The simulator sickness questionnaire, and the erroneous zero baseline assumption

Phillip Brown*, Pieter Spronck and Wendy Powell

Department of Cognitive Science and Artificial Intelligence, Tilburg University, Tilburg, Netherlands

Cybersickness assessment is predominantly conducted *via* the Simulator Sickness Questionnaire (SSQ). Literature has highlighted that assumptions which are made concerning baseline assessment may be incorrect, especially the assumption that healthy participants enter with no or minimal associated symptoms. An online survey study was conducted to explore further this assumption amongst a general population sample (N = 93). Results for this study suggest that the current baseline assumption may be inherently incorrect.

Introduction

Virtual reality (VR) use is more prevalent than ever, with many disciplines exploring its capabilities, and the potential benefits as a health technology (Mahrer & Gold, 2009; Malloy & Milling, 2010; Powers, 2008; Riva et al., 2019), as a traditional entertainment medium (Tan et al., 2015), and increasingly as a practical application for training purposes (Gavish et al., 2013; Sarig Bahat et al., 2015).

Whilst VR has become broader in its applications, a persistent concern is cybersickness (McCauley & Sharkey, 1992), a common, negative side effect of VR use. Cybersickness is described as being similar to motion sickness (Rebenitsch & Owen, 2016), with symptoms including eye strain, headache, disorientation, and nausea (LaViola, 2000). However, unlike motion sickness, cybersickness can occur in the absence of physical motion, and is thus considered distinct (Stanney et al., 1997). Prevalence of cybersickness during or post VR immersion has previously been reported to be as high as 80% (Cobb et al., 1999; Kim et al., 2005), with recent literature indicating that this is still a widespread issue (Yildirim, 2020). Whilst discomfort to the user is undesirable, greater implications exists when considering, for example, the reported effect of increased cybersickness symptoms on reducing the efficacy of treatment in medical and rehabilitation applications (Hoffman et al., 2004; Wiederhold et al., 2014; Weech et al., 2019).

Cybersickness is described in the literature by a variety of terms, including simulator sickness (Kennedy et al., 1993), visually induced motion sickness (Hettinger, 1992), and virtual reality induced symptoms (Cobb et al., 1999). For the purposes of this work, we will be using the term cybersickness. While the cause of cybersickness is still debated, a

number of prominent theories have been proposed, with possibly the most prominent being the sensory mismatch theory (Stanney et al., 2020; Stoffregen & Smart, 1998). This theory postulates that visual-vestibular conflicts occur when visual stimuli prompt bodily movement, although often no movement is occurring during VR immersion. Other theories including the evolutionary hypothesis, which postulates that the brain recognises certain sensorimotor conflicts as toxins which it must rid (Stanney et al., 2020), although this has been challenged in recent literature (Lawson, 2014). In lieu of a definitive conclusion, or applicable preventative measure, it is important that we are able to measure cybersickness adequately.

The severity of cybersickness symptoms can be determined by several factors, such as those of the intervention itself, including the content and hardware being used, but also subjective traits. Specific to the hardware, factors can include refresh rate, input lag and visual quality (LaViola, 2000). Additionally, individual user factors such as age and ethnicity (Edwards & Fillingim, 2001; Knight & Arns, 2006; Davis et al., 2014) have been suggested to be linked to higher susceptibility and stronger overall effects of cybersickness. Gender has also been postulated as an influencing factor, although results have been inconsistent (Davis et al., 2014; Saredakis et al., 2020).

Traditionally, when testing VR applications for their suitability regarding the potential of induced cybersickness, self-assessment questionnaires, such as the standardised Simulator Sickness Questionnaire (SSQ), are typically administered (Kennedy et al., 1993). Administration of the SSQ commonly takes place post-intervention, as the authors initially proposed (Kennedy et al., 1993). Pre-intervention administration has been previously argued against, with key justifications being over concerns with regards to participant priming (Young et al., 2006), and that in general, difference scoring between pre and post is unreliable (Cronbach & Furby, 1970). However, recent literature suggests that this may not be the case (Bimberg et al., 2020), highlighting especially that fear of participant priming should not be a barrier to pre-exposure assessment. This is particularly relevant to within-subjects designs, which already require repeated administration.

In lieu of SSQ administration pre-exposure, susceptibility questionnaires have been and are still used (Golding, 1998; Kim et al., 2005; Recenti et al., 2021), although their success at predicting cybersickness-like symptoms vary (Golding et al., 2021). However, it may be argued that susceptibility questionnaires are more a measure of general tendency, not of current symptoms, and whilst it may be a predictor of cybersickness, it tells us nothing about the participants' current state. Aside from the SSQ, other questionnaires exist for assessing the impact of cybersickness symptoms, including the Cybersickness Questionnaire (CSQ) (Stone et al., 2017) and Virtual reality sickness questionnaire (VRSQ) (Kim et al., 2018), although these are far less frequently used. Validity for the CSQ and VRSQ has been reported to be greater than that of the SSQ (Sevinc, 2020), although their lack of implementation, and therefore the

potential inability to compare between studies rendered them undesirable for this work. However, both questionnaires share similar traits as they are derived from the SSQ, so it may be assumed that a similar baseline assumption would exist between all the variants, and that any inferences from the results of the SSQ may be applicable to the CSQ, and VRSQ, respectively. Ifi sub heading.

Although the SSQ was initially intended for simulator sickness assessment, it has been used in recent times more frequently for cybersickness assessment (Bouchard et al., 2021; Brown et al., 2021; Munafo et al., 2017), and has been administered in conflict of the initial interpretation by its authors, with evidence existing that it is also beneficial under certain circumstances for pre-exposure administration (Bliss et al., 1997; Zanbaka et al., 2004; Bouchard et al., 2007, 2021; Munafo et al., 2017). While the benefits of pre-exposure assessment have been previously explored, the justification and application of its use often contradict the benefits that pre-exposure assessment may provide, and all too often, dissemination of pre-exposure SSQ results is lacking (Kim et al., 2005; Min et al., 2004; Munafo et al., 2017; Zanbaka et al., 2004). Also, the majority of pre-exposure assessment is conducted amongst medical populations, as it is often assumed that participants defined as "healthy" enter with a zero-baseline score (Kennedy et al., 1993).

Furthermore, the initial validation of the SSQ was conducted with persons defined as healthy, and who were likely more adaptive to VR than the general population (Kennedy et al., 1993). Literature has previously highlighted why this sample specifically is not representative of the general population (Bouchard et al., 2007). This highlights a key notion that the current knowledge base is not informative as to how responses to the SSQ may vary between populations, such as people with medical conditions, as has been reported in the literature previously (Brown & Powell, 2021).

Of the work documenting non-zero SSQ baseline scores, the majority is concerned with medical populations, for example, persons with acute or chronic pain (Brown & Powell, 2021), or persons who suffer with anxiety disorders (Powers, 2008; Bouchard et al., 2009). Although symptoms of medical conditions may explain non-zero SSQ baseline scores amongst specific populations, what is not so determinable is what participants of a more general, non-medical population may report as an SSQ baseline, and how this may be comparable to medical populations.

In addition, the majority of research is not concerned with measuring cybersickness at baseline, as the zero-baseline assumption is retained in the absence of pre-intervention susceptibility questioning, which would often exclude participants if they answer anything other than feeling healthy (Garrett et al., 2017; Tashjian et al., 2017; Spiegel et al., 2019). However, the literature does suggest that assuming a zero baseline may be incorrect (Bolte et al., 2011; Freitag et al., 2018; Kim et al., 2021; Mostajeran et al., 2021; Young et al., 2006), although limited

literature exists on healthy participant baseline (pre-intervention) SSQ administration. Of the work which does report healthy participant baseline SSQ scores, the literature does suggest that non-zero scores of some concern have also been reported (Freeman et al., 2008; Langbehn et al., 2018; Oberdörfer and Games, 2019; Oberdörfer et al., 2021; Park et al., 2008).

However, absence of baseline SSQ scores could result in inaccurate interpretations. In studies which are not using the SSQ at baseline, it may be said that they are assuming it is the VR intervention which is influencing their post-intervention SSQ scores. It has also been postulated that assumptions of unknown and potentially erroneous baseline SSQ scores may be a reason for high scores being observed (Bimberg et al., 2020), or more practically, it may be that the attribution of a high post-exposure SSQ score may be misleading if the participant started the intervention with a non-zero score. Furthermore, it has been highlighted that identifying baseline symptoms would be useful to "further understand the theoretical underpinning of simulated motion sickness" (Bruck & Watters, 2009).

Populations of concern are those who, in regards to their baseline cybersickness scores, are either being screened out by virtue of exhibiting symptoms of an illness, which may correlate to symptoms of cybersickness, or *via* pre-exposure screening questionnaires. Both of which are uninformative to the intended population, as excluding participants as such is not informing about how applicable or beneficial a proposed intervention for said population may be. Therefore, rather than assuming the zero baseline when conducting the SSQ, understanding whether the zero baseline assumption is correct better informs post-exposure SSQ recordings.

Study objectives

In this work, we examined the assumption of a zero baseline, investigating non-exposure scores in a general population, and compared the results of defined healthy and medical sub-populations. Firstly, we suspected from the literature that the zero-baseline assumption is indeed incorrect (Bolte et al., 2011; Freitag et al., 2018; Kim et al., 2021; Mostajeran et al., 2021; Park et al., 2008; Young et al., 2006). Secondly, we aimed to obtain a better understanding on what a normal baseline range for the general populations looks like, as well as, understand how baselines of medical and healthy sub-populations differ.

Therefore, the primary objective of this work was to investigate the SSQ zero baseline assumption amongst a general population. As such, we first hypothesised that:

H1: The study sample SSQ scores will be significantly greater than the zero baseline assumption.

Secondary, this work investigated how cybersickness may differ between healthy and medical populations at baseline. We anticipated that both medical and healthy samples would score greater than the assumed zero-baseline, although our medical population sample was likely to score significantly higher than our healthy population sample, as we propose that symptoms of cybersickness and medical conditions may be confounded. Thus, we hypothesised (H2) that:

H2.1: Our healthy sub-population sample SSQ score will be significantly greater than the zero baseline assumption.
H2.2: Our medical sub-population sample will score significantly greater than the zero baseline assumption.
H2.3: Our medical sub-population sample will score significantly greater than that of our healthy sub-population on total score.

Additionally, how participants of the respective sub-populations may score on SSQ sub-categories would be informative, as although we hypothesised that the medical group would score greater than the healthy population on total score, the distribution of symptom scores may be disproportionate between the nausea and oculomotor sub scales. Therefore, we compared the sub-scale scores of our sub-populations, hypothesising that:

H3.1: Our medical sub-population sample will score significantly greater than that of our healthy sub-population on nausea sub-scale.
H3.2: Our medical sub-population sample will score significantly greater than that of our healthy sub-population on oculomotor sub-scale.

Post-hoc testing was conducted to determine how the respective populations scored on individual items of the SSQ. We present these results, and discuss, in the context of any potential differences between medical and healthy participants, whether frequent medication usage could influence SSQ scores, as proposed previously in the literature (Kruk, 1992; McCauley & Sharkey, 1992). Results of the SSQ will be presented using Bouchard et al. (2007, 2021) unweighted scoring approach. We discuss the benefits to an unweighted scoring approach, versus the traditional weighted scoring approach (Kennedy et al., 1993), including how the application of either could influence respective outputs and conclusions.

Materials and methods

Participants

Ninety three participants aged 18–80 (M = 29 ± 15) (Table 1) were recruited from the University of Tilburg participant pool, as well as *via* social media. Snowball sampling was also used,

TABLE 1 Participant demographics.

		n (%)
Gender		
	Male	39 (42%)
	Female	53 (57%)
	Prefer not to say	1 (1%)
Group		
	Medical	38 (41%)
	Healthy	55 (59%)
Nationality		
	Dutch	49 (53%)
	British	17 (18%)
	Other	27 (29%)

whereby any potential participants could contact the lead researcher for further information. Participation inclusion criteria was being aged 18 years or above, and being competent in the English language. Participants recruited from the participant pool who were also students were compensated for their time with course credit.

Participants were assigned to either the healthy or medical sub-population for analytical purposes. Persons in the healthy sub-population were included because they answered "no" to having a diagnosed medical condition, and "no" to taking regular or prescribed medication. Persons in the medical sub-population were assigned because they answered "yes" to either of the questions. All participants were assigned to the normal population group, as this was determined to represent the general population. For a complete dissemination of participant demographics, see Supplementary Appendix A.

Design

A cross-sectional, observational study design was used.

Data has primarily been presented using Bouchard's proposed revised factor structure (Bouchard et al., 2007), as opposed to Kennedy et al. (1993 initial factor structure. Bouchard's structure was used because an unweighted approach is a more suitable method of determining differences between groups (Bouchard et al., 2007), and better reflects users of the general population (Bouchard et al., 2021), in contrast to Kennedy's which was conducted amongst specified personnel (Kennedy et al., 1993). Results using Kennedy's weighted scoring approach have also been included in Supplementary Appendices B,C for transparency and replicability.

A preliminary power analysis indicated that 71 participants would be required. Further powering took place after data collection commenced, and collection would continue until adequately powered.

Procedure

Participants were asked to complete an online survey, consisting of a demographic questionnaire and the SSQ. Data collection took place during the months of restricted movement during COVID-19. In order to roughly mimic conditions which are commonly used prior to participating in VR studies, participants were asked to complete the online survey at a time of day roughly an hour after eating a meal and having drunk water. Participants were also asked to complete the study in a home based environment, as aside from COVID-19, it is common for VR use to take place within a home based environment, which would be similar to where daily VR usage is likely in area such as rehabilitation (Miller et al., 2014; Garcia et al., 2021).

On the demographic questionnaire, information on gender, age, and nationality, as well as medical information post questionnaire (any diagnosed condition or disorder, and any medication taken regularly or prescribed).

The maximum time allotted for completion was 15 min, although the average time taken to complete was 4 m 01 s ± 2 m 53 s.

The survey was administered via the secure Qualtrics platform (Qualtrics, 2005).

Data analysis

We examined the distribution of demographic characteristics, particularly between what is considered the general population (no exclusions), medical (participants who reported having a medical condition or prescribed medication), and healthy (participants who reported having no medical condition or prescribed medication).

A series of tests were performed between the healthy and medical sub-populations, and against the general populations zero baseline assumption, to determine if differences existed between total scoring, as well as the oculomotor and nausea sub-scales of the SSQ.

The following is concerned with unweighted scorings, as it will more clearly demonstrate any potential differences between groups without weightings manipulating the interpretation of scores.

Results

Statistical

Kolmogorov-Smirnov tests of normality were conducted for all sub-sets of data being compared, with all showing a significant departure from normality, which is common for the SSQ (Helland et al., 2016; Lucas et al., 2020). Therefore, non-parametric tests were conducted for each hypothesis, presented below.

TABLE 2 Mann–Whitney test—individual SSQ question comparison between medial and healthy sub-population.

	SSQ1	SSQ2	SSQ3	SSQ4	SSQ5	SSQ6	SSQ7	SSQ8	SSQ9	SSQ10	SSQ11	SSQ12	SSQ13	SSQ14	SSQ15	SSQ16
U	864	870	878	988	901	1,032	845	966	1,020	979	1,002	1,000	960	1,034	1,022	1,015
Z	−1.52	−1.45	−1.53	−0.47	−1.17	−0.13	−2.01	−0.94	−0.20	−0.54	−0.42	−0.49	−0.93	−0.18	−0.22	−0.35
Sig	0.13	0.15	0.13	0.64	0.24	0.90	**0.04***	0.35	0.84	0.59	0.68	0.63	0.35	0.86	0.83	0.73

The * and bold response is to indicate a significant result.

H.1: Baseline SSQ scores in our general population displayed a median SSQ score of 8 (IQR = 4–14.5). A one-sample Wilcoxon signed rank test was run to determine whether baseline SSQ scores in our general population were different to the assumed baseline, defined as a baseline score of 0. Post-test ranks, median = 8, were statistically significantly higher than the assumed baseline score of 0 ($T = 4,095$, $z = 8.24$, $p < .001$).

H.2.1: Baseline SSQ scores in the recruited healthy sub-population displayed a median SSQ score of 9 (IQR = 4–14). A one-sample Wilcoxon signed rank test was run to determine whether baseline SSQ scores in the recruited healthy sub-population were different to the assumed baseline, defined as a baseline score of 0. Post-test ranks, median = 9, were statistically significantly higher than the assumed baseline score of 0 ($T = 1,485$, $z = 6.40$, $p < .001$).

H.2.2: Baseline SSQ scores in the recruited medical sub-population displayed a median SSQ score of 8 (IQR = 4–15). A one-sample Wilcoxon signed rank test was run to determine whether baseline SSQ scores in the recruited medical sub-population were different to the assumed baseline, defined as a baseline score of 0. Post-test ranks, median = 8, were statistically significantly higher than the assumed baseline score of 0 ($T = 666$, $z = 5.24$, $p < .001$).

H.2.3: A Mann-Whitney test was run to determine whether total SSQ scores between the medical and healthy sub-populations were different. It was indicated that a non-significant difference exists between the groups ($U = 1,024$, $p = 0.869$).
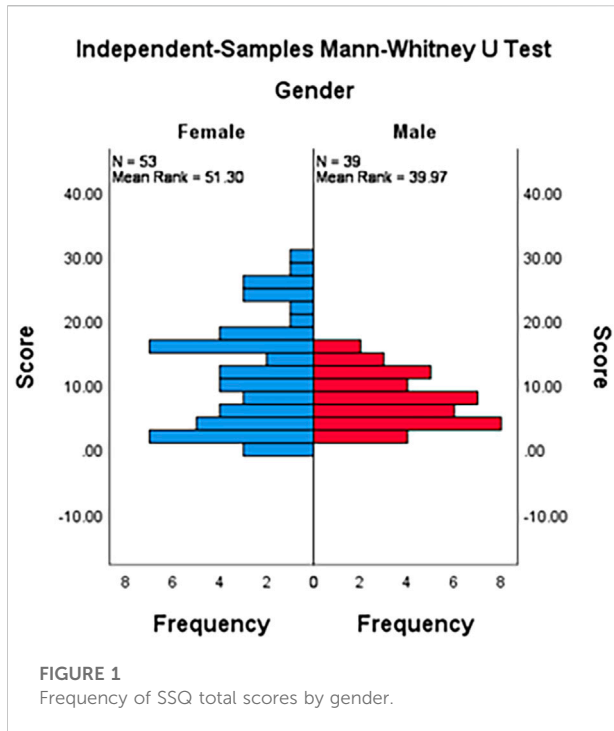
H.3.1: A Mann-Whitney test was run to determine whether nausea SSQ scores between the medical and healthy sub-populations were different. It was indicated that a non-significant difference exists between the groups ($U = 1,004.5$, $p = 0.748$).

H.3.2: A Mann-Whitney test was run to determine whether oculomotor SSQ scores between the medical and healthy sub-populations were different. It was indicated that a non-significant difference exists between the groups ($U = 1,021.5$, $p = 0.854$).

Post-hoc analysis of individual simulator sickness questionnaire items

Following the non-significant findings between our medical and healthy sub-populations (see result of H.2.3), we were interested to discover how individual questions on the SSQ were answered, and whether any significant differences between the sub-populations existed.

To determine this, a Mann-Whitney test was performed, comparing the 16 SSQ questions between the medical and healthy sub-populations.

FIGURE 1
Frequency of SSQ total scores by gender.

TABLE 3 Participant responses and categorisation—SSQ total score.

SSQ Score	Categorisation	Number of responses (%)
0	No symptoms	3 (3.2%)
<5	Negligible symptoms	4 (4.3%)
5–10	Minimal symptoms	2 (2.2%)
10–15	Significant symptoms	11 (11.8%)
15–20	Symptoms are a concern	5 (5.4%)
>20	A bad intervention	68 (73.1%)

line with our initial hypothesis. More specifically, the study sample mean was much larger than the zero-baseline assumption, with an average score of 9.66 ± 7.11, as it was observed that although some participants did enter with minimal symptoms, the majority exhibited symptoms which may be considered significant (see Table 3).

Basing categorisation from Stanney et al. (1997) paper, and using a weighted scoring approach for comparison, the authors looked at distinguishing between simulator sickness and cybersickness, the responses when no intervention has taken place clearly shows that if 73.1% of participants gave these responses post-immersion, it would have been determined that the intervention was "bad", and that 90.3% of participants would have experienced "significant symptoms" or greater (Table 3). Although we do not propose this to be a direct comparison to Stanney & Kennedy's work, as we are talking about two different phenomena, as their work outlined, the principles of the symptoms share the same motion sickness profile, and thus what can be derived from their categorisation is comparable.

Whilst the above is scored using the Kennedy's weighted approach (Kennedy et al., 1993), we do propose future work determine severity categorisation using an Bouchard's unweighted approach (Bouchard et al., 2021).

A recent meta-analysis has suggested that using scores above 20 to describe a "bad" intervention is outdated, as it was reported amongst work reviewed that the withdrawal rate of approximately one third of users was with a weighted SSQ score of 40 or higher (Caserman et al., 2021). It was therefore suggested that a score of 40 would be more indicative of a "bad" intervention. Using this categorisation, 58.1% of our participants still scored greater than 40, and if conducted as a post-immersion questionnaire, the intervention would have been considered unsuitable.

Overall, a non-zero baseline has been observed in this work in the majority of participants, with only 3 (3.2%) of the 93 participants reporting a zero baseline (no symptoms). Of those 3 participants, 2 of the participants were classified as part of the medical sub-population, with the other participant being categorised under the healthy sub-population.

Tested against a Bonferroni-adjusted alpha level of 003 (0.05/16), it was determined that no significant differences exist between the sub-populations answers for all questions (Table 2).

Effects of individual factors on simulator sickness questionnaire scores

As literature suggests that individual factors of participants may linked to changes in SSQ scores, tested to determine if differences between genders exist in a general collected dataset.

A Mann-Whitney test was run to determine if differences existed between male and female participants. The result for the test indicated that a significant difference exists between the groups (U = 1,288, p = 044). Additionally it was observed that female participant scores were more dispersed between low and high answers, whilst male scores were concentrated between low-medium results (Figure 1).

Discussion

H1—Investigating the simulator sickness questionnaire zero baseline assumption

The primary objective of this work was to investigate the validity of the SSQ zero baseline assumption. Our results indicated that our sample population scored significantly greater than the general populations assumed zero baseline, in

TABLE 4 SSQ total scores.

Total score (non-weighted)	Mean	SD	SE	n
Normal population	9.656	7.107	0.737	93
Healthy population	9.545	7.031	0.948	55
Medical population	9.816	7.307	1.185	38

TABLE 5 SSQ sub-scale scores.

Sub-scale scores (non-weighted)	Mean	SD	SE
Nausea			
Normal population	3.237	3.792	0.393
Healthy population	3.200	3.899	0.526
Medical population	3.289	3.683	0.597
Oculomotor			
Normal population	6.419	4.111	0.426
Healthy population	6.345	4.052	0.546
Medical population	6.526	4.247	0.689

FIGURE 2
Distribution of total SSQ scores by sub-population, displaying median and interquartile ranges. Individual datapoints are shown in blue.

This work clearly indicates that amongst our normal population sample, the zero baseline assumption is unfounded. This finding extends previous literature which has carried out similar pre-condition SSQ assessment (Bliss et al., 1997; Zanbaka et al., 2004; Munafo et al., 2017), supports previous literature which has reported similar non-zero baseline scores (Freeman et al., 2008; Oberdörfer and Games, 2019; Brown & Powell, 2021), and thus, it is reasonable to reject the zero baseline assumption.

H2 and H3—Do pre-immersion simulator sickness questionnaire scores differ between healthy and medical sub-populations

Our second hypothesis was concerned with how medical and healthy sub-populations may individually differ from the zero baseline assumption. Our results indicated that a Significant departure from the zero baseline assumption exists for both medical and healthy sin-populations. Therefore, the zero baseline assumption is not a realistic expectation for pre-VR baseline SSQ self-score.

Furthermore, we hypothesised that our medical sub-population would score significantly greater than the healthy sub-population. Our results indicated that the medical population total SSQ score was not significantly greater than that of the healthy population. As Table 4 shows, all groups scored similarly. Standard deviation from the sample means was large, which could indicate that the responses between the
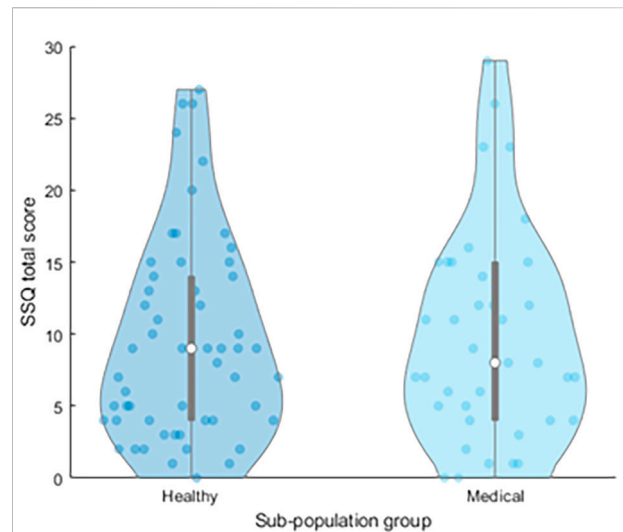
participants in their respective groups was highly variable. However, the standard error may be more indicative of how our samples means are dispersed relative to the population. With this reasoning, it would still be expected that the non-zero baseline would be sustained within the wider population E (within 95% CI) (Table 4).

Likewise, for the SSQ oculomotor and nausea sub-scale scores, our results indicated that no statistically significant difference existed between the groups on either sub-scale scoring. Similarly, to the total score interpretation, the variability in responses to questions pertaining to both the oculomotor and nausea sub categories is high, both for healthy and medical responses (Table 5). Likewise, how this may be expected to reflect the wider population would likely result in non-zero baselines for both sub-scales.

Literature suggests that persons with medical conditions or taking medication may score more highly on the SSQ (Kruk, 1992; McCauley & Sharkey, 1992), as symptoms of condition and the SSQ are similar. However, the results of this work indicate that no statistical significance exists between the healthy and medical populations, both of whom presented non-zero baseline results. Furthermore, healthy and medical distributions were similar, with a large variety of responses regardless of sub-population group (Figure 2).

Collectively, these results could indicate that healthy and medical participants alike may enter a VR intervention with pre-existing cybersickness symptoms. It may be determined from these findings that pre-immersion testing would be beneficial if using the SSQ as an indicator for cybersickness, regardless of the participants' medical background. Furthermore, it also could

TABLE 6 Individual question SSQ scores.

Normal	SSQ1	SSQ2	SSQ3	SSQ4	SSQ5	SSQ6	SSQ7	SSQ8	SSQ9	SSQ10	SSQ11	SSQ12	SSQ13	SSQ14	SSQ15	SSQ16
Mean	0.85	1.17	0.48	0.91	1.30	0.39	0.35	0.23	1.18	1.01	0.35	0.28	0.28	0.16	0.42	0.28
SE	0.09	0.09	0.07	0.09	0.10	0.08	0.07	0.06	0.10	0.10	0.06	0.06	0.06	0.06	0.07	0.07
Healthy																
Mean	0.727	1.055	0.418	0.891	1.400	0.364	0.455	0.291	1.200	1.055	0.327	0.255	0.236	0.164	0.400	0.309
SE	0.105	0.114	0.099	0.126	0.131	0.091	0.096	0.092	0.133	0.133	0.078	0.079	0.073	0.077	0.092	0.097
Medical																
Mean	1.026	1.342	0.579	0.947	1.158	0.421	0.211	0.132	1.158	0.947	0.395	0.316	0.342	0.158	0.447	0.237
SE	0.153	0.152	0.111	0.141	0.162	0.134	0.094	0.056	0.167	0.160	0.104	0.101	0.102	0.089	0.117	0.096

suggest that in many studies the post-exposure SSQ score may not in fact be indicating an adverse effect of the VR exposure, but may simply reflect an existing SSQ score.

Individual question, and sub-scale interpretation

When concerned with the individual items of the SSQ, the medical sub-population scored a greater mean score on general discomfort (SSQ1), fatigue (SSQ2), headache (SSQ3), eye strain (SSQ4), salivation increasing (SSQ6), blurred vision (SSQ11), dizziness with eyes open (SSQ12), dizziness with eyes closed (SSQ13), and stomach awareness (SSQ15), and interestingly, the healthy sub-population scored higher average scores on all of the other questions (Table 6; Figure 3).

The implications of this may be that sub-scales which feature these symptoms in their scoring would indicate a greater prevalence of that effect in certain populations.

Just taking the normal population means into consideration, it may be expected from our results that some questions would reliably score a non-zero baseline score in the wider population. For example, fatigue (SSQ2), and difficulty concentrating (SSQ9), it may be expected that these questions would score at least "slight" (1) on the SSQ. If these questions score for the majority of participants, then a new non-zero baseline, or expected baseline, is already established. The likelihood of an example like this occurring is not unreasonable, as we have found that 51 participants (55%) of our study sample answered at least "slight" (1) to these two questions. Of those 51 participants, 23 (45%) were classified as belonging to the medical population, meaning that there over half of the respondents (28 participants) who many report symptoms on the most common categories to score are in fact regarded as being healthy, with no medical condition or prescribed medication.

Concerning how scores are distributed (Figure 4), the majority of responses are skewed towards participants answering with "none" or "slight", with some exceptions prominently being in the two questions highlighted previously (questions 2, and 9), as well as general discomfort (question 1). These three questions (excluding difficulty focussing, see below) were the only ones which more responses were given for having minimum "slight" (1) as opposed to "none" (0).

Overall, questions 8, 12, 13, 14, 15, 16 were answered by the majority of participants with "none" (0). The Virtual Reality Sickness Questionnaire (VRSQ) (Kim et al., 2018) had attempted to address this issue by redesigning the SSQ, but without a lot of the questions belonging to the nausea subscale, of which the majority of the questions highlighted above belong to. It may be advised that a questionnaire such as this may be more suited for measuring cybersickness as a self-report method, although only preliminary validation has taken place (Sevinc, 2020), and widespread implementation is lacking.

FIGURE 3
Individual questions SSQ mean scores.
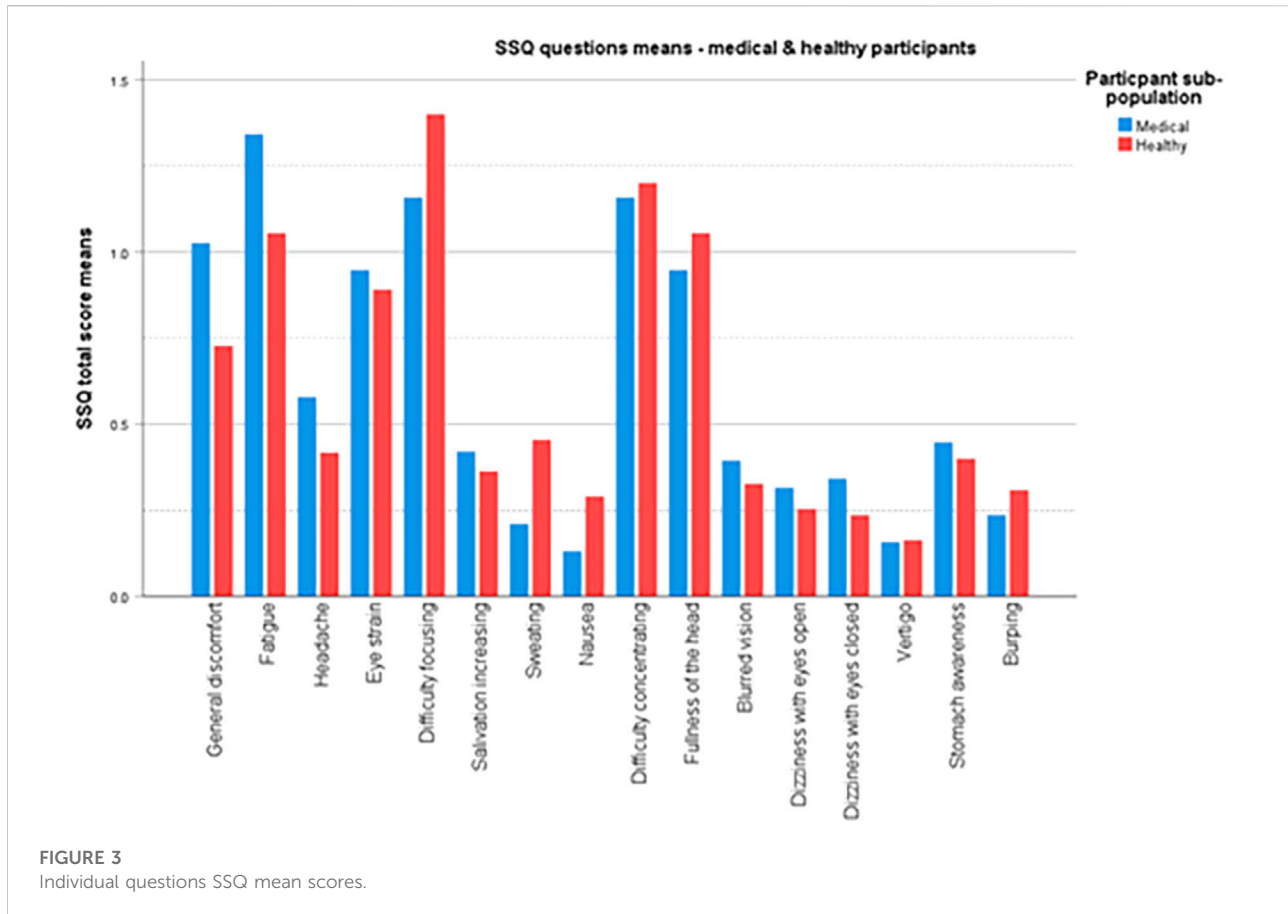
Question 5 (difficulty focusing) is intended to be a measure of visual accommodation, rather than a test of mental focus, and this was not explicitly explained to participants. Therefore, it could reasonable to assume that some participants scored it as such, which would explain why it scored amongst the highest in 'severe' across all questions. Difficulty focussing has been removed from the above dissemination for this reason. However, if we remove this question from our total score calculation, and base categorisation as similarly displayed in Table 3, the percentage of non-zero baseline scores remains the same (3.2%), and remaining category distributions similar (<5 = 6.5%; 5–10 = 4.3%; 10–15 = 4.3%; 10–15 = 10.7%; 15–20 = 6.5%; >20 = 68.8%).

Weighted vs. unweighted, and total score interpretation

For the purposes of comparing between participant sub-groups, we have done so based upon Bouchard et al. (2021) unweighted scoring approach. This has been done specifically for its ease of understanding, as well as its total score calculation being free of some items being counted twice,

as is the case with Kennedy et al. (1993) method. A distinction worth highlighting, is that the unweighted scoring approach implies that all symptoms equally contribute to the strength of felt-sickness.

Where this is particularly problematic is how total scores are interpreted. Five items of the SSQ are loaded on more than 1 factor, which give those items twice the weight of the other 11 items, and has previously been highlighted in the literature as being unintuitive (Bouchard et al., 2007, 2009). In Bouchard's unweighted approach, two participants could each score, for example, 6, where as if calculated using an weighted approach each score could be calculated to be anywhere between 22.44 and 44.88. With such varying scores for what is essentially a relatively comparable sickness, any statistical analysis made is not informative, as well as comparisons between interventions of sample groups tough to justify.

With that in mind, and as highlighted excellently by Bouchard et al. (2021), we propose that any future dissemination of SSQ results use an unweighted approach also.

For comparisons between weighted and unweighted total scores, see Supplementary Appendix B. For comparisons between weighted and unweighted sub scale scores, see Supplementary Appendix C. For a complete dissemination of
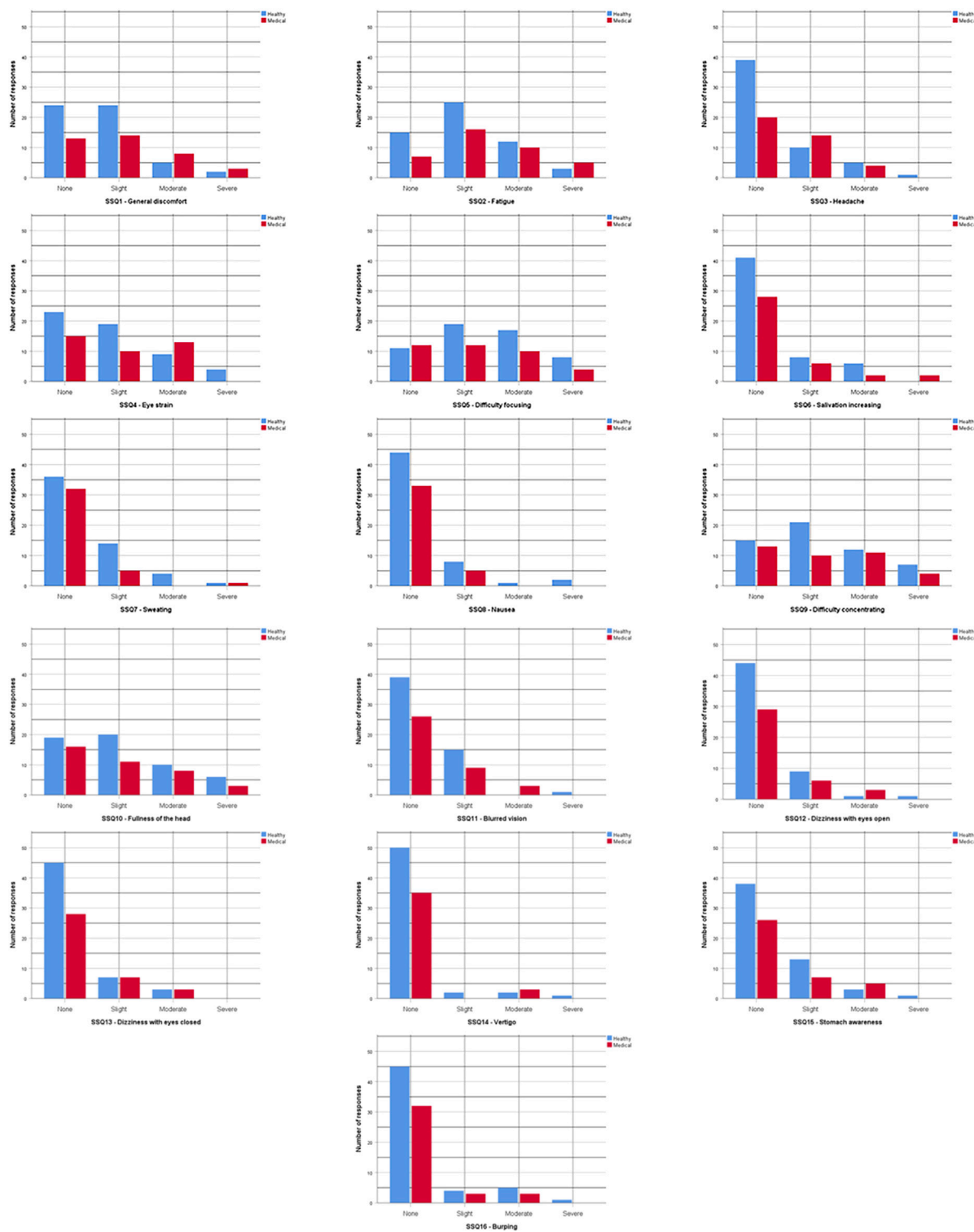
FIGURE 4
Left—right, top—bottom. SSQ questions 1–16 score distributions.

participant SSQ responses, both weighted and unweighted, see data availability statement.

This work also highlights that individual factors such as age and gender may be associated with how cybersickness associated symptoms are reported pre-exposure. For example, the literature postulates a number of factors as to why female participants may report heightened susceptibility to cybersickness (Davis et al., 2015; Saredakis et al., 2020), although in lieu of VR intervention, significantly different scores have been observed between males and females. Regardless of what the cause of this disparity is, acknowledging that disparity may exist, and that participants of different backgrounds exhibit variable, and potentially non-zero, baseline symptoms, should be considered as a cause to work with informed baselines and not assumptions.

As suggested by Stanney et al. (1997), the profiles of simulator and cybersickness are distinct, and it may be proposed therefore that the criteria for assessment should be also. In lieu of an established observational measurement for cybersickness, it is proposed that alternative measurement techniques should be considered. Such techniques could include physiological measuring, such as heart-rate variability (Garcia-Agundez et al., 2019), and electrodermal activity (Caserman et al., 2021), which has been demonstrated as a potential solution to undesirable self-report methods (Mavridou et al., 2018).

Conclusion and future work

This work has presented a case for rejecting a zero baseline assumption, and has demonstrated that it may be not just persons with medical conditions who would report a non-zero baseline, but also participants who would usually be defined as healthy. Future work should look to expand upon this result with a larger sample population, and could even explore participants who were expecting a VR intervention also, as opposed to just those knowingly not participating in a further or continued experiment.

The exploration of alternatives to the SSQ should also be considered, as literature highlights that self-report methods are undesirable for their potential lack of consistency, and inability to measure response in real time (Chang et al., 2020).

Behavioural measures such as physiological responses should be considered further, with evidence in the literature to suggest that measures such as Electrodermal Activity (Magaki & Vallance, 2019), Heart Rate (Cebeci et al., 2019), Heart Rate Variability (Magaki & Vallance, 2019), and electroencephalogram (H. Kim et al., 2021), may be appropriate measurement techniques, in lieu of current self-report approaches. Physiological assessment provides a non-invasive, non-subjective ground truth which would be invaluable in instances where self-report methods are either unfeasible, or undesirable.

Limitations

It is important to note that this survey was partly conducted amongst student participants who were re-imbursed *via* course credits for their participation, as well as non-student participants who were not reimbursed. Also, as this was conducted remotely, there is no way to eliminate guessing answers by the student population, as they had incentive to complete the questionnaire. We did however compare completion times between the groups, and found there no be significant differences between how long participants took in answering the questionnaire, nor were any answer times regarded as unreasonably short (indicating guessing for accreditation did not take place). However, some higher than expected SSQ scores were observed, so this cannot be ruled out.

Although precise splits between the medical and healthy sub-populations would have been ideal for drawing comparisons between them, this work was concerned with sampling the general population, so it could not be expected that precise splits would be achieved. For more precise comparisons between the groups, future work could aim to recruit more stringent to the applicable sub-populations. Furthermore, distinctions between, and severity of medical conditions were not included as part of this analysis, which could have implications for how a general medical population, or participants of specific medical sub-populations report baseline SSQ. Further work is warranted to explore how specific medical sub-populations answer baseline SSQ, how these participants compare to single medical condition groups, and comparatively to defined healthy participants.

Further individual factor tests concerning the demographics of participants could not be performed because of the difference in group sizes. A test on gender differences was warranted, as although the groups were not even, they did not have a large disparity. The same was not true for demographics such as age and ethnicity, as the majority of participants were younger and of white European backgrounds.

Also of note, is that this work has been conducted during the COVID-19 pandemic, and the impact as such could have an effect on how people feel day to day. Questions of the SSQ ask of fatigue, difficulty focusing, and difficulty concentrating, all of which could realistically be abnormal during abnormal times. A study conducted after the main effects of the COVID-19 pandemic have diminished could clarify these concerns.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Tilburg School of Humanities and Digital Sciences Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

Author contributions

The study was designed by PB and WP. The data acquisition was conducted by PB. Formal analysis was conducted by PB. Supervision for this study was completed by WP. The author who took lead on writing the manuscript was PB, and main feedback and edits were provided by PS and WP. All authors contributed feedback to the final version of the manuscript.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2022.945800/full#supplementary-material

References

Bimberg, P., Weissker, T., and Kulik, A. (2020). "On the Usage of the Simulator Sickness Questionnaire for Virtual Reality Research," in Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interface. doi:10.1109/VRW50115.2020.00098

Bliss, J. P., Tidwell, P. D., and Guest, M. A. (1997). The Effectiveness of Virtual Reality for Administering Spatial Navigation Training to Firefighters. *Presence. (Camb).* 6, 73–86. doi:10.1162/pres.1997.6.1.73

Bolte, B., Bruder, G., and Steinicke, F. (2011). "The Jumper Metaphor: An Effective Navigation Technique for Immersive Display Setups," in Proceedings of Virtual Reality International Conference (VRIC).

Bouchard, S., Berthiaume, M., Robillard, G., Forget, H., Daudelin-Peltier, C., Renaud, P., et al. (2021). Arguing in Favor of Revising the Simulator Sickness Questionnaire Factor Structure when Assessing Side Effects Induced by Immersions in Virtual Reality. *Front. Psychiatry* 12, 739742. doi:10.3389/FPSYT.2021.739742

Bouchard, S., Robillard, G., and Renaud, P. (2007). Revising the Factor Structure of the Simulator Sickness Questionnaire. *Annu. Rev. Cybetherapy Telemedicine* 5, 117–122.

Bouchard, S., St-Jacques, J., Renaud, P., and Wiederhold, B. K. (2009). Side Effects of Immersions in Virtual Reality for People Suffering from Anxiety Disorders. *J. Cyber Ther. Rehabilitation* 2 (2), 127–137.

Brown, P., Powell, W., Dansey, N., Al-Abbadey, M., Stevens, B., and Powell, V. (2021). Virtual Reality as a Pain Distraction Modality for Experimentally Induced Pain in a Chronic Pain Population. *Cyberpsychology, Behav. Soc. Netw* 25 (1), 66–71.

Brown, P., and Powell, W. (2021). Pre-Exposure Cybersickness Assessment within a Chronic Pain Population in Virtual Reality. *Front. Virtual Real.* 2. doi:10.3389/frvir.2021.672245

Bruck, S., and Watters, P. A. (2009). "Estimating Cybersickness of Simulated Motion Using the Simulator Sickness Questionnaire (SSQ): A Controlled Study," in Proceedings of the 2009 6th International Conference on Computer Graphics, Imaging and Visualization: New Advances and Trends, CGIV2009. doi:10.1109/CGIV.2009.83

Caserman, P., Garcia-Agundez, A., Gámez Zerban, A., and Göbel, S. (2021). Cybersickness in Current-Generation Virtual Reality Head-Mounted Displays: Systematic Review and Outlook. *Virtual Real.* 25 (4), 1153–1170. doi:10.1007/s10055-021-00513-6

Cebeci, B., Celikcan, U., and Capin, T. K. (2019). A Comprehensive Study of the Affective and Physiological Responses Induced by Dynamic Virtual Reality Environments. *Comput. Animat. Virtual Worlds* 30, 3–4. doi:10.1002/cav.1893

Chang, E., Kim, H. T., and Yoo, B. (2020). Virtual Reality Sickness: A Review of Causes and Measurements. *Int. J. Human–Computer. Interact.* 36 (17), 1658–1682. doi:10.1080/10447318.2020.1778351

Cobb, S. V. G., Nichols, S., Ramsey, A., and Wilson, J. R. (1999). Virtual Reality-Induced Symptoms and Effects (VRISE). *Presence. (Camb).* 8, 169–186. doi:10.1162/105474699566152

Cronbach, L. J., and Furby, L. (1970). How We Should Measure "Change": Or Should We? *Psychol. Bull.* 74, 68–80. doi:10.1037/h0029382

Davis, S., Nesbitt, K., and Nalivaiko, E. (2014). "A Systematic Review of Cybersickness," in Proceedings of the 2014 Conference on Interactive Entertainment, 1–9. doi:10.1145/2677758.2677780

Davis, S., Nesbitt, K., and Nalivaiko, E. (2015). "Comparing the Onset of Cybersickness Using the Oculus Rift and Two Virtual Roller Coasters," in 11th Australasian Conference on Interactive Entertainment (IE 2015).

Edwards, R. R., and Fillingim, R. B. (2001). Age-associated Differences in Responses to Noxious Stimuli. *Journals Gerontology Ser. A Biol. Sci. Med. Sci.* 56 (3), M180–M185. doi:10.1093/gerona/56.3.M180

Freeman, D., Pugh, K., Antley, A., Slater, M., Bebbington, P., Gittins, M., et al. (2008). Virtual Reality Study of Paranoid Thinking in the General Population. *Br. J. Psychiatry* 192 (4), 258–263. doi:10.1192/bjp.bp.107.044677

Freitag, S., Weyers, B., and Kuhlen, T. W. (2018). "Interactive Exploration Assistance for Immersive Virtual Environments Based on Object Visibility and Viewpoint Quality," in 25th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2018 - Proceedings. doi:10.1109/VR.2018.8447553

Garcia, L., Birckhead, B., and Krishnamurthy, P. (2021). An 8-Week Self-Administered at-Home Behavioral Skills-Based Virtual Reality Program for Chronic Low Back Pain: Double-Blind, Randomized, Placebo. *Jmir. Org.* 10 (1), 59–73. doi:10.2196/25291

Garcia-Agundez, A., Reuter, C., Caserman, P., Konrad, R., and Gobel, S. (2019). Identifying Cybersickness through Heart Rate Variability Alterations. *Int. J. Virtual Real.* 19 (1), 1–10. doi:10.20870/ijvr.2019.19.1.2907

Garrett, B., Taverner, T., and McDade, P. (2017). Virtual Reality as an Adjunct Home Therapy in Chronic Pain Management: An Exploratory Study. *JMIR Med. Inf.* 5, e11. doi:10.2196/medinform.7271

Gavish, N., Gutiérrez, T., Webel, S., Rodríguez, J., Peveri, M., Bockholt, U., et al. (2013). Evaluating Virtual Reality and Augmented Reality Training for Industrial Maintenance and Assembly Tasks. *Interact. Learn. Environ.* 23 (6), 778–798. doi:10.1080/10494820.2013.815221

Golding, J. F. (1998). Motion Sickness Susceptibility Questionnaire Revised and its Relationship to Other Forms of Sickness. *Brain Res. Bull.* 47, 507–516. doi:10.1016/S0361-9230(98)00091-4

Golding, J. F., Rafiq, A., and Keshavarz, B. (2021). Predicting Individual Susceptibility to Visually Induced Motion Sickness by Questionnaire. *Front. Virtual Real.* 0, 3. doi:10.3389/FRVIR.2021.576871

Helland, A., Lydersen, S., Lervåg, L. E., Jenssen, G. D., Mørland, J., and Slørdal, L. (2016). Driving Simulator Sickness: Impact on Driving Performance, Influence of Blood Alcohol Concentration, and Effect of Repeated Simulator Exposures. *Accid. Analysis Prev.* 94, 180–187. doi:10.1016/J.AAP.2016.05.008

Hettinger, L. (1992). Visually Induced Motion Sickness in Virtual Environments. *Presence. (Camb).* 1 (3), 306–310. doi:10.1162/pres.1992.1.3.306

Hoffman, H. G., Sharar, S. R., Coda, B., Everett, J. J., Ciol, M., Richards, T., et al. (2004). Manipulating Presence Influences the Magnitude of Virtual Reality Analgesia. *Pain.* doi:10.1016/j.pain.2004.06.013

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *Int. J. Aviat. Psychol.* 3, 203–220. doi:10.1207/s15327108ijap0303_3

Kim, H., Kim, D. J., Chung, W. H., Park, K.-A., Kim, J. D. K., Kim, D., et al. (2021). Clinical Predictors of Cybersickness in Virtual Reality (VR) Among Highly Stressed People. *Sci. Rep.* 11 (1), 12139. doi:10.1038/s41598-021-91573-w

Kim, H. K., Park, J., Choi, Y., and Choe, M. (2018). Virtual Reality Sickness Questionnaire (VRSQ): Motion Sickness Measurement Index in a Virtual Reality Environment. *Appl. Ergon.* 69, 66–73. doi:10.1016/j.apergo.2017.12.016

Kim, Y. Y., Kim, H. J., Kim, E. N., Ko, H. D., and Kim, H. T. (2005). Characteristic Changes in the Physiological Components of Cybersickness. *Psychophysiology* 4 (5), 616–625. doi:10.1111/j.1469-8986.2005.00349.x

Knight, M. M., and Arns, L. L. (2006). "The Relationship Among Age and Other Factors on Incidence of Cybersickness in Immersive Environment Users," in Proceedings - APGV 2006: Symposium on Applied Perception in Graphics and Visualization, doi:10.1145/1140491.1140539

Kruk, R. V. (1992). "Simulator Sickness Experience in Simulators Equipped with Fiber Optic Helmet Mounted Display Systems," in Flight Simulation Technologies Conference, doi:10.2514/6.1992-4135

Langbehn, E., Lubos, P., and Reality, F. (2018). Evaluation of Locomotion Techniques for Room-Scale Vr: Joystick, Teleportation, and Redirected Walking. *VRIC '18 Proc. Virtual Real. Int. Conf.* 18. doi:10.1145/3234253.3234291

LaViola, J. J. (2000). A Discussion of Cybersickness in Virtual Environments. *SIGCHI Bull.* 32, 47–56. doi:10.1145/333329.333344

Lawson, B. (2014). Motion Sickness Symptomatology and Origins. https://www.academia.edu/download/43998820/Chapter_23_Motion_Sickness_Symptomatology_and_Origins.pdf.

Lucas, G., Kemeny, A., and Paillot, D. (2020). A Simulation Sickness Study on a Driving Simulator Equipped with a Vibration Platform. *Elsevier* 68, 15–22.

Magaki, T., and Vallance, M. (2019). Developing an Accessible Evaluation Method of VR Cybersickness. 26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings, 1072–1073. doi:10.1109/VR.2019.8797748

Mahrer, N., and Gold, J. (2009). *The Use of Virtual Reality for Pain Control: A Review*. Springer, 100–109.

Malloy, K. M., and Milling, L. S. (2010). "The Effectiveness of Virtual Reality Distraction for Pain Reduction: A Systematic Review," in *Clinical Psychology Review.* doi:10.1016/j.cpr.2010.07.001

Mavridou, I., Seiss, E., Kostoulas, T., Nduka, C., and Balaguer-Ballester, E. (2018). "Towards an Effecve Arousal Detecon System for Virtual Reality," in Proceedings of the Human-Habitat for Health (H3): Human-Habitat Multimodal Interaction for Promoting Health and Well-Being in the Internet of Things Era - 20th ACM International Conference on Multimodal Interaction. doi:10.1145/3279963.3279969

McCauley, M. E., and Sharkey, T. J. (1992). Cybersickness: Perception of Self-Motion in Virtual Environments. *Presence. (Camb).* 1, 311–318. doi:10.1162/pres.1992.1.3.311

Miller, K., Adair, B., Pearce, A., Said, C. M., Ozanne, E., and Morris, M. M. (2014). Effectiveness and Feasibility of Virtual Reality and Gaming System Use at Home by Older Adults for Enabling Physical Activity to Improve Health-Related Domains: a Systematic Review. *Age Ageing* 43 (2), 188–195. doi:10.1093/ageing/aft194

Min, B. C., Chung, S. C., Min, Y. K., and Sakamoto, K. (2004). Psychophysiological Evaluation of Simulator Sickness Evoked by a Graphic Simulator. *Appl. Ergon.* 35 (6), 549–556. doi:10.1016/J.APERGO.2004.06.002

Mostajeran, F., Krzikawski, J., Steinicke, F., and Kühn, S. (2021). Effects of Exposure to Immersive Videos and Photo Slideshows of Forest and Urban Environments. *Sci. Rep.* 11 (1), 3994. doi:10.1038/s41598-021-83277-y

Munafo, J., Diedrick, M., and Stoffregen, T. A. (2017). The Virtual Reality Head-Mounted Display Oculus Rift Induces Motion Sickness and Is Sexist in its Effects. *Exp. Brain Res.* 235, 889–901. doi:10.1007/s00221-016-4846-7

Oberdörfer, S., and Games, M. (2019). Knowledge Encoding in Game Mechanics: Transfer-Oriented Knowledge Learning in Desktop-3d and Vr. *Int. J. Comput. Games* 2019.

Oberdörfer, S., Heidrich, D., Birnstiel, S., and Latoschik, M. E. (2021). Enchanted by Your Surrounding? Measuring the Effects of Immersion and Design of Virtual Environments on Decision-Making. *Front. Virtual Real.* 2. doi:10.3389/frvir.2021.679277

Park, J. R., Lim, D. W., Lee, S. Y., Lee, H. W., Choi, M. H., and Chung, S. C. (2008). Long-term Study of Simulator Sickness: Differences in EEG Response Due to Individual Sensitivity. *Int. J. Neurosci.* 118, 857–865. doi:10.1080/00207450701239459

Powers, M. (2008). Virtual Reality Exposure Therapy for Anxiety Disorders: A Meta-Analysis. *J. Anxiety* 22 (3), 561–569.

Qualtrics (2005). *Qualtrics*. Available at https://www.qualtrics.com.

Rebenitsch, L., and Owen, C. (2016). Review on Cybersickness in Applications and Visual Displays. *Virtual Real.* 20, 101–125. doi:10.1007/s10055-016-0285-9

Recenti, M., Ricciardi, C., Aubonnet, R., Picone, I., Jacob, D., Svansson, H. Á. R., et al. (2021). Toward Predicting Motion Sickness Using Virtual Reality and a Moving Platform Assessing Brain, Muscles, and Heart Signals. *Front. Bioeng. Biotechnol.* 9, 635661. doi:10.3389/fbioe.2021.635661

Riva, G., Wiederhold, B. K., and Mantovani, F. (2019). Neuroscience of Virtual Reality: From Virtual Exposure to Embodied Medicine. *Cyberpsychology, Behav. Soc. Netw.* 22 (1), 82–96. doi:10.1089/cyber.2017.29099.gri

Saredakis, D., Szpak, A., Birckhead, B., Keage, H. A. D., Rizzo, A., and Loetscher, T. (2020). Factors Associated with Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis. *Front. Hum. Neurosci.* 14, 96. doi:10.3389/fnhum.2020.00096

Sarig Bahat, H., Takasaki, H., Chen, X., Bet-Or, Y., and Treleaven, J. (2015). Cervical Kinematic Training with and without Interactive VR Training for Chronic Neck Pain - a Randomized Clinical Trial. *Man. Ther.* 20, 68–78. doi:10.1016/j.math.2014.06.008

Sevinc, V. (2020). Psychometric Evaluation of Simulator Sickness Questionnaire and its Variants as a Measure of Cybersickness in Consumer Virtual Environments. *Appl. Ergon.* 82, 102958. doi:10.1016/j.apergo.2019.102958

Spiegel, B., Fuller, G., Lopez, M., Dupuy, T., Noah, B., Howard, A., et al. (2019). Virtual Reality for Management of Pain in Hospitalized Patients: A Randomized Comparative Effectiveness Trial. *PLoS ONE* 14, e0219115. doi:10.1371/journal.pone.0219115

Stanney, K., Lawson, B. D., Rokers, B., Dennison, M., Fidopiastis, C., Stoffregen, T., et al. (2020). Identifying Causes of and Solutions for Cybersickness in Immersive Technology: Reformulation of a Research and Development Agenda. *Int. J. Human–Computer. Interact.* 36, 1783–1803. doi:10.1080/10447318.2020.1828535

Stanney, K. M., Kennedy, R. S., and Drexler, J. M. (1997). Cybersickness Is Not Simulator Sickness. *Proc. Hum. Factors Ergonomics Soc. Annu. Meet.* 41, 1138–1142. doi:10.1177/107118139704100292

Stoffregen, T. A., and Smart, L. J. (1998). Postural Instability Precedes Motion Sickness. *Brain Res. Bull.* 47, 437–448. doi:10.1016/S0361-9230(98)00102-6

Stone, W. B., Iii, Bruce, W., Froelich, A., Credé, M., Kelly, J., and Krizan, Z. (2017). Psychometric Evaluation of the Simulator Sickness Questionnaire as a Measure of

Cybersickness. Available at https://search.proquest.com/openview/c2d59f535cdea2e705131a3b052a1b05/1?pq-origsite=gscholar&cbl=18750.

Tan, C., Leong, T., and Shen, S. (2015). "Exploring Gameplay Experiences on the Oculus Rift," in Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play, 253–264. doi:10.1145/2793107.2793117

Tashjian, V., Mosadeghi, S., Howard, A. R., Lopez, M., Dupuy, T., Reid, M., et al. (2017). Virtual Reality for Management of Pain in Hospitalized Patients: Results of a Controlled Trial. *JMIR Ment. Health* 4 (1), e9. doi:10.2196/mental.7387

Weech, S., Kenny, S., and Barnett-Cowan, M. (2019). Presence and Cybersickness in Virtual Reality Are Negatively Related: A Review. *Front. Psychol.* 10, 158. doi:10.3389/fpsyg.2019.00158

Wiederhold, B. K., Gao, K., Sulea, C., and Wiederhold, M. D. (2014). Virtual Reality as a Distraction Technique in Chronic Pain Patients. *Cyberpsychology, Behav. Soc. Netw.* 17, 346–352. doi:10.1089/cyber.2014.0207

Yildirim, C. (2020). Don't Make Me Sick: Investigating the Incidence of Cybersickness in Commercial Virtual Reality Headsets. *Virtual Real.* 24, 231–239. doi:10.1007/s10055-019-00401-0

Young, S. D., Adelstein, B. D., and Ellis, S. R. (2006). "Demand Characteristics of a Questionnaire Used to Assess Motion Sickness in a Virtual Environment," in IEEE Virtual Reality Conference (VR 2006), 97–102. doi:10.1109/VR.2006.44

Zanbaka, C., Lok, B., Babu, S., Xiao, D., Ulinski, A., and Hodges, L. F. (2004). "Effects of Travel Technique on Cognition in Virtual Environments," in IEEE Virtual Reality 2004, 149–286. doi:10.1109/VR.2004.1310068