both be required to induce sickness in the current study.² If both are present together, then the theory predicts that sickness severity should increase with pitch head velocity and the amount of pitch lag. Yaw head velocity and yaw lag would not be expected to influence cybersickness. The theory therefore predicts that: (1) *Pitch-Lag + Pitch-Move* conditions should cause significantly more sickness than *Yaw-Lag + Yaw-Move* conditions (as the former should produce only pitch DVP,

Footnote 1 (continued)

induced sickness when they made “no” head-movements instead. However, with such a manipulation, it would still be important to use 1-DOF (as opposed to 6-DOF) lag in order to strongly test this theory.

² According to the theory, sickness could also be induced if there was some roll lag and some roll head movement. However, we did not include *Roll-Lag* conditions in the current study (only *Pitch-Lag* and *Yaw-Lag* conditions).

whereas the latter should produce only yaw DVP); and (2) *Pitch-Lag + Yaw-Move* conditions might cause more sickness than *Yaw-Lag + Pitch-Move* conditions (since involuntary pitch head movements made in the former condition should generate some pitch DVP, but deliberate pitch head movements made during the latter condition would not).

By contrast, the *DVP hypothesis* predicts that sickness should increase when lag magnitudes and head velocities increase along the same head axis. Unlike the *subjective vertical conflict theory*, it predicts that sickness should also be induced in the *Yaw-Lag + Yaw-Move* conditions. This sickness could be quite similar to that experienced in the *Pitch-Lag + Pitch-Move* conditions (e.g. when the lag magnitudes and head velocities along the respective axes are matched). According to the *DVP hypothesis*, both of these ('congruent' head-and-lag axis) conditions would be expected to produce more DVP, and thus more severe sickness, than the other ('incongruent') *Pitch-Lag + Yaw-Move* and *Yaw-Lag + Pitch-Move* conditions.

2 Method

2.1 Participants

Thirty-two participants were initially recruited from the University of Wollongong and the general population. These 16 females and 16 males ranged in age from 18 to 28 years ($M = 22.3$ years; $SD = 2.4$ years). All were healthy (i.e. with no self-reported visual, neurological, or vestibular impairments) and had either normal or corrected-to-normal vision (if they were required, participants wore contacts during the experiment). Twelve of these participants (10 males and 2 females) were regular gamers (defined as playing video games for more than 10 h per week). Only 4 of the 32 participants had used HMD VR previously. The remaining 28 participants were HMD VR novices. Two of these 32 participants dropped out early due to sickness (after recording sickness severity ratings above 15 out of 20 during their first block of trials³). The experiment was approved by the University of Wollongong Human Research Ethics Committee prior to testing. All participants provided written informed consent before commencing the study.

³ One participant dropped out on the third trial, and the other on the fifth trial, of their first experimental block (after experiencing an extra 89 ms *Pitch-Lag* under 1 Hz *Pitch-Move* conditions or an extra 222 ms *Yaw-Lag* under 0.5 Hz *Yaw-Move* conditions).

2.2 Materials

We used an Oculus Rift CV1 HMD and custom code to present our participants with a basic virtual room simulation. This (OLED) HMD had a resolution of 1080×1200 pixels per eye and a refresh rate of 90 Hz.⁴ Its horizontal and vertical fields of view were 87° and 88° , respectively, with 71.15° of binocular overlap. The interpupillary distance (IPD) of this HMD was set to 63.5 mm for all participants.⁵ The HMD had an outside-in tracking system which provided 6-DOF head tracking data sampled at 90 Hz. Our experimental code was run on a high-performance Microsoft Windows 10 Dell Precision 5820 computer, which had a NVidia GeForce GTX1080 graphics card and an Intel 7th generation CPU. The software presented seated participants with dynamic, first person views of a simulated environment. This virtual room consisted of a blue wireframe ground plane and a blue wireframe ceiling plane. These two surfaces—each simulated to be 16 m wide by 12 m deep—were presented on an otherwise black background. Participants were simulated to be seated at the very start of the ground plane. The ground was simulated to lie 1 m below their eye level, and the ceiling was simulated to lie 4 m above the ground.

Using previously published methods (Feng et al. 2019), the baseline lag of our PC VR system was estimated to be ~ 4 ms when running our custom code. This was the average effective display lag in our study when no extra lag was imposed. On different trials, small constant increments in display lag were added on top of this ~ 4 ms baseline lag by: (1) constructing a circular memory array to store the participant's head tracking data; and (2) increasing the number of elements in that array (from $N = 1$ to $N = 8$ or 20) to increase the added lag from 0 ms to either 89 or 222 ms.⁶ In previous studies (Feng et al. 2019; Kim et al. 2020; Palmisano et al. 2019, 2023), we used this method of lag injection to delay the visual consequences of *all* user head movements made during a trial. That is, we imposed 6-DOF display lag onto the simulation, where display updates were delayed for all 3 types of linear, and all 3 types of angular, head movement. However, in the current study, we only delayed the visual consequences of *one type* of (angular) head movement per trial. We imposed 1-DOF display lag into our simulations by buffering one of the Euler angle components (either yaw or pitch) of the participant's head pose and using

⁴ The obtained number of frames/second were the same as the prescribed frame rate (there were no frame drops).

⁵ This could have increased the likelihood of cybersickness in female participants — as poor IPD fit has been shown to be a cause of sickness (Stanney et al. 2020) and female IPDs are smaller on average than males (Dodgson 2004). However, sex was not found to have a significant effect on sickness in this study, $t(28) = -0.005$, $p = 1.00$.

⁶ As baseline lag was ~ 4 ms, the total lag imposed in these other conditions would have been 93 ms or 226 ms.

current head tracking estimates for the remainder of the head pose. During *Pitch-Lag* trials, we delayed display updates by an extra 0, 89 or 222 ms for pitch (but not yaw and roll) head movements. During *Yaw-Lag* trials, we delayed display updates by an extra 0, 89 or 222 ms for yaw (but not pitch and roll) head movements. Display updates associated with head translations were not artificially delayed for any trial. The time between making any head translation—or any roll head rotation—and the display updating was the same as for the baseline conditions (i.e. ~ 4 ms) in these *Pitch-Lag* and *Yaw-Lag* conditions.

In this within-subjects designed study, the main experimental manipulations occurred across trials and blocks. Thus, after each trial, we obtained the participant's ratings of sickness severity using the *Fast Motion Sickness* (FMS) scale [from 0 = "no sickness at all" to 20 = "frank sickness"; see Keshavarz and Hecht (2011)]. We also obtained their ratings of perceived scene instability during the trial ["how stable did the environment appear to be?" from 0 = "stable" to 20 = "completely unstable"; see Kim et al. (2021)]. These sickness and instability ratings were each made using a partially filled 2-D virtual rectangular figure, which was presented in their HMD. Pressing the up and down arrow keys on the computer's keyboard either filled or emptied this rectangular figure, respectively (to a maximum value of 20 or a minimum value of 0, in steps of ± 1 with each key press).

We were also interested in the sickness symptoms generated by this study. Thus, directly before their first, and directly after their last, HMD VR exposure block, participants provided information about their sickness symptoms by completing the *Simulator Sickness Questionnaire* (SSQ) (see Kennedy et al. 1993). They indicated their degree of general discomfort, fatigue, headache, eye strain, difficulty focussing, increased salivation, sweating, nausea, difficulty concentrating, fullness of head, blurred vision, dizziness with eyes open, dizziness with eyes closed, vertigo, stomach awareness, and burping (as none/slight/moderate/severe in each case). Their responses were then used to calculate a total sickness score (SSQ-T) and 3 sub-scores: (1) a disorientation score (SSQ-D), (2) a nausea score (SSQ-N), and (3) an oculomotor discomfort score (SSQ-O).

2.3 Design

This experiment had a 2 (HEAD MOVEMENT AXIS) \times 2 (HEAD SPEED) \times 2 (LAG AXIS) \times 3 (LAG MAGNITUDE) within-subjects design. It was comprised of four blocks, each of which consisted of 6 experimental trials. On each of these trials, participants made continuous oscillatory head movements while viewing the virtual room (see Supplementary Movie #1). HEAD MOVEMENT AXIS and LAG AXIS factors were manipulated across blocks as follows. Participants were either instructed to make only

pitch head movements (*Pitch-Move* blocks) or only yaw head movements (*Yaw-Move* blocks) (see Fig. 1). In half of these blocks, display lag was only added to the visual consequences of head movements made along that same axis (i.e. the *Pitch-Lag + Pitch-Move* and *Yaw-Lag + Yaw-Move* blocks). In the remaining blocks, display lag was only added to the visual consequences of unintended head movements made along an orthogonal axis (i.e. the *Pitch-Lag + Yaw-Move* and *Yaw-Lag + Pitch-Move* blocks). Block order was randomised across participants. In contrast to the above manipulations, HEAD SPEED and LAG MAGNITUDE factors were manipulated within (as opposed to across) blocks. Trial order was randomised within each block. Depending on the trial in the block: (1) the requested head movements were either slow (0.5 Hz) or fast (1.0 Hz); and (2) the visual consequences of these head movements were either delayed by an extra 0, 89, or 222 ms (on top of the ~ 4 ms baseline system lag).

In this study, we selectively introduced DVP about only one of the three cardinal head axes. For example, consider the situation in a *Pitch-Lag + Pitch-Move* trial with a LAG MAGNITUDE of 222 ms. On this trial, display updates would have been delayed by an extra 222 ms for any tracked pitch head movements [estimated total lag = ~ 4 ms (baseline lag) + 222 ms (extra lag) = 226 ms]. However, display updates would have been identical to those in the baseline conditions for any other tracked head movements (i.e. these would only have been delayed by ~ 4 ms for unintended yaw and roll head-rotations and all head translations). Now let us consider a *Pitch-Lag + Yaw-Move* trial with a LAG MAGNITUDE of 222 ms. On this trial, display updates for the instructed yaw head movements should have occurred after only ~ 4 ms. Only display updates for the participant's unintended pitch head movements would have been significantly delayed on this trial. Examples of the DVP produced by this amount of *Pitch-Lag* are shown for both pitch and yaw head movements in supplementary movie #2. Similarly, examples of the DVP produced by 222 ms of *Yaw-lag* are shown for yaw and pitch head movements in supplementary movie #3.

2.4 Procedure

Before each experimental block, our seated participants first observed the experimenter make the type of (pitch/yaw) head movement required for all of the trials in that block. They then attempted to recreate that head movement (matching their head movement amplitude to that of the experimenter). After donning their HMD, they completed two practice trials—one fast and one slow HEAD SPEED condition, both with 0 ms of added display lag. On each of these trials, participants made continuous oscillatory head movements for 35 s (timed to the sound of an audible metronome). During this time, the virtual room was only

visible for the last 30 s of each trial (note: for the first 5 s, the screen was completely black, except for a green fixation point). After 35 s had elapsed, and the virtual room had disappeared from view, they were then prompted by the display to provide ratings of: (1) the severity of their sickness⁷; and (2) the perceived stability of their virtual environment during the simulation (using the virtual rating scales described in Sect. 2.2).

After a two-minute break, and confirmation that they had fully recovered, participants then commenced their experimental trials. These were identical to their practice trials—except that the LAG MAGNITUDE now varied from trial-to-trial. After completing their sickness and stability ratings for each trial, there was a delay of at least 90 s before the next trial. During each block of 6 trials, participants remained in HMD VR. When it was time for the next trial, this was indicated to them via a text prompt on their display and there was also a check that they felt well enough to continue. It took approximately 14 min to complete one block of trials. Each block roughly approximated the common situation faced by HMD users, where periods of minimal conflict are interspersed by brief periods of salient conflict (e.g. due to tracking problems or an under-performing VR system). At the end of each block, there was a minimum 10-min break before the next block of trials began. This study design (with its brief exposures to conditions of imposed display lag, its interstimulus intervals that were much longer than these brief exposures, and its significant breaks between blocks) was aimed at minimising the likelihood of sickness contamination across trials and blocks.

2.5 DVP estimation

After testing had completed, we estimated each participant's instantaneous DVP throughout each trial. We first obtained their yaw, pitch, and roll head orientation at each instant. Head orientation in the Oculus Rift CV1 HMD is reported in terms of yaw–pitch–roll rotations in a right-handed coordinate system (yaw rotation occurs around the y-axis, pitch rotation occurs around the x-axis, and roll rotation occurs around the z-axis—see <https://developer.oculus.com/documentation/native/pc/dg-sensor/>). In this study, the virtual room was visible for 30 s during each 35 s trial. At each instant during that 30 s exposure, we estimated the participant's physical head orientation to be their recorded head orientation in yaw, pitch and roll at that particular time (based on the HMD's sensors). We then estimated their virtual head

orientation⁸ throughout the trial using these same recorded head orientation data. This was done by selectively phase shifting their recorded head orientation data along one axis (either pitch or yaw) by the LAG MAGNITUDE for that trial (either 0, 89 or 222 ms) to obtain the head pose used for rendering the virtual environment at each instant.

Figure 2 shows the 3-D head orientation and DVP data over time for one participant (P2) in a *Yaw-Lag + Yaw-Move* trial. On this trial: (1) P2 was instructed to make only yaw head movements at 1 Hz; and (2) display updates based on their tracked yaw movements were artificially delayed by an extra 222 ms. The *solid red, green and blue traces* in Fig. 2 (top) show P2's recorded head orientations in yaw, pitch and roll at each instant, whereas the *dotted lines* show the actual pose used for rendering throughout that trial. The minimum rotation angles between quaternion representations of these recorded (solid line) and virtual (dotted line) head pose traces at each instant were used to form the DVP time series data for that trial—shown in Fig. 2 (bottom). For this particular trial, only the red dotted virtual yaw head orientation data can be seen (as the undelayed green and blue dotted virtual data lines perfectly overlap the green and blue solid recorded data lines). While P2 made unintended pitch and roll head movements during this trial, their virtual pitch and roll head orientations were not affected by the application of *Yaw-Lag*. Thus, the data shown in Fig. 2 (bottom) actually reflect the resulting pure unsigned yaw-DVP in degrees.

Figure 3 (top) provides another example—this time of a *Yaw-Lag + Pitch-Move* trial. On this trial: (1) P2 was instructed to make only pitch head movements at 1 Hz; and (2) display updates to any unintended yaw movements were artificially delayed by an extra 222 ms. As can be seen from Fig. 3 (top), this 222 ms of *Yaw-Lag* did not alter P2's virtual pitch or roll head orientation. This extra lag only delayed display updates based on P2's unintended yaw head movements (note: a red dotted line is visible on close inspection of this figure). Again, the data shown in Fig. 3 (bottom) reflect the pure unsigned yaw-DVP (in degrees) generated during this trial.

The bottom plots of Figs. 2 and 3 provide quantitative estimates of P2's instantaneous sensory input conflict during each trial. When P2 was instructed to make continuous oscillatory yaw head movements, 222 ms of extra *Yaw-Lag* was found to generate large amplitude, time-varying DVP (see Fig. 2, bottom). However, the same 222 ms *Yaw-Lag* generated very little DVP under pitch head movement instruction conditions (see Fig. 3, bottom), since P2's physical yaw head movements were much smaller under these conditions (as they were incidental and involuntary rather than deliberate).

⁷ As physiological variables have yet to be established as standalone objective measures of sickness (Keshavarz et al. 2022), subjective measures of cybersickness (the FMS and SSQ) were used instead in this study.

⁸ i.e. how they perceived their head to be oriented based on what they were seeing.

Using these estimates of the DVP time series data, we then calculated summary measures of the DVP for each trial. First, we calculated the peak and standard deviation of the DVP for each trial (i.e. *peakDVP* and *stdDVP*). These provided estimates of the *spatial magnitude* of the DVP for that trial. Second, we examined how the DVP evolved over time on each trial [i.e. its *temporal dynamics*—see Peng et al. (1995), Ihlen (2012)]. We initially conducted (monofractal) detrended fluctuation analyses (DFA) on these DVP time series data (Peng et al. 1995). The scaling exponents (α_{DVP}) of these DFAs provide an index of how self-similar the DVP data are across different time scales. Time series data can differ in terms of their degree of multifractality (Koslucher et al. 2016). Thus, we also examined the multifractal spectrum width of the DVP (*SpecWidthDVP*) using multifractal detrended fluctuation analysis (Ihlen 2012). The wider the multifractal spectrum, the more multifractal the DVP. The *peak*, *standard deviation*, *DFA α* and *multifractal spectrum widths* of these DVP data were then considered as potential predictors of the sickness severity ratings obtained for each trial.

3 Results

3.1 Checks on participant head movement compliance

Head movement frequencies and amplitudes were estimated for each trial by fitting a sinusoidal function to each trace (note: all values in yaw, pitch and roll were initially shifted so that their mean values for the trial were 0). As can be seen in Fig. 4, head movements were confirmed to occur primarily along the axis of the instructed head movement in each block (with considerably smaller head movement amplitudes observed along the other two orthogonal axes). When participants were instructed to make only pitch head movements, their yaw and roll head movement amplitudes were on average 85% less than those made in pitch. Similarly, when they were instructed to make only yaw head movements, their pitch and roll head movement amplitudes were on average 83% less than those made in yaw.

3.1.1 Effects of experimental conditions on user head movements

We were interested in how participant head movements varied across our different experimental conditions. Thus, we conducted separate 2 (HEAD MOVEMENT AXIS) \times 2 (HEAD SPEED) \times 2 (LAG AXIS) \times 3 (LAG MAGNITUDE) repeated measures ANOVAs on the frequencies and the amplitudes of their recorded head movements.

3.1.1.1 Head movement amplitude There was a significant main effect of HEAD MOVEMENT AXIS, $F(1, 29)=49.934$, $p<0.001$, $\eta^2=0.633$. On average, head movements were significantly larger during *Yaw-Move* ($M=24.8^\circ$) compared to *Pitch-Move* ($M=15.9^\circ$) conditions (see Fig. 5, left). There was also a significant main effect of HEAD SPEED, $F(1, 29)=43.49$, $p<0.001$, $\eta^2=0.600$ —indicating that head movements were smaller on average during the 1 Hz ($M=17.7^\circ$), compared to the 0.5 Hz ($M=23.0^\circ$), conditions. However, the main effects of LAG AXIS and LAG MAGNITUDE were both non-significant, $F(1, 29)=0.054$, $p=0.817$, $\eta^2=0.002$ and $F(2, 58)=1.318$, $p=0.276$, $\eta^2=0.043$. None of the 2-, 3- or 4-way interactions were significant.

3.1.1.2 Head movement Frequency As expected, there was a significant main effect of HEAD SPEED, $F(1, 29)=913.496$, $p<0.001$, $\eta^2=0.969$ —confirming that head movement frequencies were higher during the 1 Hz ($M=0.8$ Hz), compared to the 0.5 Hz ($M=0.41$ Hz), conditions (see Fig. 5, right). The main effects of HEAD MOVEMENT AXIS and LAG AXIS were not significant, $F(1, 29)=0.034$, $p=0.854$, $\eta^2=0.001$ and $F(1, 29)=0.091$, $p=0.765$, $\eta^2=0.003$. However, there was a significant main effect of LAG MAGNITUDE, $F(2, 58)=5.415$, $p<0.007$, $\eta^2=0.157$. Bonferroni-corrected pairwise comparisons revealed that head movement frequency was significantly higher on average in the 89 ms ($M=0.61$ Hz), compared to the 222 ms ($M=0.59$ Hz), added lag conditions ($p=0.009$). However, head movement frequency was not different between the 89 ms added lag condition and the baseline lag condition ($M=0.61$ Hz) ($p=0.773$).

3.2 Cybersickness data

3.2.1 Overall experience of sickness and symptomology

Directly before the first, and after the last, HMD VR exposure, sickness symptoms were measured using the SSQ. Responses to the post-exposure SSQ were used to assess the overall amount of sickness generated by our study (whereas responses to the pre-exposure SSQ provided a baseline). Across our participants, post-exposure SSQ-T scores ranged from 3.74 to 123—with the mean and standard deviation being 46.4 and 32.7, respectively (see Fig. 6, top). This indicated that a significant number of participants experienced functionally significant sickness in our study [defined as having SSQ scores greater than 20.1—Stanney et al. (2014), see Lawson and Stanney (2021) for a discussion]. The SSQ's sub-scores were also used to assess their nausea (SSQ-N), disorientation (SSQ-D) and oculomotor (SSQ-O) symptoms (see Fig. 6, top). As expected for HMD VR (e.g. Rebenitsch and Owen 2016), we found that our

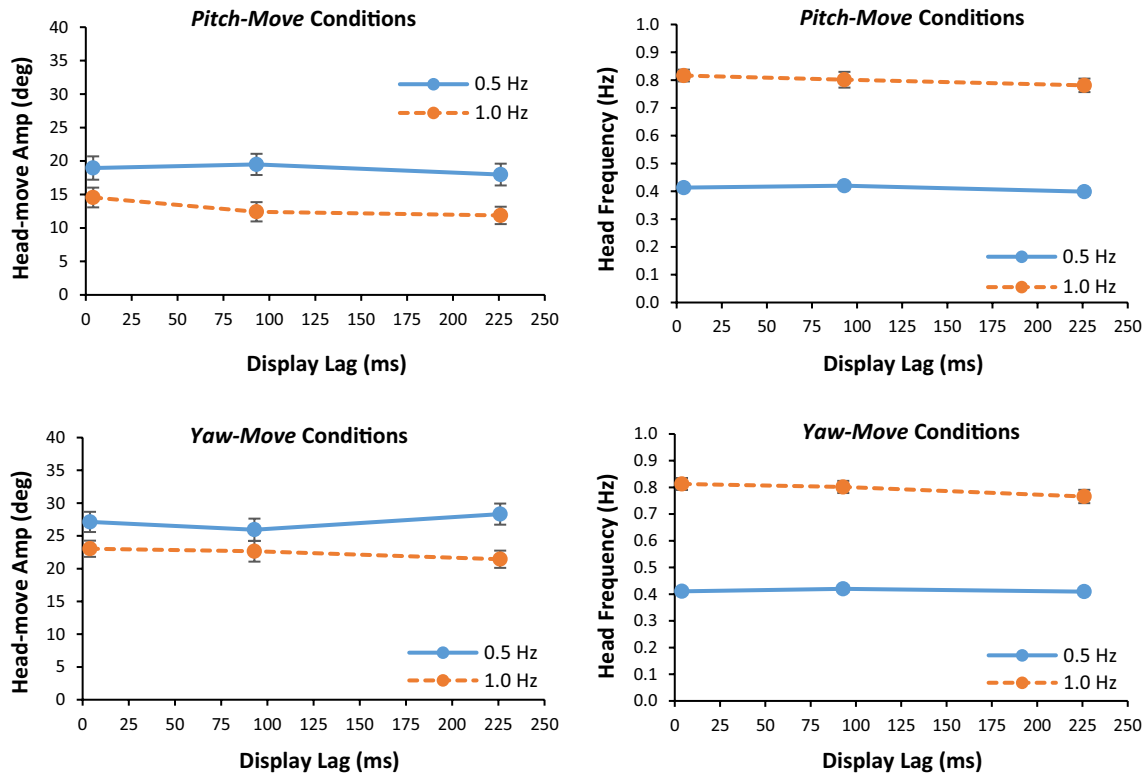


Fig. 5 Mean head movement amplitudes (Left in deg) and frequencies (Right in Hz) for *Pitch-Move* (Top) and *Yaw-Move* (Bottom) conditions. In each plot, data are shown for each lag magnitude (0, 89, or

222 ms + 4 ms baseline lag) and head speed (0.5 or 1.0 Hz) condition. Error bars represent standard errors of the mean

participants' post-exposure scores were higher on average for the SSQ-D ($M = 58.9$, $SD = 52.4$) than for the SSQ-N ($M = 31.2$, $SD = 23.1$) and the SSQ-O ($M = 37.9$, $SD = 24.0$).

3.2.2 Effects of experimental manipulations on sickness severity

Separate 2 (HEAD MOVEMENT AXIS) × 2 (HEAD SPEED) × 2 (LAG AXIS) × 3 (LAG MAGNITUDE) repeated measures ANOVAs were conducted on our participants' sickness severity ratings (as measured by the FMS). According to our *DVP hypothesis*, LAG MAGNITUDE effects should depend on both the HEAD MOVEMENT AXIS and the LAG AXIS. Consistent with this prediction, we found a significant 2-way interaction between HEAD MOVEMENT AXIS and LAG AXIS for sickness severity [$F(1, 29) = 23.774$, $p < 0.001$, $\eta^2 = 0.450$], as well as a significant 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and LAG MAGNITUDE [$F(1.645, 47.702) = 15.822$, $p < 0.001$, $\eta^2 = 0.353$]. While the main effects of HEAD MOVEMENT AXIS and LAG AXIS were not significant [$F(1, 29) = 2.014$, $p = 0.166$, $\eta^2 = 0.065$ and $F(1, 29) = 0.13$, $p = 0.909$, $\eta^2 = 0.000$], there was a significant main effect of LAG MAGNITUDE on sickness severity

[$F(1.278, 37.054) = 26.658$, $p < 0.001$, $\eta^2 = 0.479$]. Bonferroni-corrected pairwise comparisons revealed that: (1) adding 222 ms lag ($M = 4.4$) produced significantly higher sickness ratings than adding 89 ms lag ($M = 3.5$); and (2) adding 89 ms lag produced significantly higher sickness ratings than the baseline lag condition ($M = 2.2$) ($p < 0.001$ in both cases). However, consistent with the expected three-way interaction, these LAG MAGNITUDE effects varied across blocks. From Fig. 6 (bottom left), it can be seen that sickness increased more with LAG MAGNITUDE when the LAG AXIS for the trial was the same as the HEAD MOVEMENT AXIS (i.e. sickness was generally worse, and increased more with the amount of added lag, in the *Pitch-Lag + Pitch-Move* and *Yaw-Lag + Yaw-Move* conditions compared to the *Pitch-Lag + Yaw-Move* and *Yaw-Lag + Pitch-Move* conditions).

Our *DVP hypothesis* also predicted that sickness should increase with the participants' HEAD SPEED, but that such effects would depend on the HEAD MOVEMENT AXIS and LAG AXIS as well. While we did find a significant main effect of HEAD SPEED [$F(1, 29) = 10.466$, $p = 0.003$, $\eta^2 = 0.265$], the expected 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and HEAD SPEED was not significant, $F(1, 29) = 0.565$, $p = 0.458$, $\eta^2 = 0.019$. As expected, the main effect of HEAD SPEED indicated that

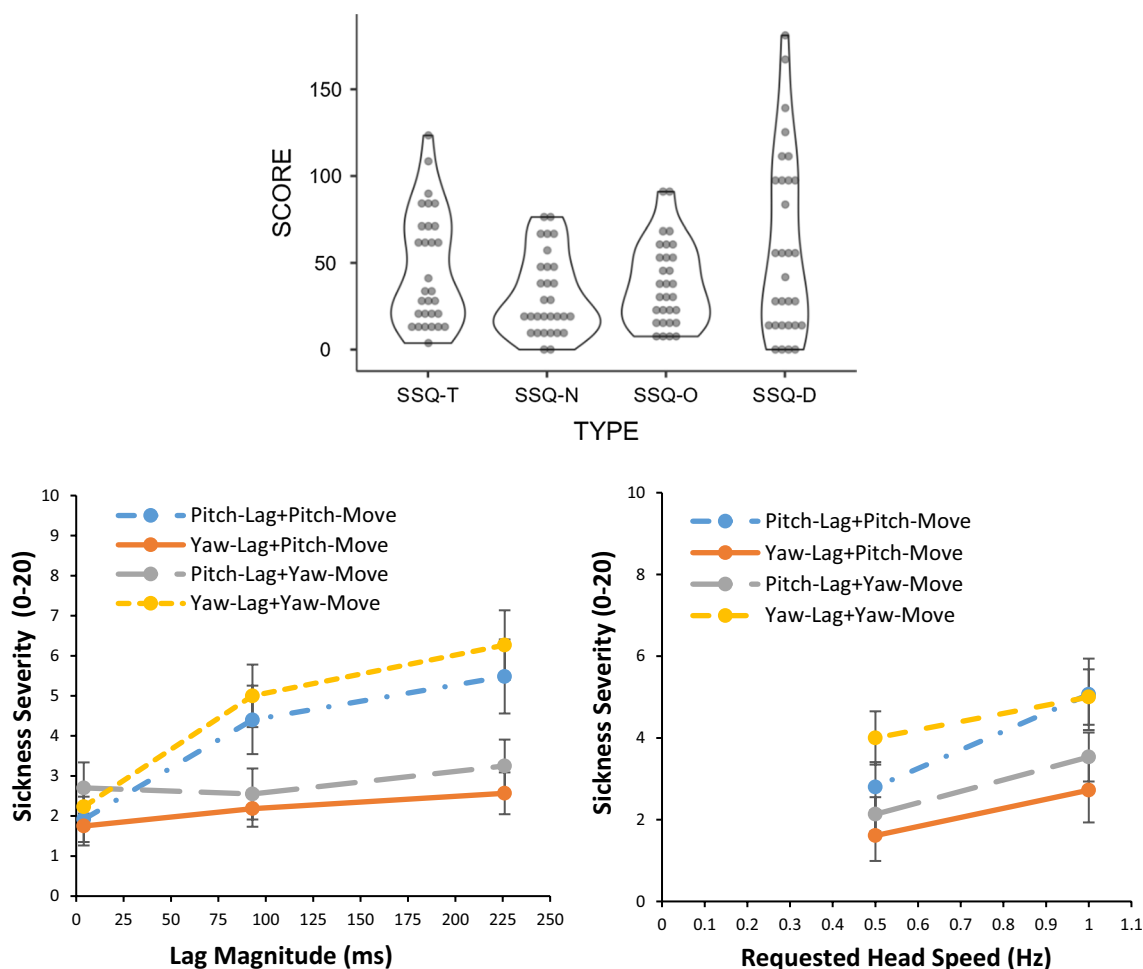


Fig. 6 (Top) Post-exposure SSQ-T (Total), SSQ-N (Nausea), SSQ-O (Oculomotor) and SSQ-D (Disorientation), scores for our 30 participants. (Bottom left) Shows the effects of HEAD MOVEMENT AXIS (*Pitch-Move* or *Yaw-Move*), LAG AXIS (*Pitch-Lag* or *Yaw-Lag*) and LAG MAGNITUDE (0–222 ms+4 ms baseline lag) on mean FMS

ratings (0–20). (Bottom right) Shows the effects of HEAD MOVEMENT AXIS (*Pitch-Move* or *Yaw-Move*), LAG AXIS (*Pitch-Lag* or *Yaw-Lag*) and HEAD SPEED (1.0 or 0.5 Hz) on mean FMS ratings (0–20). When present, the error bars represent standard errors of the mean

sickness was more severe during the 1.0 Hz ($M=4.08$), compared to the 0.5 Hz ($M=2.64$), conditions. We also found a significant 2-way interaction between LAG AXIS and HEAD SPEED, $F(1, 29)=4.865, p=0.035, \eta p^2=0.144$ —see Fig. 6 (bottom right). This indicated that sickness increased more with head speed during *Pitch-Lag* conditions (where it went from 2.5 to 4.3 on average) than during *Yaw-Lag* conditions (where it went from 2.8 to 3.9 on average).

According to the *subjective vertical conflict theory*: (1) laggy trials in the *Pitch-Lag + Pitch-Move* block should have been the most provocative for sickness; and (2) laggy trials in the *Pitch-Lag + Yaw-Move* block should have been more provocative than those in the two *Yaw-Lag* blocks.⁹

To test these predictions, we ran a one-way repeated measures ANOVA on our sickness severity data, with BLOCK TYPE as the factor (note: only data from the 89 ms and 222 ms added lag conditions were included in this analysis). We found a significant main effect of BLOCK TYPE, $F(3, 87)=9.378, p<0.001, \eta p^2=0.244$. Contrary to the predictions of *subjective vertical conflict theory*, Bonferroni-corrected pairwise comparisons revealed that cybersickness was not different in: (1) *Pitch-Lag + Pitch-Move* ($M=4.9$) compared to *Yaw-Lag + Yaw-Move* ($M=5.6$) conditions ($p=0.436$); and (2) *Pitch-Lag + Yaw-Move* ($M=2.9$) compared to *Yaw-Lag + Pitch-Move* ($M=2.4$) conditions

⁹ This is because participants still made involuntary pitch head movements during *Yaw-Move* conditions. So in *Pitch-Lag + Yaw-Move* conditions, when the visual consequences of these unintended

Footnote 9 (continued)
pitch head movements were lagged, this could still have generated provocative *Pitch-DVP* (see Fig. 3).

($p=0.349$). However, consistent with the *DVP hypothesis*, *Yaw-Lag + Yaw-Move* conditions did generate more severe sickness than the *Pitch-Lag + Yaw-Move* ($p=0.009$) and *Yaw-Lag + Pitch-Move* ($p<0.001$) conditions.

3.2.3 Checks for sickness contamination due to repeated exposure

Studies which repeatedly expose participants to potentially provocative motion stimuli run the risk of sickness contamination across trials and blocks. This is why presentation orders were randomised for each of our within subjects factors (e.g. BLOCK TYPE, HEAD SPEED and LAG MAGNITUDE). We also attempted to minimise sickness contamination in our study by using short 30 s exposures to increased display lag, much longer 90 s interstimulus intervals between exposures, and long 10-min breaks between the 4 experimental blocks. We also performed a repeated measures ANOVA on the FMS data to check for BLOCK NUMBER (1–4) and TRIAL NUMBER (1–6) effects. We found a significant main effect of TRIAL NUMBER, $F(2.880, 83.530) = 3.776$, $p = 0.015$, $\eta^2 = 0.115$ —indicating that sickness severity increased modestly from trial 1 ($M = 2.4$) to trial 6 ($M = 3.9$) within blocks. However, the main effect of BLOCK NUMBER was not significant, $F(3, 87) = 1.40$, $p = 0.248$, $\eta^2 = 0.046$ —this confirmed that sickness was not significantly more severe in the later blocks. The BLOCK NUMBER by TRIAL NUMBER interaction [$F(7.405, 214.737) = 1.422$, $p = 0.194$, $\eta^2 = 0.047$] was also not significant.

3.3 Effects of experimental manipulations on proposed sickness predictors

According to the *DVP hypothesis*, we should be able to predict a participant's experience of sickness on a trial based on their objectively estimated DVP (and possibly also based on their scene instability ratings). Thus, we were interested in how DVP and scene instability varied across our experimental conditions. In order to explore these effects, we conducted separate 2 (HEAD MOVEMENT AXIS) \times 2 (HEAD SPEED) \times 2 (LAG AXIS) \times 2 or 3 (LAG MAGNITUDE) repeated measures ANOVAs on our *peakDVP*, *stdDVP*, *α DVP*, *SpecWidthDVP* and *scene instability* data.¹⁰

¹⁰ Note: only data from the 89 and 222 ms added lag conditions were included in the DVP analyses (as DVP could not be objectively estimated for our 0 ms baseline conditions). By contrast, data from all 3 LAG MAGNITUDE conditions (i.e. 0, 89 and 222 ms added lag) were included in our *scene instability* analysis.

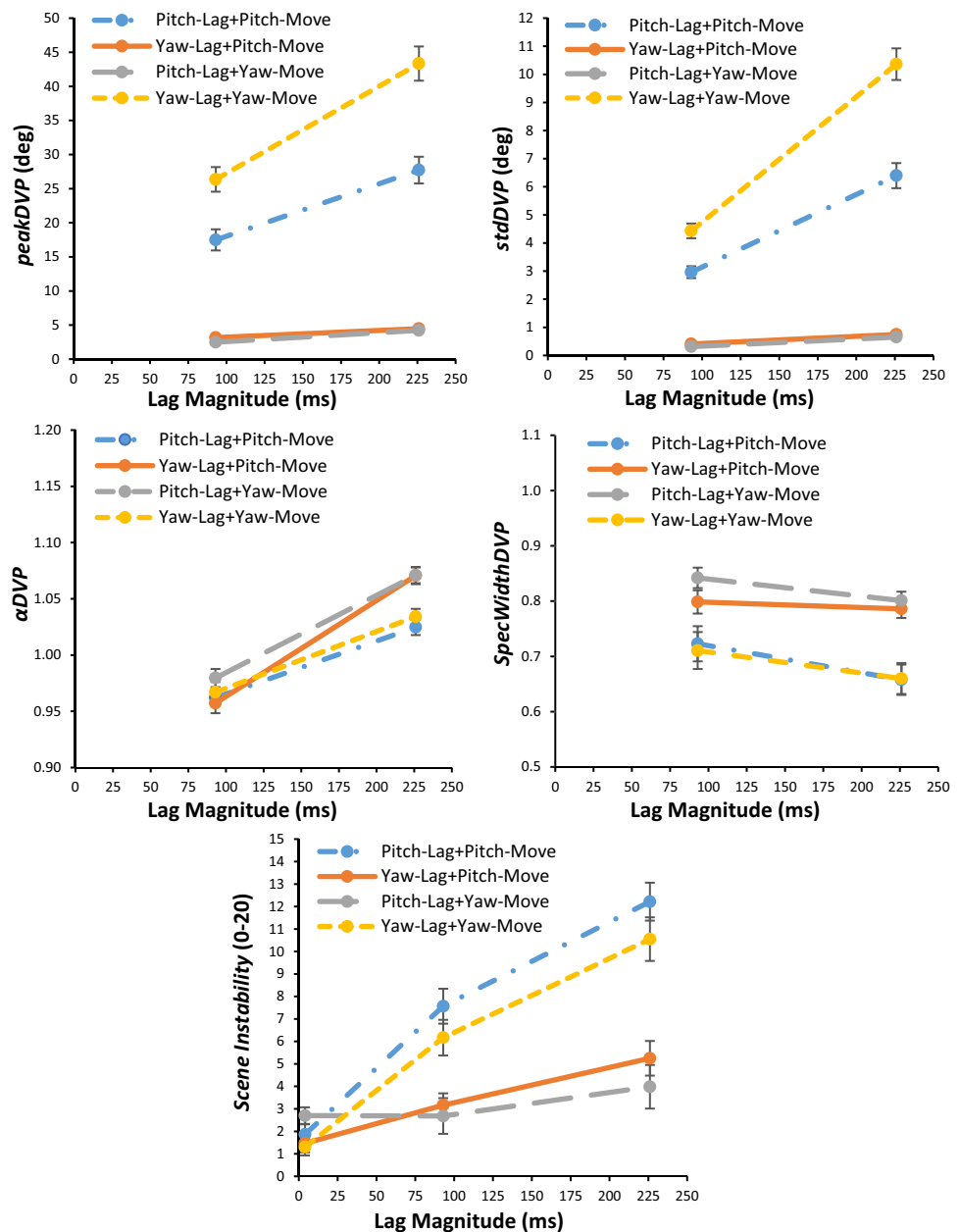
3.3.1 Effects of experimental manipulations on peakDVP

As expected, there was a significant 2-way interaction between HEAD MOVEMENT AXIS and LAG AXIS [$F(1, 29) = 312.540$, $p < 0.001$, $\eta^2 = 0.915$], and a significant 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and LAG MAGNITUDE for *peakDVP* [$F(1, 29) = 115.920$, $p < 0.001$, $\eta^2 = 0.800$]. There was also a significant main effect of LAG MAGNITUDE on *peakDVP*, $F(1, 29) = 144.683$, $p < 0.001$, $\eta^2 = 0.833$ —where adding 222 ms lag ($M = 19.9^\circ$) produced larger *peakDVP* than adding 89 ms lag ($M = 12.4^\circ$). However, as can be seen in Fig. 7 (top left), this LAG MAGNITUDE effect was marginal to the interaction, and primarily driven by the two conditions where the LAG AXIS was the same as the HEAD MOVEMENT AXIS (i.e. the *Pitch-Lag + Pitch-Move* and *Yaw-Lag + Yaw-Move* conditions). There was also a significant 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and HEAD SPEED for *peakDVP*, $F(1, 29) = 34.652$, $p < 0.001$, $\eta^2 = 0.554$. In addition, there was a main effect of HEAD SPEED, $F(1, 29) = 37.731$, $p < 0.001$, $\eta^2 = 0.565$ —where 1.0 Hz conditions ($M = 18.8^\circ$) produced larger *peakDVP* than 0.5 Hz conditions ($M = 13.6^\circ$). However, the effect of HEAD SPEED on *peakDVP* was again driven by conditions where the LAG AXIS was the same as the HEAD MOVEMENT AXIS (see Fig. 8, top left).

3.3.2 Effects of experimental manipulations on stdDVP

As expected, there was a significant 2-way interaction between HEAD MOVEMENT AXIS and LAG AXIS [$F(1, 29) = 327.459$, $p < 0.001$, $\eta^2 = 0.919$], and a significant 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and LAG MAGNITUDE for the *stdDVP*, $F(1, 29) = 354.418$, $p < 0.001$, $\eta^2 = 0.924$. There was also a significant main effect of LAG MAGNITUDE, $F(1, 29) = 388.645$, $p < 0.001$, $\eta^2 = 0.931$ —where adding 222 ms lag ($M = 4.5^\circ$) produced larger *stdDVP* than adding 89 ms lag ($M = 2.0^\circ$). This LAG MAGNITUDE effect was again primarily driven by conditions where the LAG AXIS was the same as the HEAD MOVEMENT AXIS (i.e. *Pitch-Lag + Pitch-Move* and *Yaw-Lag + Yaw-Move*; see Fig. 7, top right). There was also a significant 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and HEAD SPEED for *stdDVP*, $F(1, 29) = 35.037$, $p < 0.001$, $\eta^2 = 0.547$. In addition, there was also a main effect of HEAD SPEED on *stdDVP*, $F(1, 29) = 41.529$, $p < 0.001$, $\eta^2 = 0.589$ —where 1.0 Hz conditions ($M = 3.7^\circ$) produced larger *stdDVP* than 0.5 Hz conditions ($M = 2.9^\circ$). However, this HEAD SPEED effect on *stdDVP* was again driven by conditions where the LAG AXIS was the same as the HEAD MOVEMENT AXIS (see Fig. 8, top right).

Fig. 7 These plots show HEAD MOVEMENT AXIS (*Pitch-Move* or *Yaw-Move*), LAG AXIS (*Pitch-Lag* or *Yaw-Lag*) and LAG MAGNITUDE (0–222 ms + 4 ms baseline lag) effects on *peakDVP* (top left), *stdDVP* (top right), *αDVP* (middle left), *SpecWidthDVP* (middle right), and *scene instability ratings* (bottom). Error bars in each case represent standard errors of the mean

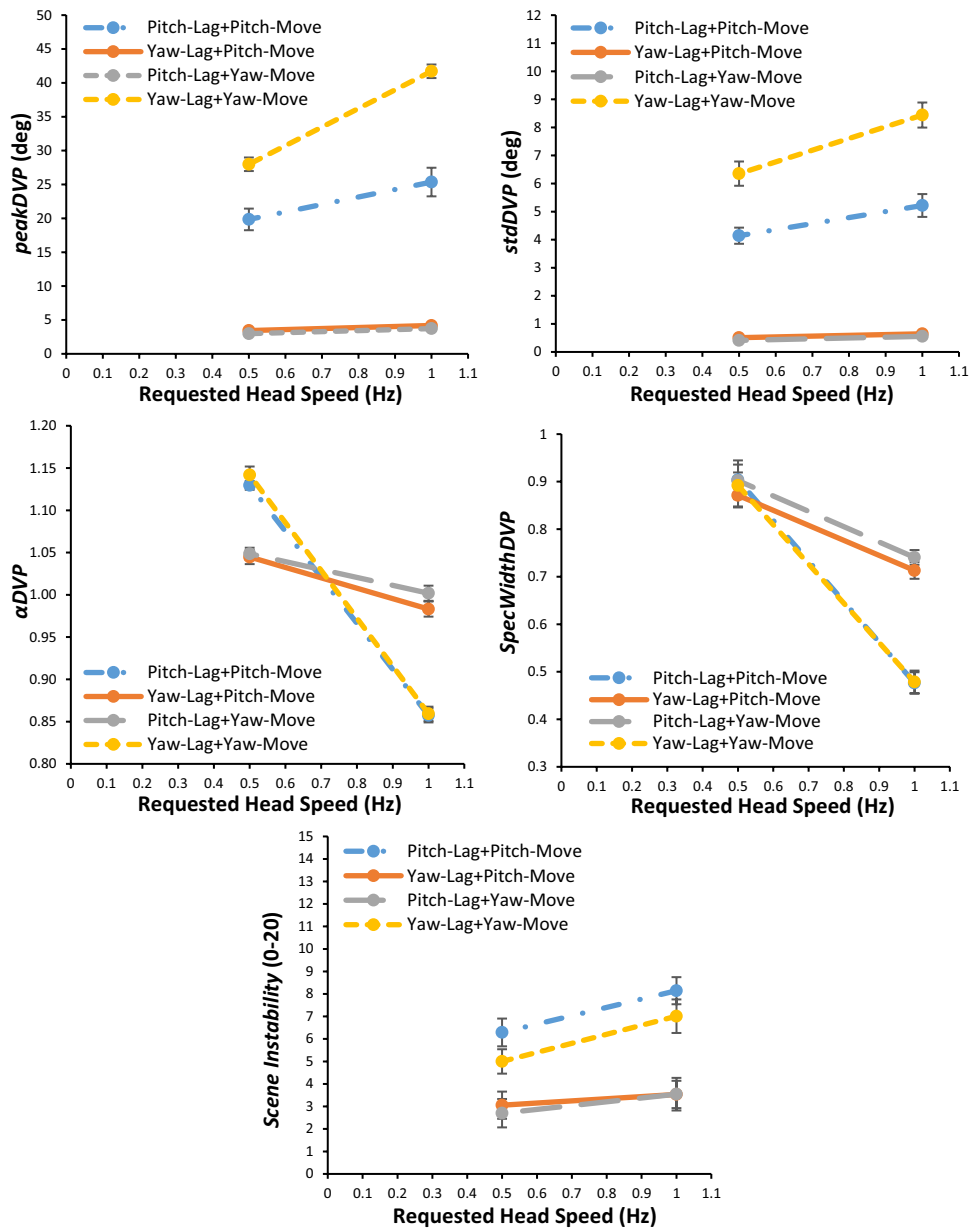


3.3.3 Effects of experimental manipulations on αDVP

There was a significant 2-way interaction between HEAD MOVEMENT AXIS and LAG AXIS [$F(1, 29) = 9.651, p < 0.004, \eta^2 = 0.250$], and a significant 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and LAG MAGNITUDE for αDVP [$F(1, 29) = 25.174, p < 0.001, \eta^2 = 0.465$] (see Fig. 7, middle left). There was also a significant main effect of LAG MAGNITUDE on $\alpha DVP, F(1, 29) = 554.855, p < 0.001, \eta^2 = 0.950$ —where adding 222 ms lag ($M = 1.1$) produced higher αDVP than adding 89 ms lag ($M = 0.97$). However, consistent with the 3-way interaction, the effect of LAG MAGNITUDE

varied by block. When 222 ms lag was added, αDVP was significantly lower in trials where the LAG AXIS was the same as the HEAD MOVEMENT AXIS ($M = 1.03$) compared to trials where it was different ($M = 1.07$) ($F(2.328, 67.515) = 16.452, p < 0.001, \eta^2 = 0.362$). However, a similar effect was not found when 89 ms lag was added ($F(2.297, 66.625) = 1.678, p < 0.191, \eta^2 = 0.055$)—under these conditions, αDVP was similar for all combinations of LAG AXIS and HEAD MOVEMENT AXIS. There was also a significant 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and HEAD SPEED for $\alpha DVP, F(1, 29) = 676.444, p < 0.001, \eta^2 = 0.959$. In addition, there was a main effect of HEAD SPEED,

Fig. 8 These plots show HEAD MOVEMENT AXIS (*Pitch-Move* or *Yaw-Move*), HEAD SPEED (0.5 or 1.0 Hz) and LAG AXIS (*Pitch-Lag* or *Yaw-Lag*) effects on *peakDVP* (top left), *stdDVP* (top right), *αDVP* (middle left), *SpecWidthDVP* (middle right), and *scene instability ratings* (bottom). Error bars in each case represent standard errors of the mean

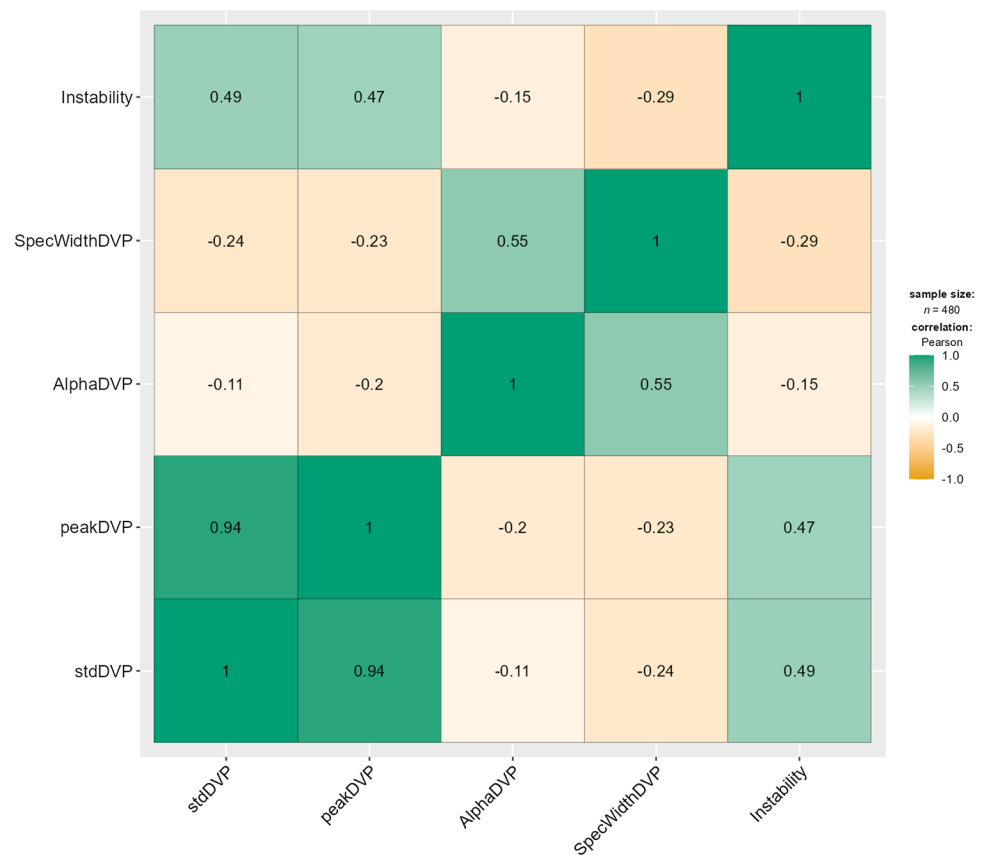


$F(1, 29) = 932.391, p < 0.001, \eta^2 = 0.970$, where 1.0 Hz conditions ($M = 0.93$) produced lower αDVP than 0.5 Hz conditions ($M = 1.1$). However, as can be seen in Fig. 8, middle left, in the 0.5 Hz conditions, trials where the LAG AXIS was the same as the HEAD MOVEMENT axis ($M = 1.14$) produced higher αDVP compared to trials where it was different ($M = 1.05$) ($p < 0.05$). By contrast, in the 1.0 Hz conditions, trials where the LAG AXIS was the same as the HEAD MOVEMENT AXIS ($M = 0.86$) produced lower αDVP compared to trials where it was different ($M = 0.99$) ($p < 0.05$).

3.3.4 Effects of experimental manipulations on the *SpecWidthDVP*

We found a significant 2-way interaction between HEAD MOVEMENT AXIS and LAG AXIS for *SpecWidthDVP*, $F(1, 29) = 24.483, p < 0.001, \eta^2 = 0.458$ – *SpecWidthDVP* was lower for trials where the HEAD MOVEMENT AXIS was the same as ($M = 0.71$), as opposed to different to ($M = 0.82$), the LAG AXIS (see Fig. 7, middle right). There was also a significant main effect of LAG MAGNITUDE for *SpecWidthDVP*, $F(1, 29) = 13.105, p < 0.001$,

Fig. 9 Correlation matrix showing the associations between the standard deviation, the peak, the monofractal DFA α and the multifractal spectrum width of the DVP and perceived scene instability. Pearson's r values are provided for each relationship. In addition, the colour and shading of each cell indicates the direction (positive/negative) and the strength (0–1) of each association



$\eta^2 = 0.311$ —where adding 222 ms lag ($M = 0.73$) produced lower *SpecWidthDVP* than adding 89 ms lag ($M = 0.77$). We did not find a significant 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and LAG MAGNITUDE for *SpecWidthDVP*, $F(1, 29) = 1.976$, $p = 0.170$, $\eta^2 = 0.064$. However, there was a significant 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and HEAD SPEED on *SpecWidthDVP*, $F(1, 29) = 50.131$, $p < 0.001$, $\eta^2 = 0.634$. There was also a main effect of HEAD SPEED, $F(1, 29) = 173.786$, $p < 0.001$, $\eta^2 = 0.857$ —where 1.0 Hz conditions ($M = 0.60$) produced lower *SpecWidthDVP* than 0.5 Hz conditions ($M = 0.89$). However, as can be seen in Fig. 8, middle right, these effects of HEAD SPEED were much greater in conditions where the LAG AXIS was the same as the HEAD MOVEMENT AXIS (i.e. the *Pitch-Lag + Pitch-Move* and *Yaw-Lag + Yaw-Move* conditions).

3.3.5 Effects of experimental manipulations on scene instability ratings

As expected, there was a significant 2-way interaction between HEAD MOVEMENT AXIS and LAG AXIS [$F(1, 29) = 49.076$, $p < 0.001$, $\eta^2 = 0.629$], and a 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and LAG MAGNITUDE for *scene instability* ratings [$F(1.429, 41.448) = 66.162$, $p < 0.001$, $\eta^2 = 0.695$]. There was also

a significant main effect of LAG MAGNITUDE, $F(1.391, 40.345) = 87.486$, $p = 0.001$, $\eta^2 = 0.751$ —indicating that *scene instability* increased from 1.83 to 8 as the added lag increased from 0 to 222 ms. However, consistent with the 3-way interaction, this LAG MAGNITUDE effect was primarily driven by conditions where the LAG AXIS was the same as the HEAD MOVEMENT AXIS (i.e. the *Pitch-Lag + Pitch-Move* and *Yaw-Lag + Yaw-Move* conditions; see Fig. 7, bottom). There was also a significant 3-way interaction between HEAD MOVEMENT AXIS, LAG AXIS and HEAD SPEED for these *scene instability* ratings, $F(1, 29) = 4.233$, $p < 0.049$, $\eta^2 = 0.127$. In addition, there was a main effect of HEAD SPEED, $F(1, 29) = 10.536$, $p < 0.003$, $\eta^2 = 0.266$ —where 1.0 Hz conditions ($M = 5.6$) produced higher instability ratings than 0.5 Hz conditions ($M = 4.3$). This effect of HEAD SPEED on *scene instability* ratings was again primarily driven by conditions where the LAG AXIS was the same as the HEAD MOVEMENT AXIS (see Fig. 8, bottom).

3.4 Relationships between proposed sickness predictors

Before we examined whether *peakDVP*, *stdDVP*, α *DVP*, *SpecWidthDVP* and *scene instability* could be used to predict

Table 1 Exploratory factor analysis of the 5 proposed sickness predictor variables

Factor loadings	Factor		Uniqueness
	1 (spatial magnitude)	2 (temporal dynamics)	
<i>stdDVP</i>	0.993		0.00776
<i>peakDVP</i>	0.935		0.11372
<i>Instability</i>	0.477		0.72799
<i>SpecWidthDVP</i>		0.986	0.00478
<i>αDVP</i>		0.544	0.69288

'Minimum residual' extraction method was used in combination with a 'varimax' rotation. This table only includes factor loadings above the threshold of 0.3. KMO for the model was 0.57

sickness severity, we first explored the associations between these five variables.

3.4.1 Correlational analysis

We initially performed a correlational analysis on these proposed sickness predictor variables using jamovi v. 2.2.5. A very strong positive correlation was found between *peakDVP* and *stdDVP* (Pearson's $r = +0.94$ —see Fig. 9). Strong positive associations were also found between these two DVP measures and *scene instability* ratings. The Pearson's r values for correlations between *stdDVP* and *scene instability*, and between *peakDVP* and *scene instability*, were $+0.49$ and $+0.47$, respectively. By contrast, *αDVP* and *SpecWidthDVP* were only found to have weak negative associations with the other 3 variables. Pearson's r values were -0.11 , -0.2 , and -0.15 for correlations between *αDVP* and *stdDVP*, between *αDVP* and *peakDVP*, and between *αDVP* and *scene instability*, respectively. Similarly, Pearson's r values were -0.24 , -0.23 , and -0.29 for correlations between *SpecWidthDVP* and *stdDVP*, between *SpecWidthDVP* and *peakDVP*, and between *SpecWidthDVP* and *scene instability*, respectively. *SpecWidthDVP* was, however, found to have a stronger positive association with *αDVP* (Pearson's $r = +0.55$).

3.4.2 Exploratory factor analysis on the five predictor variables

We next performed an exploratory factor analysis (using jamovi v. 2.2.5) on these *peakDVP*, *stdDVP*, *αDVP*, *SpecWidthDVP* and *scene instability* data (see Table 1). Three of these variables loaded onto Factor 1 (*peakDVP*, *stdDVP* and *scene instability*). This first factor was therefore interpreted as representing the spatial magnitude of the sensory conflict. Factor 2 included both *αDVP* and

SpecWidthDVP. Thus, this was interpreted as representing the temporal dynamics of the sensory conflict. Together these two factors accounted for 69.1% of the overall variance—with Factor 1 (*Spatial Magnitude*) accounting for 43.0% of the variance and Factor 2 (*Temporal Dynamics*) accounting for 26.1% of the variance.

3.5 Predicting sickness severity ratings from objectively estimated DVP

According to the *DVP hypothesis*, cybersickness is triggered by large amplitude, time-varying patterns of DVP. This proposal suggests that the spatial magnitude and the temporal dynamics of the DVP should both contribute significantly to experiences of cybersickness. Thus, in the sections below, we examined the relationships between our DVP predictor variables and sickness severity ratings using linear mixed model analyses (jamovi v. 2.2.5, GAMLj package). When we inspected the correlations between our four DVP measures, we found that: (1) *peakDVP* and *stdDVP* were very highly correlated; and (2) there was a strong correlation between *αDVP* and *SpecWidthDVP* (see Fig. 9). Thus, we decided to only include *stdDVP* and *SpecWidthDVP* in the sickness prediction models below—as *stdDVP* had the highest factor loading for *spatial magnitude* and *SpecWidthDVP* had the highest factor loading for *temporal dynamics* (see Table 1).

3.5.1 Overall sickness prediction based on DVP

The sickness severity data from all four blocks of trials were first modelled in terms of *stdDVP* and *SpecWidthDVP*, with random intercepts and slopes across participants. This overall linear mixed model was found to account for 54% of the variance in sickness severity ratings:

$$\text{Sickness Severity} \sim 1 + \text{stdDVP} + \text{SpecWidthDVP} + (1 + \text{stdDVP} + \text{SpecWidthDVP} | \text{Participant}) \quad (1)$$

According to this model, sickness was significantly predicted by both *stdDVP* [$F(1, 28) = 23.5$, $p < 0.001$] and *SpecWidthDVP* [$F(1, 58.1) = 15.6$, $p < 0.001$]. Figure 10 (left) shows the modelled linear relationships between *stdDVP* and cybersickness for each participant [as well as the mean slope of $+0.34$ and the standard error of 0.071 for these relationships]. On average, an increase of 1° in *stdDVP* resulted in a 0.34 increase in sickness severity. Figure 10 (right) also shows the modelled linear relationships between *SpecWidthDVP* and cybersickness for each participant [as well as the mean slope of -3.05 and the standard error of 0.77 for these relationships]. On average, an increase of 0.1

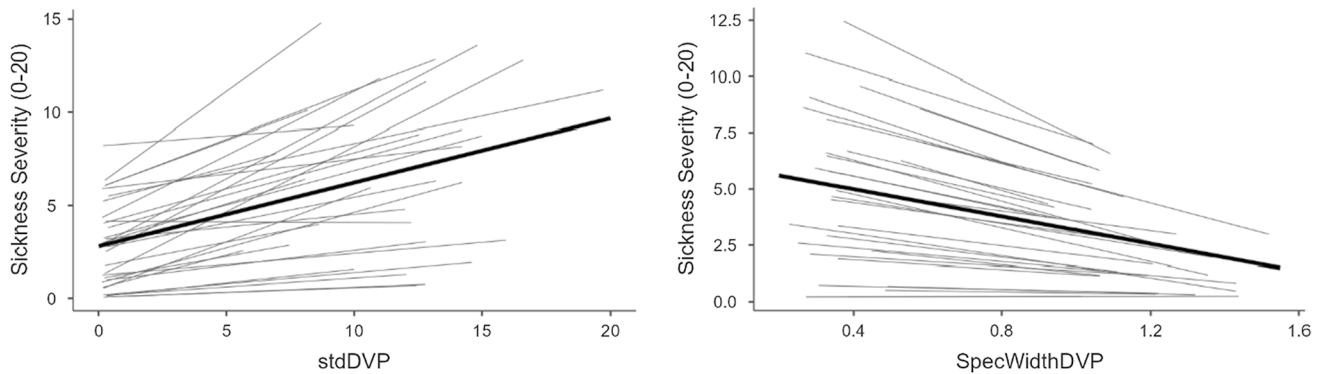


Fig. 10 Cybersickness predictions based on DVP [using data from all 4 blocks]. (Left) Shows the significant positive linear relationship between *stdDVP* and sickness severity. (Right) Shows the significant

negative linear relationship between *SpecWidthDVP* and sickness severity. Mean slopes are provided (bold lines) and random effects are plotted by participant (all other lines)

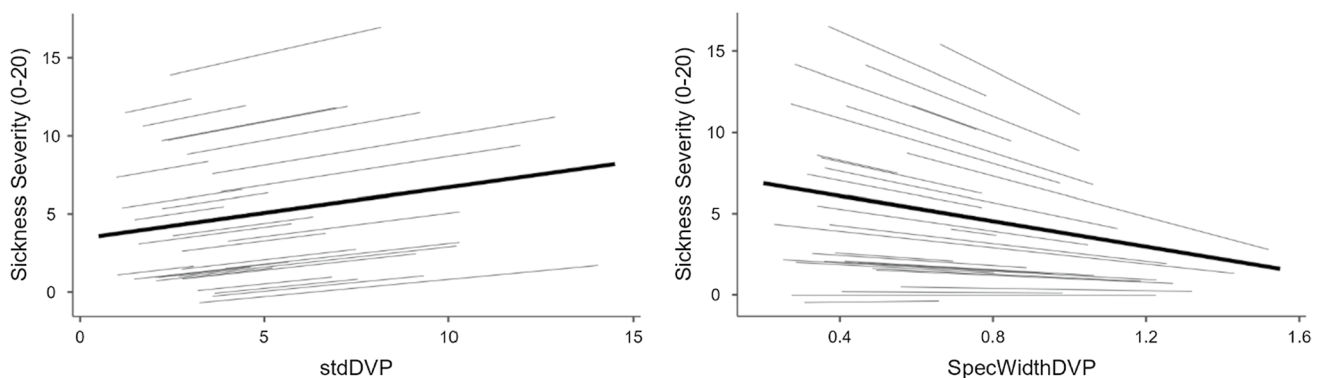


Fig. 11 Cybersickness predictions based on DVP in the *Pitch-Lag + Pitch-Move* block. (Left) Shows the significant positive linear relationship between *stdDVP* and sickness severity for these trials. (Right) Shows the significant negative linear relationship between

SpecWidthDVP and sickness severity for the same trials. Mean slopes are provided (bold lines) and random effects are plotted by participant (all other lines)

in *SpecWidthDVP* resulted in a 0.31 decrease in sickness severity.

3.5.2 Block-based sickness predictions based on DVP

Below we also examined sickness predictions based on Eq. (1) for each of the 4 experimental blocks separately.

3.5.2.1 Predicting sickness based on DVP in the Pitch-Lag + Pitch-Move block According to both the *DVP hypothesis* and the *subjective vertical conflict theory*, DVP should have been provocative for sickness on the laggy trials in this block (as the lag was added along the same axis as the instructed pitch head movements). Thus, we re-examined the relationships between our DVP measures and cybersickness using data from this block alone. When we applied Eq. (1) to the *Pitch-Lag + Pitch-Move* data, the model was found to account for 76% of the variance in participants' sickness ratings. Cybersickness was again significantly

predicted by both *stdDVP* [$F(1, 62.2)=6.47, p < 0.01$] and *SpecWidthDVP* [$F(1, 37.6)=11.30, p < 0.002$]. Figure 11 (left) shows the significant positive relationship between *stdDVP* and sickness severity [the mean slope was +0.33 and the standard error was 0.13 for this relationship]. Figure 11 (right) also shows the significant negative relationship between *SpecWidthDVP* and sickness severity [the mean slope was -3.90 and the standard error was 1.16 for this relationship].

3.5.2.2 Predicting sickness based on DVP in the Yaw-Lag + Yaw-Move block According to the *DVP hypothesis* (but not the *subjective vertical conflict theory*), DVP should have been provocative for sickness on the laggy trials in this block (as the lag was added along the same axis as the instructed yaw head movements). When we applied Eq. (1) to only these *Yaw-Lag + Yaw-Move* data, the model was found to account for 73% of the variance in participants' sickness ratings. Figure 12 (right) shows the sig-

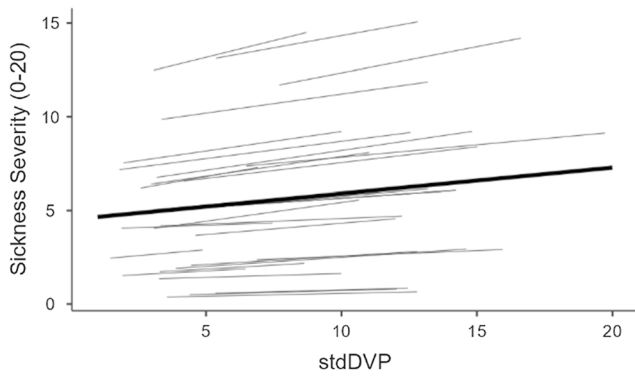
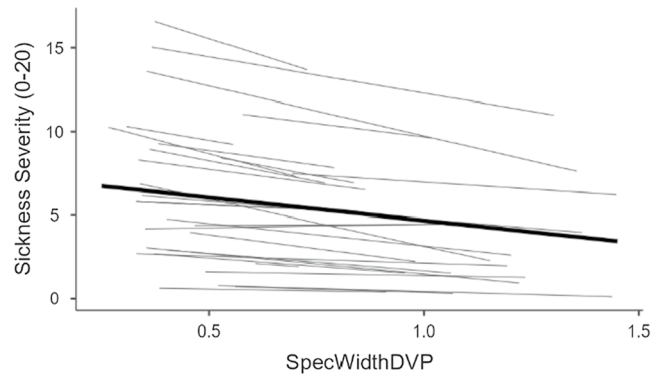


Fig. 12 Cybersickness predictions based on DVP in the *Yaw-Lag + Yaw-Move* block. (Left) Shows the non-significant relationship between *stdDVP* and sickness severity for these trials. (Right) Shows



the significant negative relationship between *SpecWidthDVP* and sickness severity for the same trials. Mean slopes are provided (bold lines) and random effects are plotted by participant (all other lines)

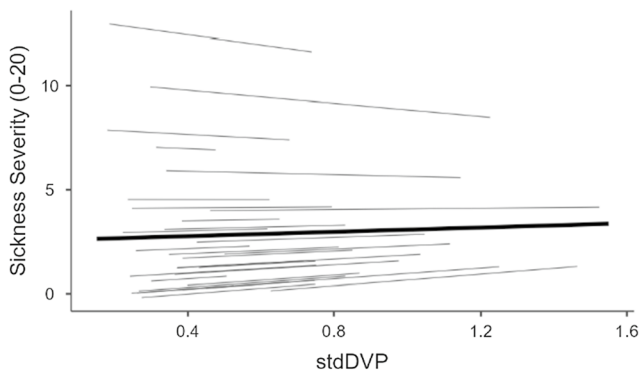
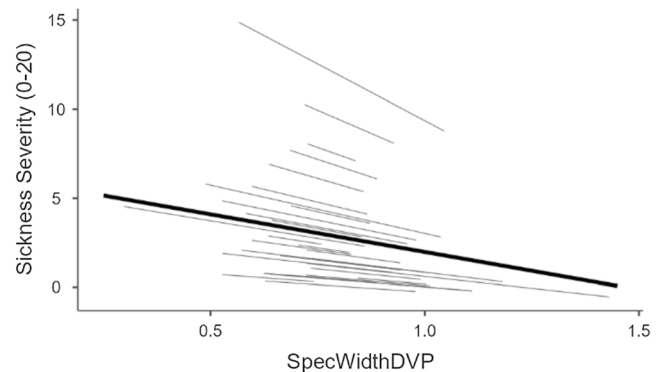


Fig. 13 Cybersickness predictions based on DVP in the *Pitch-Lag + Yaw-Move* block. (Left) Shows the non-significant relationship between *stdDVP* and sickness severity for these trials. (Right) Shows



the significant negative relationship between *SpecWidthDVP* and sickness severity. Mean slopes are provided (bold lines) and random effects are plotted by participant (all other lines)

nificant negative relationship between *SpecWidthDVP* and sickness severity [the mean slope was -2.73 and the standard error was 1.13 for this relationship], $F(1, 10.9)=5.88$, $p=0.03$. However, the relationship between *stdDVP* and sickness severity failed to reach significance in this case, $F(1, 41.1)=3.45$, $p=0.07$ (see Fig. 12, left).

3.5.2.3 Predicting sickness in the Pitch-Lag+Yaw-Move block According to the *subjective vertical conflict theory*, DVP might have also been provocative for sickness on the laggy trials in this block as well (as the user's involuntary pitch head movements would have produced some pitch DVP with this type of lag). When we applied Eq. (1) to only these *Pitch-Lag + Yaw-Move* data, the model was found to account for 63% of the variance in participants' sickness ratings. Figure 13 (right) shows the significant negative relationship between *SpecWidthDVP* and sickness severity [the mean slope was -4.23 and the standard error was 1.69 for this relationship], $F(1, 77.8)=6.301$, $p=0.014$. How-

ever, the relationship between *stdDVP* and sickness severity failed to reach significance, $F(1, 102.9)=0.269$, $p=0.605$ (see Fig. 13, left).

3.5.2.4 Predicting sickness in the Yaw-Lag+Pitch-Move block It was not expected that DVP would be provocative in the *Yaw-Lag + Pitch-Move* block. As expected, when we applied Eq. (1) to only these data, sickness severity was not significantly predicted by either *stdDVP* [$F(1, 41.1)=2.35$, $p=0.133$] or *SpecWidthDVP* values [$F(1, 21.3)=1.92$, $p=0.181$].

3.6 Predicting sickness severity ratings from perceived scene instability

Perceived scene instability might provide confirmatory evidence of provocative DVP/sensory conflicts (see Kim et al. 2020). Thus, we also examined relationships between

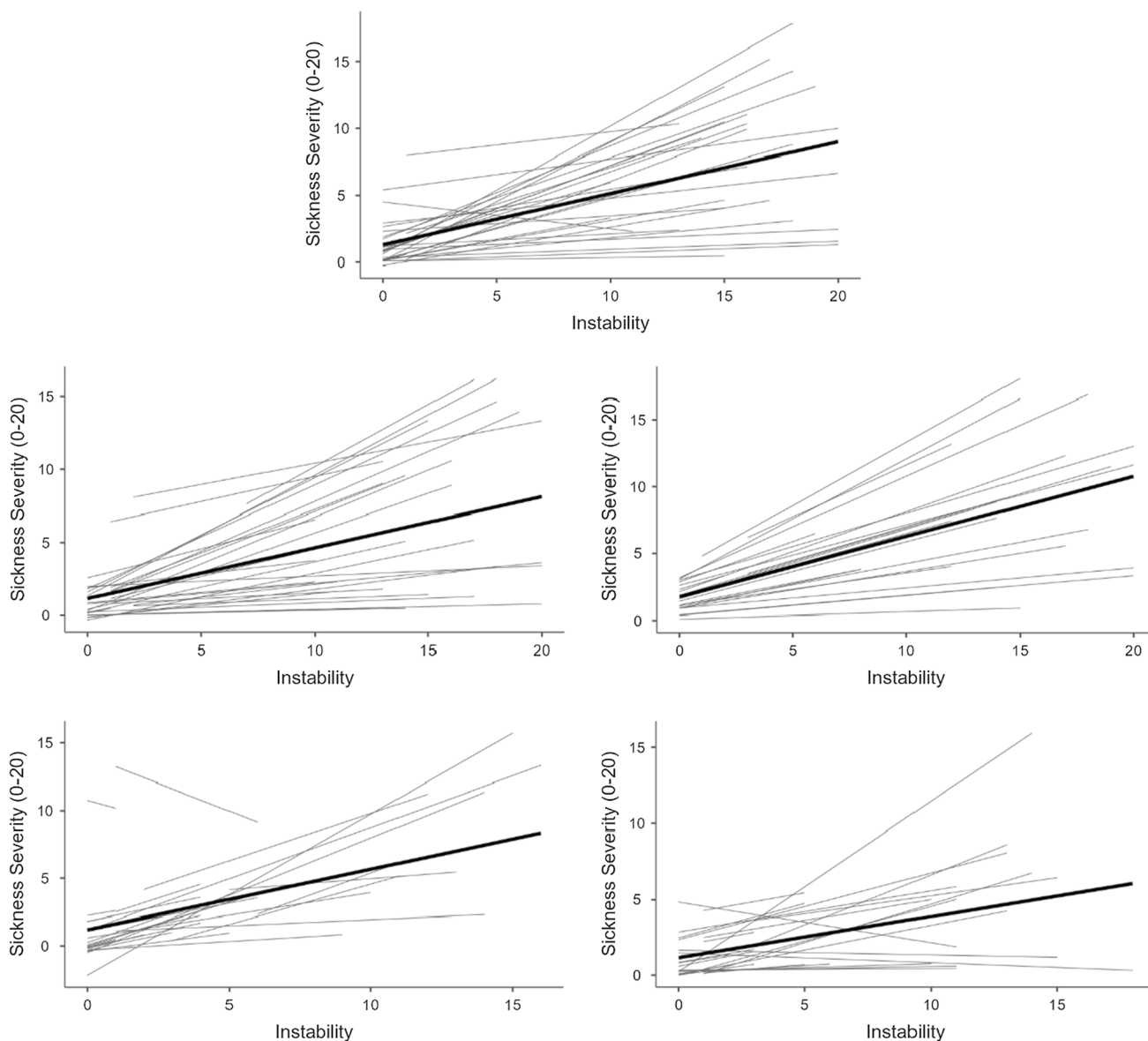


Fig. 14 These plots show (mostly) positive relationships between *scene instability* and sickness severity ratings for participants in: (1) all four blocks (top), (2) the *Pitch-Lag + Pitch-Move* block (middle Left), (3) the *Yaw-Lag + Yaw-Move* block (middle right), (4)

the *Pitch-Lag + Yaw-Move* block (bottom left); and (5) the *Yaw-Lag + Pitch-Move* block (bottom right). Mean slopes are provided (bold lines) and random effects are plotted by participant (all other lines)

scene instability and sickness severity using linear mixed model analyses (jamovi v. 2.2.5, GAMLj package).

3.6.1 Overall sickness prediction based on perceived scene instability

The linear mixed model below, which has random intercepts and slopes, was found to account for 62.1% of the variance in sickness severity ratings:

$$\text{Sickness Severity} \sim 1 + \text{SceneInstability} + (1 + \text{SceneInstability} | \text{Participant}) \tag{2}$$

This model revealed a significant positive relationship between *scene instability* and sickness ratings, $F(1, 27.8) = 38.0, p < 0.001$. Figure 14 (top) shows the positive relationships between them for (all but one) of our 30 participants [the mean slope was +0.39 and the standard error was 0.06 for this relationship]. On average, an increase in *scene instability* ratings of 1 (out of 20) was accompanied by a 0.39 increase in sickness severity ratings. This model used data from all four blocks of trials. That is, unlike the DVP-based models presented in Sect. 3.5, it also included

rating data from the baseline lag conditions (i.e. as well as data from the 89 and 222 ms added lag conditions).

3.6.2 Block-based sickness predictions based on scene instability

When we examined cybersickness predictions separately for each block, Eq. (2) revealed significant positive relationships between *scene instability* and sickness severity for all four HEAD MOVEMENT AXIS by LAG AXIS combinations [i.e. *Pitch-Lag + Pitch-Move* conditions, $F(1, 27.7) = 27.1, p < 0.001$; *Yaw-Lag + Yaw-Move* conditions, $F(1, 25.2) = 50.2, p < 0.001$; *Pitch-Lag + Yaw-Move* conditions, $F(1, 24.8) = 18.6, p < 0.001$; *Yaw-Lag + Pitch-Move* conditions, $F(1, 25.8) = 8.71, p < 0.007$]. In each case (see Fig. 14, middle left to bottom right), the models were found to explain between 60 and 86% of the variance in our participants' sickness ratings.

4 Discussion

In this study, our participants were asked to make continuous fast or slow head movements in either pitch or yaw while in HMD VR. Checks confirmed that their physical movements were primarily along the axis of the instructed head movement and close to the requested head speed for each trial. On different trials, extra display lag (0, 89, or 222 ms) was applied to the VR simulation along either the same or an orthogonal axis to the instructed head movement. This allowed us to selectively increase the sensory input conflict (or DVP) associated with one type of voluntary/involuntary head movement (e.g. head rotation in pitch), while leaving display updates based on all other head movements unaffected. Thus, in the study, we examined the effects of both lag magnitude and head speed in the following four conditions: *Pitch-Lag + Pitch-Move*, *Yaw-Lag + Yaw-Move*, *Pitch-Lag + Yaw-Move* and *Yaw-Lag + Pitch-Move*. Using these different experimental manipulations, we tested past proposals that: (1) laggy pitch head movements should be more provocative for cybersickness than laggy yaw head movements; and (2) sickness severity should scale with the spatial magnitude and the temporal dynamics of the HMD user's DVP. While we generally found the expected scaling relationships between our participants' objectively estimated DVP and their sickness severity ratings, there was little evidence that laggy pitch head movements were more provocative for cybersickness than laggy yaw head movements during HMD VR. Instead of pitch DVP being more provocative for cybersickness than yaw DVP, both types of DVP appeared to produce similar experiences of cybersickness in the current study. These results are discussed in more detail in the following sections.

4.1 Testing the subjective vertical conflict theory

Contrary to the predictions of the *subjective vertical conflict theory*: (1) *Yaw-Lag + Yaw-Move* conditions were still found to induce significant cybersickness in our participants (despite generating no more pitch or roll DVP than the baseline conditions); (2) sickness severity in these *Yaw-Lag + Yaw-Move* conditions was similar to that found in *Pitch-Lag + Pitch-Move* conditions (which were supposed to be the most provocative conditions in our study according to this theory, because they generated the most pitch DVP); and (3) sickness was not more severe in *Pitch-Lag + Yaw-Move* conditions compared to *Yaw-Lag + Pitch-Move* conditions (even though the former conditions generated more pitch DVP than the baseline conditions, whereas the latter conditions did not). As expected based on this theory, we did find significant linear relationships between pitch DVP and sickness severity in both the *Pitch-Lag + Pitch-Move* and *Pitch-Lag + Yaw-Move* conditions. However, we also found significant linear relationships between sickness severity and yaw DVP in the *Yaw-Lag + Yaw-Move* conditions as well. This latter finding is difficult to explain based on the *subjective vertical conflict theory* due to our use of 1-DOF yaw lag.¹¹ The greater sickness in these *Yaw-Lag + Yaw-Move* conditions (compared to the baseline conditions) must have been due to the increase in yaw DVP. However, according to the *subjective vertical conflict theory*, an increase in yaw DVP should not be provocative for cybersickness (only increases in pitch or roll DVP should induce, and exacerbate, this sickness).

4.2 Testing the DVP hypothesis

While our experimental findings were not supportive of the *subjective vertical conflict theory*, they were largely consistent with the *DVP hypothesis*. Cybersickness was quite similar (not markedly different) during the pitch and yaw head rotation conditions in this study. The severity of this sickness generally increased with both the participant's head speed and the amount of display lag added to their VR simulation. This was expected based on the *DVP hypothesis*, since both increases in head speed and display lag tended to increase the spatial magnitude of the DVP for the trial—as can be seen in the peak and standard deviation DVP data provided in Figs. 7 and 8 (top left and right). Both of these experimental manipulations also significantly altered the temporal dynamics of our participants' DVP. As can be seen in Figs. 7 and 8 (middle right), the multifractal spectrum widths of

¹¹ Although it might be possible to explain these findings based on a variant of this theory—the *subjective vertical-horizontal conflict theory* (see Khalid et al. 2011a, 2011b).

the DVP decreased as both the head speed and display lag magnitude increased—meaning that participants' DVP generally exhibited less multifractality when they made faster head movements and were exposed to more display lag. In addition, the DFA α data in Fig. 7 (middle left) show that DVP switched from being persistent to anti-persistent as the amount of added display lag increased from 89 to 222 ms. Thus, it was likely that the effects of display lag and head speed on the temporal dynamics of the DVP also contributed significantly to the sickness experienced in this study (i.e. in addition to the effects of these experimental manipulations on the spatial magnitude of the DVP). This possibility will be examined later in Sect. 4.2.1.

Also as expected based on the *DVP hypothesis*, cybersickness was found to be more severe in conditions where the 1-DOF display lag was added along the same axis as the instructed head movement (i.e. the 'congruent', as opposed to 'incongruent', head-and-lag axis conditions).¹² This interaction, shown in Fig. 6 (bottom left), can be explained by the patterns of DVP produced by the trials in the four different experimental blocks. In the top rows of Figs. 7 and 8, it can be seen that the *Yaw-Lag + Yaw-Move* and the *Pitch-Lag + Pitch-Move* conditions produced the largest amplitude, most variable patterns of DVP. The peaks and the standard deviations of the DVP in these 'congruent' conditions were considerably larger than those in the other 'incongruent' conditions (i.e. the *Yaw-Lag + Pitch-Move* and the *Pitch-Lag + Yaw-Move* conditions). Thus, according to our *DVP hypothesis*, cybersickness was more severe in the 'congruent' conditions because the spatial magnitudes of their sensory input conflicts were larger than those in the 'incongruent' conditions. The effects of lag magnitude on the temporal dynamics of the DVP were also noticeably different for 'congruent' and 'incongruent' conditions. For example, the multifractal spectrum widths of the DVP were significantly lower for the 'congruent' conditions (see Fig. 7, middle right). The DFA α of the DVP was also significantly lower for these 'congruent' conditions when 222 ms lag was added (see Fig. 7, middle left). Thus, the DVP generally displayed less multifractality and were more likely to be anti-persistent in the 'congruent' (compared to the 'incongruent') conditions. These observed differences in *SpecWidthDVP* and αDVP therefore provide another possible DVP-based explanation for why cybersickness was more severe in the 'congruent' conditions (in addition to the DVP in these

conditions having larger peaks and standard deviations than those in the 'incongruent' conditions).

While display lag effects on cybersickness were consistently found to be more provocative in the 'congruent' (compared to the 'incongruent') conditions, we did not find a similar 3-way interaction involving head speed. As expected, the spatial magnitude of the DVP and the severity of the cybersickness both increased with the participant's head speed. However, these effects did not appear to be significantly worse in the 'congruent', compared to the 'incongruent', conditions – see Fig. 6 (bottom right). Sickness was more severe in the *Yaw-Lag + Yaw-Move* and *Pitch-Lag + Pitch-Move* conditions. However, the increases in this sickness with head speed were quite similar in 'congruent' ($M = 1.6$) and 'incongruent' ($M = 1.3$) conditions. One possible reason for this unexpected null finding was that: (1) head speed was manipulated within blocks in this experiment; (2) sickness severity was found to increase with trial number in each block; and (3) head speeds were requested by the experimenter and then self-generated by the participant. Thus, the modest cross trial sickness contamination observed in this study, and the individual differences in participants' self-generated head speed, might have obscured the expected 3-way interaction for sickness involving head speed.

4.2.1 Can DVP be used to predict the severity of cybersickness?

In the above section, we described how the objectively estimated DVP could explain the cybersickness findings of the current study. However, to be most useful, our theory actually needs to be able to predict the HMD user's experience of cybersickness. Thus, in order to assess the predictive power of the *DVP hypothesis*, we examined the relationships between the objectively estimated DVP and our participants' sickness severity ratings using linear mixed model analyses. We considered four potential DVP-based predictors of cybersickness: the peak, the standard deviation, the DFA α , and the multifractal spectrum width of the DVP. When these DVP measures were examined on their own, using the data from all four blocks of trials, our preliminary analyses confirmed that they were each capable of predicting a significant amount of variance in our participants' sickness severity ratings.¹³ However, when we inspected the associations between these four DVP measures, we found that *peakDVP* was very highly correlated with *stdDVP*, and αDVP was also significantly correlated with *SpecWidthDVP*. So, in order to avoid problems with multicollinearity, we decided

¹² Participant debriefing confirmed that when lag was applied along the orthogonal axis to the instructed head-movement, it was often not detected at all. In cases where it was detected, its effects on the simulation were perceived to be relatively minor. This was in stark contrast to conditions where the lag was applied along the same axis as the instructed head-movement — these effects were often very large and obvious to all participants.

¹³ When *peakDVP* was included as the sole predictor in a linear mixed model it significantly predicted sickness severity, $F(1,23.7) = 25.4, p < 0.001$. Similarly, when αDVP was included as

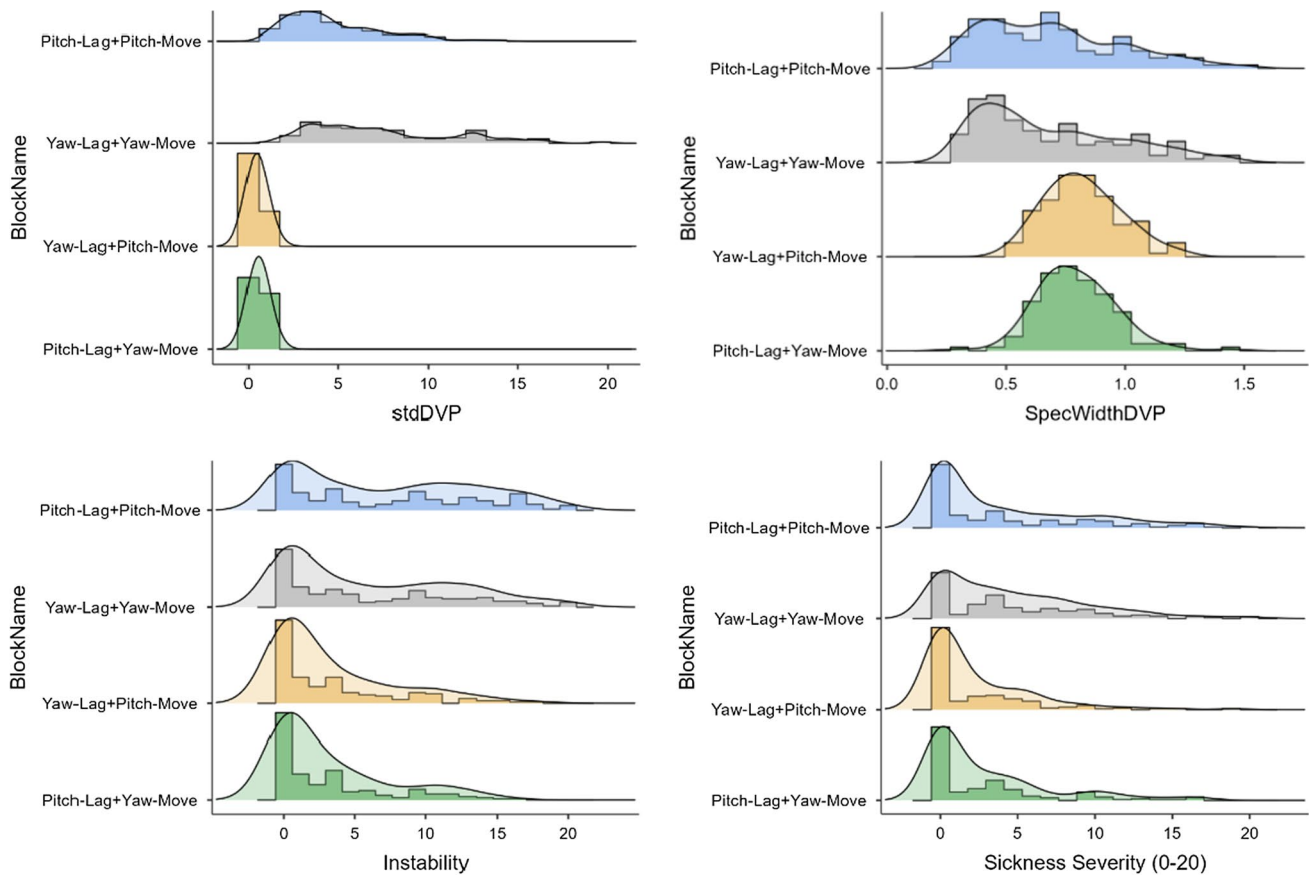


Fig. 15 Histograms and density plots showing the differences in *stdDVP*, *SpecWidthDVP*, *scene instability* and sickness severity ratings, across the 4 different blocks of trials

to only include *stdDVP* (an index of the spatial magnitude of the DVP) and *SpecWidthDVP* (an index of the temporal dynamics of the DVP) in our sickness prediction models (i.e. *peakDVP* and *α DVP* were excluded from further analysis).

We first examined overall sickness predictions based on the DVP. When data were included from all four blocks of trials, a linear mixed model based on *stdDVP* and *SpecWidthDVP* was found to account for 54% of the variance in our participants' sickness ratings. We found a significant positive linear relationship between *stdDVP* and sickness severity, as well as a significant negative linear relationship between *SpecWidthDVP* and sickness severity. This confirmed that cybersickness was influenced by both the spatial magnitude and the temporal dynamics of the DVP. However, it is important to note that there were also substantial individual differences in these DVP-sickness relationships—as

is clearly shown in Figs. 10, 11, 12 and 13, where the random effects in each case were plotted by participant.

We next examined sickness predictions based on DVP for each of the four experimental blocks separately. The predictive performance of the model was found to improve substantially when we focussed only on the data from the (more provocative) 'congruent' conditions. Using the same two DVP measures, this model predicted 76% of the variance in sickness ratings during the *Pitch-Lag + Pitch-Move* conditions (based on significant relationships involving both *SpecWidthDVP* and *stdDVP*). This model also predicted 73% of the variance in sickness ratings during the *Yaw-Lag + Yaw-Move* conditions (based in this case on only a significant negative relationship involving *SpecWidthDVP*). However, sickness predictions based on DVP appeared to be less reliable for the 'incongruent' conditions. While the model was still found to predict 63% of the variance in sickness ratings during the *Pitch-Lag + Yaw-Move* conditions (based again on a significant negative relationship involving *SpecWidthDVP*), neither of the two DVP measures were found to significantly predict sickness severity during the 'incongruent' *Yaw-Lag + Pitch-Move* conditions.

Footnote 13 (continued)

the sole predictor it also significantly predicted sickness severity, $F(1,37.3) = 15.1, p < 0.001$.

Why were sickness predictions based on DVP significant for both ‘congruent’ conditions, but only one of the two ‘incongruent’ conditions, in the current study? As can be seen from the top row of Fig. 15, the descriptive statistics of *stdDVP* and *SpecWidthDVP* were quite different for ‘congruent’ and ‘incongruent’ conditions. While the maximum *stdDVP* was either 14° or 20° in the *Pitch-Lag + Pitch-Move* and *Yaw-Lag + Yaw-Move* conditions, it was only 2° in the *Pitch-Lag + Yaw-Move* and *Yaw-Lag + Pitch-Move* conditions (see Fig. 15, top left). It is therefore not surprising that the ‘incongruent’ conditions, with their less variable DVP, produced many more ‘no sickness’ trials than the ‘congruent’ conditions (see Fig. 15, bottom right). Also, while the *SpecWidthDVP* and sickness severity distributions appeared similar for the two ‘congruent’ blocks, they were markedly different for the ‘incongruent’ blocks (see Fig. 15, top and bottom right). Presumably these large differences in spatial magnitudes and temporal dynamics explain why the relationships between DVP and cybersickness were robust and reliable for the ‘congruent’ experimental blocks, but not always present for the other ‘incongruent’ blocks, in the study.

4.2.2 What are the relationships between DVP, perceived scene instability, and cybersickness severity?

When display lag generates large-amplitude, highly variable DVP during active HMD VR, this is often accompanied by user perceptions of scene instability. Thus, Kim et al. (2020) proposed that perceived scene instability might provide confirmatory evidence of potentially provocative sensory conflicts. Consistent with this idea, *scene instability* ratings were found to be strongly correlated with both *peakDVP* and *stdDVP* (but not *αDVP* and *SpecWidthDVP*) in this study (see Fig. 9). In an exploratory factor analysis (see Table 1), where a two-factor structure appeared to best fit the data, we found that *peakDVP*, *stdDVP* and *scene instability* all loaded onto the same factor. This suggested that our participants were consciously aware of the spatial magnitude of their DVP—with detectable increases in *peakDVP* and *stdDVP* being perceived as increases in *scene instability*. Our participants did not, however, appear to be consciously aware of important differences in the temporal dynamics of their DVP (which must therefore have only been available to their pre-conscious processing). In other words, their *scene instability* ratings appeared to primarily reflect their conscious experience of the spatial magnitude of the DVP (not its temporal dynamics). Nevertheless, we still found significant positive linear relationships between *scene instability* and sickness severity in this study. When data were used

from all four blocks of trials, a linear mixed model based only on *scene instability* was found to account for 62% of the variance in participants’ sickness ratings (see Fig. 14, top). This was somewhat better than the predictions provided by our model based on the objectively estimated DVP, which accounted for only 54% of the variance in sickness severity ratings when data from all four blocks was used. Unlike our sickness predictions based on *stdDVP* and *SpecWidthDVP* (which were only significant for the *Pitch-Lag + Pitch-Move*, *Yaw-Lag + Yaw-Move*, and *Pitch-Lag + Yaw-Move* blocks), these predictions based on *scene instability* remained significant across all four blocks (including the *Yaw-Lag + Pitch-Move* block; see Fig. 14 from middle left to bottom right).

The above results suggest that our participants’ subjective ratings of *scene instability* were more reliable predictors of their cybersickness than the DVP measures that we calculated (which in this specific case were the *stdDVP* and the *SpecWidthDVP*). However, the DVP produced by the experimental trials in this study were complex time series data. Thus, it was possible that these *scene instability* ratings were based on, or alternatively influenced by, other information about the DVP (beyond that provided by *peakDVP*, *stdDVP*, *αDVP* and *SpecWidthDVP* measures). These subjective ratings of *scene instability* might also have been influenced by individual differences in our participants’ susceptibility to motion sickness. Information about their motion sickness susceptibility (e.g. Golding 1998; Golding et al. 2021) was not obtained in our study and therefore not included in our DVP-based sickness prediction models. However, it is possible that participants who were more susceptible to motion sickness were also more sensitive to DVP. That is, given the same pattern of DVP, those participants would rate it as producing more *scene instability*, and if cybersickness was induced, then they would tend to experience it more severely. This could therefore be another reason why sickness predictions based on *scene instability* appeared to be more reliable than those based on the objectively estimated DVP.

However, there is another possible reason why sickness predictions based on *scene instability* were more reliable in the current study. By necessity, the analyses which examined sickness predictions based on *stdDVP* and *SpecWidthDVP* only used data from a subset of the experimental trials.¹⁴ Specifically, they only used the DVP and sickness severity data from the 89 ms and 222 ms added lag trials. This was because we were not able to estimate the DVP time series data for the baseline conditions (it was not appropriate to use *peakDVP* and *stdDVP* values of 0, or *αDVP* and

¹⁴ There were 720 trials in total (our 30 participants were each exposed to 24 different experimental trials). Predictions based on DVP were based only on data from 480 of these 720 trials. By contrast, predictions based on *scene instability* were based on data from all 720 trials.

SpecWidthDVP values of NaN, for these baseline conditions).¹⁵ By contrast, the analyses involving *scene instability* included data from all of the experimental trials (they included instability and sickness ratings for the baseline trials, as well as for the 89 and 222 ms added lag trials). Thus, predictions based on *scene instability* modelled the differences in sickness between the no added lag and 89 ms added lag conditions, as well as the differences between the 89 ms and the 222 ms added lag conditions. It is therefore not surprising that models based on *scene instability*, which included 33% more data and a wider range of data values, provided more reliable and robust sickness predictions than models based only on the DVP.

4.3 Limitations of the study

As can be seen in Fig. 5, while self-generated head speeds were similar, participant head movement amplitudes were smaller in the *Pitch-Move*, compared to the *Yaw-Move*, conditions. Ideally, we wanted these pitch and yaw head movement amplitudes to be the same in order to produce similar magnitudes of pitch and yaw DVP. As a result of these differences in head movement amplitude, *peakDVP* and *stdDVP* were smaller on average in the *Pitch-Lag + Pitch-Move*, compared to the *Yaw-Lag + Yaw-Move*, conditions. This might have been why pitch DVP was not more provocative for cybersickness than yaw DVP in the current study (see Fig. 6, bottom left). Pitch and yaw head movement amplitudes could perhaps be better controlled in future studies by having participants actively track a visually moving target by making head-movements (this target could move up-and-down, or left-and-right, on the HMD's screens by the same distance on *Pitch-Move* or *Yaw-Move* trials). Despite this problem with head-movement amplitudes in the current study, it is clear that—contrary to the predictions of *subjective vertical conflict theory*—pure yaw DVP was still capable of inducing cybersickness. Our selective use of 1-DOF display lag ensured that *Pitch-Lag + Pitch-Move* trials only injected pitch DVP, and *Yaw-Lag + Yaw-Move* trials only injected yaw DVP, into the VR system. Even though laggy *Yaw-Lag + Yaw-Move* trials contained no more pitch DVP or roll DVP than baseline trials, they were still found to produce functionally significant sickness (see Stanney et al. 2014).

Ideally for modelling predictive relationships between DVP and cybersickness severity, it would be useful to also include data from the baseline (i.e. no added lag) conditions (as we did when we examined the relationships between

perceived scene instability and cybersickness in Sect. 3.6). In this study, we estimated the average effective display lag for these baseline conditions to be only ~4 ms. However, we did not obtain information about how the display lag in these baseline conditions changed over time [which it would have—see Wu et al (2013), Stauffert et al. (2018)]. Estimating the DVP time series data for a trial requires information about the magnitude of its display lag, and the HMD user's head velocity, at each instant. Because we did not have information about the typical fluctuations in baseline lag over time, we could not estimate the DVP data for these baseline trials. Thus, we were not able to calculate the *stdDVP* and the *SpecWidthDVP* for those trials and use those estimated data (as well as the data from the experimental added lag trials) to predict cybersickness severity. Future studies could conduct a fine-grained analysis of the natural fluctuations in display lag for their particular HMD VR system under baseline conditions. When this information is paired with the HMD user's actual head movements during baseline trials, it could then be used to estimate the (mostly non-provocative) DVP data experienced during those trials. Including this extra baseline data in our models might result in much more reliable and robust sickness predictions based on DVP.

The current findings suggest that increases in display lag were more provocative when the DVP directly affected our participants' online control of their intended (as opposed to involuntary) head movements. However, the laggy 'congruent' conditions in our study could have further exacerbated our participants' adverse experiences in VR by increasing the difficulty and workload associated with their continuous head-movement task (see Jasper et al. 2023; Law et al. 2019; Sepich et al. 2022). These differences in task difficulty and workload might (at least partially) explain why cybersickness was less severe in the 'incongruent' conditions (e.g. where the lag-induced DVP was primarily generated by involuntary head movements that were unrelated to the participant's task performance). Thus, future research might need to find a way to manipulate and isolate pitch and yaw DVP that does not systematically alter the HMD user's task difficulty or workload.

5 Conclusions

In recent years, technology developers have made considerable efforts to reduce the amount of display lag in HMD VR. However, despite their innovations and improvements (e.g. Van Waveren 2016), some display lag still remains in modern VR systems [due to sensing, processing, data smoothing, transmission, rendering and frame rate delays—see Stauffert et al. (2018)]. This study examined the effects of display lag and head speed (and the DVP that they generate) on experiences of cybersickness during

¹⁵ Estimating the DVP in baseline (i.e. no added lag) conditions is still a work in progress. We plan to provide estimates of the DVP for these conditions in our future studies.

active HMD VR. Its findings generally support the *DVP hypothesis* for cybersickness, rather than the *subjective vertical conflict theory* of motion sickness. Contrary to the *subjective vertical conflict theory*, *Yaw-Lag + Yaw-Move* conditions were still found to induce significant cybersickness, and this sickness was similar (not less) than that induced in *Pitch-Lag + Pitch-Move* conditions. Both findings were, however, expected based on the *DVP hypothesis*. Importantly, they suggest that VR developers may not need to avoid situations where HMD users are required to make pitch (or roll) head movements.

In this study, we replicate and extend, our recent findings that cybersickness due to display lag scales with the objectively estimated DVP (see also Kim et al. 2020; Palmisano et al. 2020, 2023). When extra lag was selectively added to display updates along the same axis as the instructed head movement, DVP was found to predict between 73 and 76% of the variance in our participants' cybersickness severity ratings. As expected from the *DVP hypothesis*, the spatial magnitude and the temporal dynamics of these DVP data were both found to contribute significantly to these sickness predictions. Specifically, cybersickness severity was found to increase as the standard deviation of the DVP increased, and as the multifractal spectrum width of the DVP decreased.

The conscious experiences produced by these DVP data—i.e. user perceptions of scene instability—were also found to significantly predict sickness severity ratings in this study. Indeed, these ratings of perceived scene instability appeared to provide somewhat more reliable and robust predictions of cybersickness than our objective estimates of the DVP. These scene instability ratings were more strongly associated with the spatial magnitude of the DVP (rather than its temporal dynamics).

In this study, we have shown that objective estimates of the DVP, and subjective estimates of scene instability, can be used to predict the severity of cybersickness experienced by HMD users. The next step will be to identify the patterns of DVP (or scene instability) that occur just before the onset of cybersickness. Once identified, these patterns should have direct applications in helping to reduce the incidence (as well as the severity) of cybersickness in HMD VR. Taken together, the current findings suggest that it should be possible to: (1) further improve sickness predictions using objective estimates of the DVP; and (2) identify many provocative conditions for cybersickness based on the HMD user's subjective reports of scene instability.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10055-023-00909-6>.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. This research was supported by an Australian Research Council (ARC) Discovery Project (DP210101475).

Availability of data and materials Data will be made available upon request.

Code availability Code can be made available upon request.

Declarations

Conflict of interest None.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allison RS, Harris LR, Jenkin M, Jasiobedzka U, Zacher JE (2001) Tolerance of temporal delay in virtual environments. In: Proceedings of IEEE virtual reality. IEEE, Yokohama, pp 247–254. <https://doi.org/10.1109/VR.2001.913793>
- Bles W, Bos JE, De Graaf B, Groen E, Wertheim AH (1998) Motion sickness: only one provocative conflict? *Brain Res Bull* 47(5):481–487. [https://doi.org/10.1016/S0361-9230\(98\)00115-4](https://doi.org/10.1016/S0361-9230(98)00115-4)
- Bles W, Bos JE, Kruit H (2000) Motion sickness. *Curr Opin Neurol* 13(1):19–25. <https://doi.org/10.1097/00019052-200002000-00005>
- Bonato F, Bubka A, Story M (2005) Rotation direction change hastens motion sickness onset in an optokinetic drum. *Aviat Sp Environ Med* 76(9):823–827
- Bos JE, Bles W (1998) Modelling motion sickness and subjective vertical mismatch detailed for vertical motions. *Brain Res Bull* 47(5):537–542. [https://doi.org/10.1016/S0361-9230\(98\)00088-4](https://doi.org/10.1016/S0361-9230(98)00088-4)
- Bos JE, Bles W (2002) Theoretical considerations on canal–otolith interaction and an observer model. *Biol Cybern* 86(3):191–207. <https://doi.org/10.1007/s00422-001-0289-7>
- Bos JE, Bles W, Groen EL (2008) A theory on visually induced motion sickness. *Displays* 29(2):47–57. <https://doi.org/10.1016/j.displa.2007.09.002>
- Bubka A, Bonato F (2003) Optokinetic drum tilt hastens the onset of vection-induced motion sickness. *Aviat Space Environ Med* 74:315–319
- Caserman P, Martinussen M, Göbel S (2019) Effects of end-to-end latency on user experience and performance in immersive virtual reality applications. In: Proceedings of the joint international conference on entertainment computing and serious games. Springer, Arequipa, pp 57–69. https://doi.org/10.1007/978-3-030-34644-7_5
- Clifton J, Palmisano S (2019) Effects of steering locomotion and teleporting on cybersickness and presence in HMD-based virtual reality. *Virtual Real* 24:453–468. <https://doi.org/10.1007/s10055-019-00407-8>
- de Graaf B, Bles W, Bos JE (1998) Roll motion stimuli: sensory conflict, perceptual weighting and motion sickness. *Brain Res Bull* 47(5):489–495. [https://doi.org/10.1016/S0361-9230\(98\)00116-6](https://doi.org/10.1016/S0361-9230(98)00116-6)

- Dennison MS, Wisti AZ, D'Zmura M (2016) Use of physiological signals to predict cybersickness. *Displays* 44:42–52. <https://doi.org/10.1016/j.displa.2016.07.002>
- DiZio P, Lackner JR (1997) Circumventing side effects of immersive virtual environments. In: Salvendy G, Smith MJ, Salvendy G, Koubek RJ (eds) *Advances in human factors/ergonomics*, vol 21b. Design of computing systems: social and ergonomic considerations. Elsevier, Amsterdam, pp 893–896
- Dodgson NA (2004) Variation and extrema of human interpupillary distance. In: Woods AJ, Merritt JO, Benton SA, Bolas MT (eds) *Stereoscopic displays and virtual reality systems XI*, vol 5291. International Society for Optics and Photonics. San Jose, CA, pp 36–47
- Draper MH, Viirre ES, Furness TA, Gawron VJ (2001) Effects of image scale and system time delay on simulator sickness within head-coupled virtual environments. *Hum Factors* 43:129–146. <https://doi.org/10.1518/001872001775992>
- Ebenholtz SM, Cohen MM, Linder BJ (1994) The possible role of nystagmus in motion sickness: a hypothesis. *Aviat Space Environ Med* 65(11):1032–1035
- Feng J, Kim J, Luu W, Palmisano S (2019) Method for estimating display lag in the Oculus Rift S and CV1. In: *Proceedings of SIGGRAPH Asia, SIGGRAPH*, Brisbane, Australia, pp 1–2. <https://doi.org/10.1145/3355056.3364590>
- Gavani AM, Nesbitt KV, Blackmore KL, Nalivaiko E (2017) Profiling subjective symptoms and autonomic changes associated with cybersickness. *Auton Neurosci* 203:41–50. <https://doi.org/10.1016/j.autneu.2016.12.004>
- Golding JF (1998) Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. *Brain Res Bull* 47(5):507–516. [https://doi.org/10.1016/S0361-9230\(98\)00091-4](https://doi.org/10.1016/S0361-9230(98)00091-4)
- Golding JF (2016) Motion sickness. In: Furman JM, Lempert T (eds) *Handbook of clinical neurology*, vol 137, 3rd Series Neuro-Otology. Elsevier, Amsterdam, pp 371–390. <https://doi.org/10.1016/B978-0-444-63437-5.00027-3>
- Golding JF, Rafiq A, Keshavarz B (2021) Predicting individual susceptibility to visually induced motion sickness by questionnaire. *Front Virtual Real* 2:576871. <https://doi.org/10.3389/frvir.2021.576871>
- Howarth PA, Costello PJ (1997) The occurrence of virtual simulation sickness symptoms when an HMD was used as a personal viewing system. *Displays* 18(2):107–116. [https://doi.org/10.1016/S0141-9382\(97\)00011-5](https://doi.org/10.1016/S0141-9382(97)00011-5)
- Howarth PA, Finch M (1999) The nauseogenicity of two methods of navigating within a virtual environment. *Appl Ergon* 30(1):39–45. [https://doi.org/10.1016/S0003-6870\(98\)00041-6](https://doi.org/10.1016/S0003-6870(98)00041-6)
- Ihlen EA (2012) Introduction to multifractal detrended fluctuation analysis in Matlab. *Front Physiol* 3:141. <https://doi.org/10.3389/fphys.2012.00141>
- Jasper A, Sepich NC, Gilbert SB, Kelly JW, Dorneich MC (2023) Predicting cybersickness using individual and task characteristics. *Comput Hum Behav* 146:107800. <https://doi.org/10.1016/j.chb.2023.107800>
- Jennings S, Craig G, Reid L, Kruk R (2000) The effect of visual system time delay on helicopter control. In: *Proceedings of the human factors and ergonomics society annual meeting*, vol 44, no 13. SAGE, Los Angeles, pp 69–72
- Jennings S, Reid LD, Craig G, Kruk RV (2004) Time delays in visually coupled systems during flight test and simulation. *J Aircr* 41(6):1327–1335. <https://doi.org/10.2514/1.449>
- Kennedy RS, Lilienthal MG (1994) Measurement and control of motion sickness after-effects from immersion in virtual reality. In: *Proceedings of virtual reality and medicine: the cutting edge, SIG-advanced applications*, New York, pp 111–119
- Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG (1993) Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *Int J Aviat Psychol* 3(3):203–220. https://doi.org/10.1207/s15327108ijap0303_3
- Kennedy RS, Jones MB, Lilienthal MG, Harm DL (1994) Profile analysis of after-effects experienced during exposure to several virtual reality environments. *AGARD-CP-541: virtual interfaces. Res Appl* 2:1–9
- Keshavarz B, Hecht H (2011) Validating an efficient method to quantify motion sickness. *Hum Factors* 53:415–426. <https://doi.org/10.1177/0018720811403736>
- Keshavarz B, Hecht H, Lawson BD (2014) Visually induced motion sickness: causes, characteristics, and countermeasures. In: Hale KS, Stanney KM (eds) *Handbook of virtual environments: design, implementation and applications*. CRC Press, Boca Raton, pp 652–703
- Keshavarz B, Peck K, Rezaei S, Taati B (2022) Detecting and predicting visually induced motion sickness with physiological measures in combination with machine learning techniques. *Int J Psychophysiol* 176:14–26. <https://doi.org/10.1016/j.ijpsycho.2022.03.006>
- Khalid H, Turan O, Bos JE (2011a) Theory of a subjective vertical–horizontal conflict physiological motion sickness model for contemporary ships. *J Mar Sci Technol* 16:214–225. <https://doi.org/10.1007/s00773-010-0113-y>
- Khalid H, Turan O, Bos JE, Incecik A (2011b) Application of the subjective vertical–horizontal–conflict physiological motion sickness model to the field trials of contemporary vessels. *Ocean Eng* 38(1):22–33. <https://doi.org/10.1016/j.oceaneng.2010.09.008>
- Kim K, Rosenthal MZ, Zielinski DJ, Brady R (2014) Effects of virtual environment platforms on emotional responses. *Comput Methods Programs Biomed* 113(3):882–893. <https://doi.org/10.1016/j.cmpb.2013.12.024>
- Kim J, Luu W, Palmisano S (2020) Multisensory integration and the experience of scene instability, presence and cybersickness in virtual environments. *Comput Hum Behav* 113:106484. <https://doi.org/10.1016/j.chb.2020.106484>
- Kim J, Palmisano S, Luu W, Iwasaki S (2021) Effects of linear visual-vestibular conflict on presence, perceived scene stability and cybersickness in the Oculus Go and Oculus Quest. *Front Virtual Real* 2:582156. <https://doi.org/10.3389/frvir.2021.582156>
- Kinsella A, Mattfeld R, Muth E, Hoover A (2016) Frequency, not amplitude, of latency affects subjective sickness in a head-mounted display. *Aerosp Med Hum* 87(7):604–609. <https://doi.org/10.3357/AMHP.4351.2016>
- Koslucher F, Munafo J, Stoffregen TA (2016) Postural sway in men and women during nauseogenic motion of the illuminated environment. *Exp Brain Res* 234:2709–2720. <https://doi.org/10.1007/s00221-016-4675-8>
- LaViola JJ Jr (2000) A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin* 32(1):47–56
- Law A, Jennings S, Ellis K (2019) A comparison of control activity and heart rate as measures of pilot workload in a helicopter tracking task. In: *Proceedings of the vertical flight society's 75th annual forum and technology display*. NRC Canada, Philadelphia, pp 1–9
- Lawson B, Stanney KM (2021) Editorial: cybersickness in virtual reality and augmented reality. *Front Virtual Real* 2:2759682. <https://doi.org/10.3389/frvir.2021.759682>
- McCauley ME, Sharkey TJ (1992) Cybersickness: perception of self-motion in virtual environments. *Presence: Teleoperat Vir Environ* 1:311–318. <https://doi.org/10.1162/pres.1992.1.3.311>
- Merhi O, Faugloire E, Flanagan M, Stoffregen TA (2007) Motion sickness, console video games, and head mounted displays. *Hum Factors* 49:920–934
- Moss JD, Muth ER (2011) Characteristics of head-mounted displays and their effects on simulator sickness. *Hum Factors* 53(3):308–319. <https://doi.org/10.1177/0018720811405196>

- Moss JD, Austin J, Salley J, Coats J, Williams K, Muth ER (2011) The effects of display delay on simulator sickness. *Displays* 32(4):159–168. <https://doi.org/10.1016/j.displa.2011.05.010>
- Munafo J, Diedrick M, Stoffregen TA (2017) The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. *Exp Brain Res* 235(3):889–901. <https://doi.org/10.1007/s00221-016-4846-7>
- Nooij SA, Pretto P, Oberfeld D, Hecht H, Bühlhoff HH (2017) Vection is the main contributor to motion sickness induced by visual yaw rotation: Implications for conflict and eye movement theories. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0175305>
- Oman CM (1982) A heuristic mathematical model for the dynamics of sensory conflict and motion sickness hearing in classical musicians. *Acta Otolaryngol* 94(s392):4–44. <https://doi.org/10.3109/00016488209108197>
- Oman CM (1990) Motion sickness: a synthesis and evaluation of the sensory conflict theory. *Can J Physiol Pharmacol* 68(2):294–303. <https://doi.org/10.1139/y90-044>
- Palmisano S, Constable R (2022) Reductions in cybersickness with repeated exposure to HMD-based virtual reality are game-specific. *Virtual Real* 26:1373–1389. <https://doi.org/10.1007/s10055-022-00634-6>
- Palmisano S, Szalla L, Kim J (2019) Monocular viewing protects against cybersickness produced by head movements in the Oculus Rift. In: *Proceedings of the 25th ACM symposium on virtual reality software and technology* (Sydney, NSW: ACM), pp 1–2. <https://doi.org/10.1145/3359996.3364699>
- Palmisano S, Allison RS, Kim J (2020) Cybersickness in head-mounted displays is caused by differences in the user's virtual and physical head pose. *Front Virtual Real* 1:587698. <https://doi.org/10.3389/frvir.2020.587698>
- Palmisano S, Allison RS, Teixeira J, Kim J (2023) Differences in virtual and physical head orientation predict sickness during active head-mounted display based virtual reality. *Virtual Real*. <https://doi.org/10.1007/s10055-022-00732-5>
- Peng CK, Havlin S, Stanley HE, Goldberger AL (1995) Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* 5(1):82–87. <https://doi.org/10.1063/1.166141>
- Reason JT (1978) Motion sickness adaptation: a neural mismatch model. *J R Soc Med* 71(11):819–829
- Rebenitsch L, Owen C (2016) Review on cybersickness in applications and visual displays. *Virtual Real* 20(2):101–125. <https://doi.org/10.1007/s10055-016-0285-9>
- Risi D, Palmisano S (2019) Effects of postural stability, active control, exposure duration and repeated exposures on HMD induced cybersickness. *Displays* 60:9–17. <https://doi.org/10.1016/j.displa.2019.08.003>
- Sepich NC, Jasper A, Fieffer S, Gilbert SB, Dorneich MC, Kelly JW (2022) The impact of task workload on cybersickness. *Front Virtual Real* 3:943409. <https://doi.org/10.3389/frvir.2022.943409>
- Sharples S, Cobb S, Moody A, Wilson JR (2008) Virtual reality induced symptoms and effects (VRISE): comparison of head mounted display (HMD), desktop and projection display systems. *Displays* 29:58–69. <https://doi.org/10.1016/j.displa.2007.09.005>
- St. Pierre ME, Banerjee S, Hoover, AW, Muth ER (2015) The effects of 0.2 Hz varying latency with 20–100 ms varying amplitude on simulator sickness in a helmet mounted display. *Displays* 36:1–8. <https://doi.org/10.1016/j.displa.2014.10.005>
- Stanney KM, Mourant RR, Kennedy RS (1998a) Human factors issues in virtual environments: a review of the literature. *Presence Teleoperat Virtual Environ* 7(4):327–351. <https://doi.org/10.1162/105474698565767>
- Stanney K, Salvendy G, Deisinger J, DiZio P, Ellis S, Ellision J, Fogleman G, Gallimore J, Hettinger L, Kennedy R, Lackner J, Lawson B, Maida J, Mead A, Mon-Williams M, Newman D, Piantanida T, Reeves L, Riedel O, Singer M, Stoffregen T, Wann J, Welch R, Wilson J, Witmer R (1998b) Aftereffects and sense of presence in virtual environments: formulation of a research and development agenda. Report sponsored by the Life Sciences Division at NASA headquarters. *Int J of Hum Comput Interact* 10:135–187
- Stanney KM, Kennedy RS, Hale KS (2014) Virtual Environment Usage Protocols. In: Hale KS, Stanney KM (eds) *Handbook of virtual environments: design, implementation and applications*. CRC Press, Boca Raton, pp 532–587
- Stanney K, Fidopiastis C, Foster L (2020) Virtual reality is sexist: but it does not have to be. *Front Robot AI* 7:4. <https://doi.org/10.3389/frbot.2020.00004>
- Stauffert JP, Niebling F, Latoschik ME (2018) Effects of latency jitter on simulator sickness in a search task. In: *IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE, Reutlingen, pp 121–127. <https://doi.org/10.1109/VR.2018.8446195>
- Stoffregen TA, Riccio GE (1991) An ecological critique of the sensory conflict theory of motion sickness. *Ecol Psychol* 3(3):159–194. https://doi.org/10.1207/s15326969eco0303_1
- Stoffregen TA, Chang CH, Chen FC, Zeng WJ (2017) Effects of decades of physical driving on body movement and motion sickness during virtual driving. *PLoS ONE* 12(11):e0187120. <https://doi.org/10.1371/journal.pone.0187120>
- Teixeira J, Palmisano S (2021) The effect of dynamic field-of-view restriction on experiences of cybersickness and presence in HMD-based virtual reality. *Virtual Real* 25:433–445. <https://doi.org/10.1007/s10055-020-00466-2>
- Teixeira J, Miellel S, Palmisano S (2022) Unexpected vection exacerbates cybersickness during HMD-based virtual reality. *Front Virtual Real* 3:860919. <https://doi.org/10.3389/frvir.2022.860919>
- Van Waveren JMP (2016) The asynchronous time warp for virtual reality on consumer hardware. In: *Proceedings of the 22nd ACM conference on virtual reality software and technology*. ACM, Munich, pp 37–46. <https://doi.org/10.1145/2993369.2993375>
- Wu W, Dong Y, Hoover A (2013) Measuring digital system latency from sensing to actuation at continuous 1-ms resolution. *Presence Teleoperat Vir Environ* 22(1):20–35. https://doi.org/10.1162/PRES_a_00131
- Yildirim C (2019) Cybersickness during VR gaming undermines game enjoyment: a mediation model. *Displays* 59:35–43. <https://doi.org/10.1016/j.displa.2019.07.002>
- Yildirim C (2020) Don't make me sick: investigating the incidence of cybersickness in commercial virtual reality headsets. *Virtual Real* 24:231–239. <https://doi.org/10.1007/s10055-019-00401-0>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.