

The effect of a projected virtual reality training environment on vision symptoms in undergraduates

MERITXELL CRISTINO AMENÓS^{1,2} BSc (Hons) Orthop DOO(EC) DipTP(IP) AND PAUL C. KNOX PhD CBiol MIBiol¹

¹Eye and Vision Science, University of Liverpool, Liverpool

²Orthoptics and Optometry, Royal Liverpool University Hospital, Liverpool

Abstract

Aim: Virtual reality (VR) systems induce a range of unwelcome symptoms in a proportion of the population. A similar phenomenon has been reported with 3D presentation systems. Given the increasingly wide deployment of such systems, we investigated the effect of exposure to a projected VR training simulation on a group of undergraduates.

Methods: Two groups of students attended two teaching sessions using a 3D stereoscopic back-projector system with active stereo glasses. One group was given a full orthoptic and optometric assessment before they attended their first session.

Participants completed the Virtual Reality Symptom Questionnaire (VRSQ) before and after both sessions.

Results: While no participant reported any gross discomfort after either session, there was a statistically significant increase in VRSQ symptom scores from pre- to post-exposure in the first session that was not observed in the second session. Pre-exposure scores were statistically significantly different between sessions; analysis of the difference between individual pre- and post-exposure results from both sessions revealed no consistent effects. There was a statistically significant correlation between prism fusion amplitude and symptom scores.

Conclusions: We found no evidence of uncomfortable symptoms in a group of undergraduate students. Projected VR systems, in which participants are largely passive observers, are less likely to induce eye symptoms than head-mounted systems which make higher demands on the visual system. We also found that in a typical undergraduate class there were a number of students with no or low stereopsis who could derive no benefit from a VR system.

Key words: Stereopsis, Virtual Environment for Radiotherapy Training (VERT), Virtual reality, Virtual Reality Symptom Questionnaire (VRSQ)

Introduction

The 1990s saw the rapid deployment of virtual reality

(VR) systems in a wide range of contexts. As technology has developed, and costs have reduced, larger scale deployments in education and training have become possible. Allied to this has been the arrival of 3D cinema and TV, which seek to improve on the 2D visual experience, without the levels of interactivity that characterise more immersive VR systems. However, with the widespread use of these technologies a number of concerns have arisen. Following the release of Hollywood films such as 'Avatar' and 'Alice in Wonderland', there were numerous media reports attributing symptoms to viewing the films (see for example 'Do 3D films make you sick?'¹). Shibata *et al.*² quote the results of a survey of cinema-goers in Russia in which 30% of respondents reported 'eye tiredness' after watching a 3D film at the cinema. Many of these concerns have now attached themselves to 3D television, prompting one major manufacturer to issue a warning about certain groups who should not watch 3D TV. However, going to the cinema to watch a 3D film or buying a 3D TV are discretionary activities that are easily avoided if found to be unpleasant. Various VR and 3D systems are now being incorporated into educational and training programmes. Here, their use may become indispensable or even compulsory.

The Directorate of Medical Imaging and Radiotherapy at the University of Liverpool is one of several Higher Education providers in the UK which was funded by the UK Department of Health to introduce the use of a Virtual Environment for Radiotherapy Training (VERT). The VERT is used to simulate a clinical environment, for example simulating a clinical linear accelerator, allowing students to set treatment parameters, position a virtual patient for radiotherapy delivery and inspect the resulting distribution of radiation dose. All these technical skills can be developed without the risks inherent in using the real machine on a real patient. The VERT is a projected visual environment, similar in important respects to 3D cinema. It provides a large field of view, and shutter glasses are used to deliver appropriate frames to each eye to induce a disparity. Unlike 3D cinema (and 3D TV) it also allows a degree of interactivity on the part of the user. However, the sort of activities being undertaken, the visual stimuli involved, and the level of feedback and interactions, are some way short of what might be achieved in a more immersive VR system with a head-mounted display (HMD) and haptic or tactile feedback.

Correspondence and offprint requests to: Meritxell Cristino Amenós, 3 Stanley Terrace, Mossley Hill, Liverpool L18 5EE. e-mail: meri.cristino@gmail.com

	NONE	SLIGHT	MODERATE	SEVERE
General Symptoms				
General Discomfort	0	1 2	3 4	5 6
Fatigue	0	1 2	3 4	5 6
Boredom	0	1 2	3 4	5 6
Drowsiness	0	1 2	3 4	5 6
Headache	0	1 2	3 4	5 6
Dizziness	0	1 2	3 4	5 6
Difficulty Concentrating	0	1 2	3 4	5 6
Nausea	0	1 2	3 4	5 6
Eye Symptoms				
Tired Eyes	0	1 2	3 4	5 6
Sore/aching eyes	0	1 2	3 4	5 6
Eyestrain	0	1 2	3 4	5 6
Blurred vision	0	1 2	3 4	5 6
Difficulty focusing	0	1 2	3 4	5 6
Double vision	0	1 2	3 4	5 6

Fig. 1. Symptom questionnaire developed by Ames *et al.*⁹ and used in our study. Subjects completed one questionnaire before VERT exposure and one after, by circling a single digit against each item.

Concern arises because simulators and VR systems, particularly those using HMDs, have been shown to induce unwelcome symptoms in a proportion of the population.^{3–5} While less is known about projected systems,^{6,7} there are data that indicate their potential to induce symptoms.⁸ Therefore we investigated the visual symptoms induced by the Liverpool VERT in undergraduate radiotherapy students undergoing training.

Methods

Participants

All participants were first-year radiotherapy students, undertaking training in the BSc Radiotherapy undergraduate programme in the School of Health Sciences at the University of Liverpool. One whole first-year class, consisting of 33 students (mean age 22 years), had a full vision assessment before their first exposure to the VERT. Prior to a timetabled VERT session students completed a symptom questionnaire, and completed a second questionnaire immediately after the session. A second class of 30 students (mean age 23 years) was recruited and completed questionnaires only. This study was approved by the University of Liverpool Research Ethics Committee and followed the Tenets of the Declaration of Helsinki. All subjects provided written informed consent.

Vision assessment

The vision assessment consisted of measurement of the participant's visual acuity for near (at 35 cm) and distance (at 6 m) with their habitual correction (if any) using a logMAR near vision card and a Bailey-Lovie logMAR chart, respectively. We also measured best corrected visual acuity for near (at 35 cm) and distance (at 6 m, Bailey-Lovie chart), cover test and prism cover

test for near and distance, prism fusion amplitude for near and distance (using a prism bar), stereopsis for near (TNO), measurement of the AC/A ratio (heterophoria method) and near point of convergence (using an accommodative target).

The Virtual Reality Symptom Questionnaire (VRSQ)

To investigate the occurrence and severity of symptoms induced by VERT exposure we used the Virtual Reality Symptom Questionnaire (VRSQ⁹; see Fig. 1). This was originally designed to investigate the effects of head-mounted VR displays and comprises two lists of symptoms (General and Eye symptoms) relevant to viewing in VR environments. A scoring system from 0 (None) to 6 (Severe) is provided for each symptom. Each participant circles a single number for each symptom, and the score is calculated by summing the circled numbers. The Eye (score range of 0–36) and General (range 0–48) scores are the sum of the individual symptom scores under those categories. The Total score is the sum of Eye and General scores (range 0–84).

Of the group of 33 students who had their vision assessed, 32 completed the VRSQ before and after their first VERT session; 27 of these students also completed a second set of pre- and post-session questionnaires after a subsequent VERT session approximately 4 weeks after their first session. A second group of 30 first-year radiotherapy students also completed pre-VERT and post-VERT VRSQs for their first VERT session but did not have a vision assessment. Sixteen of this group also completed a second set of pre- and post-VERT questionnaires. Thus we had a total sample of 62 participants who completed questionnaires covering the first VERT session, 43 of whom also completed questionnaires for their second session.

Table 1. Summary of results from vision assessments of 33 students, comprising one full first-year class of radiotherapy students at the University of Liverpool undergoing training using VERT

Test	Range	Average	Variability
Habitual VA (logMAR)			
Near	-0.16 to 0.67	0.012	0.149
Distance	-0.12 to 0.89	0.051	0.185
BCVA (logMAR)			
Near	-0.16 to 0.5	-0.032	0.112
Distance	-0.12 to 0.59	-0.016	0.131
Stereo (seconds of arc)	15 to 480	45	30
AC/A ratio	0.3 to 10.4	5.893	1.698
Prism fusion amplitude (Δ)			
Near	22 to 70	55	8
Distance	14 to 45	30	11
Prism cover test (Δ)			
Near	30 BO to 12 BI	0	4
Distance	50 BO to 6 BI	0	4
Near point of convergence (cm)	6 to 18	7.27	2.18

Mean (SD) were used for continuous data, and median (IQR) for discrete data. Stereo vision was measured using the TNO test. Two of 33 participants did not exhibit any measurable stereopsis; the median shown here is for the 31 remaining participants. The AC/A ratio was measured using the heterophoria method. Prism fusion amplitude was calculated by combining the base in and base out prism fusion ranges.

VERT sessions

Students were exposed to the VERT in groups of 4 or 5, and teaching sessions lasted approximately 60 minutes. The VERT system comprises a 3D stereoscopic dual-projector back-projection system, which projects images on to a 2.4 m × 5.3 m screen. Active stereo goggles worn by users automatically shutter between stereo projector views to produce a 3D percept. The particular display used in the teaching sessions simulated a clinical linear accelerator, in which a virtual patient had to be positioned and treated. Students wore the shutter goggles throughout the session. VERT sessions consisted of an introduction to the simulated environment and instructions from a member of staff, then each student had the opportunity to actively operate the simulated clinical

machine while observed by the other students in their group. The operator stood approximately 1.5 m from the display screen, while observers sat approximately 2.5 m from it.

Data analysis

Data from the vision testing were summarised using parametric or non-parametric measures of central tendency and variability as appropriate. As the VRSQ returned categorical data, medians and interquartile ranges were used to summarise the results, and the Wilcoxon Signed Ranks test was used to compare pre-VERT and post-VERT symptom scores. For post-VERT questionnaires we also calculated the frequency of endorsement for each individual symptom. This refers to the proportion of subjects that gave a score greater than 0 for a given symptom. Data were collated using MS Excel and statistical analysis was conducted using SPSS.

Results

A summary of the results from the vision testing of the first class of 33 students is shown in Table 1. The students' mean near and distance acuities with their habitual correction were 0.012 ± 0.149 and 0.051 ± 0.185, respectively (logMAR, mean ± SD). Median stereopsis was 45" of arc; 29 of 33 subjects (88%) had 60" of arc or better. However, 4 of the 33 subjects (12%) had only gross stereopsis or suppression (2 had gross stereopsis due to anisometropic amblyopia and uncorrected refractive error, 2 had suppression due to childhood esotropia and surgical treatment of monocular congenital cataract).

We examined the pre- and post-VERT questionnaires for the first VERT session, combining data from the two groups of students, providing data from 62 participants (Fig. 2). For the Total symptom scores and both component scores (General and Eye) we observed increases in the post-VERT scores. Statistical analysis using a paired non-parametric test (Wilcoxon Signed

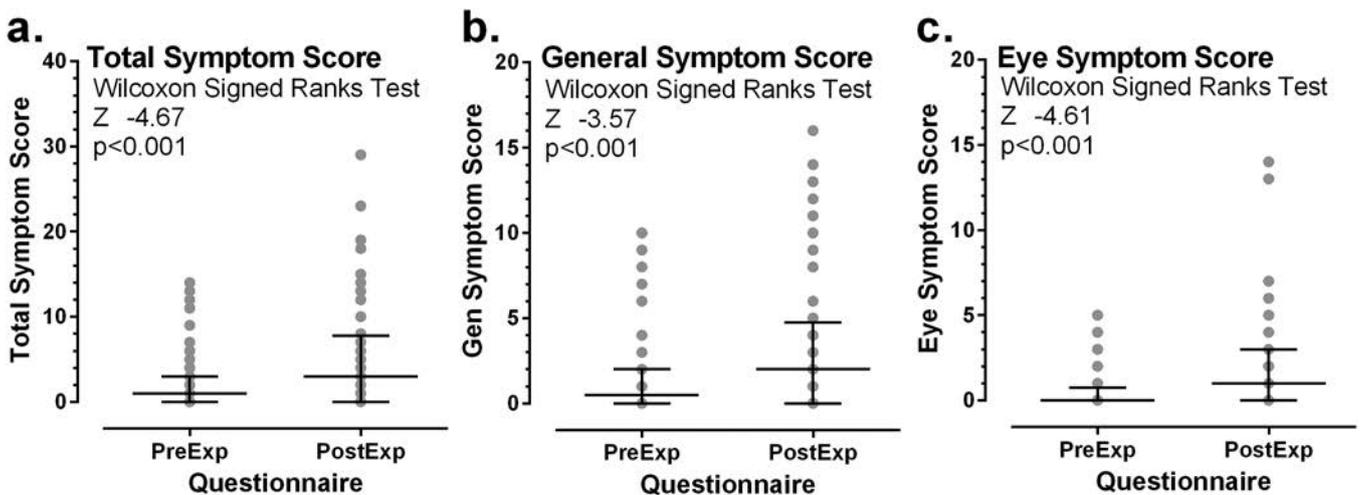


Fig. 2. Data from the first VERT session. Total (a), General (b) and Eye (c) symptom scores are shown. The points show data from each of 62 participants. Central (long) horizontal lines show median scores; the interquartile range is also shown. Note that the y-axis range in (a) is different from that in (b) and (c).

Ranks) demonstrated that the pre–post differences were statistically significant (Total, General and Eye all $p < 0.001$). When we examined the scores of individuals, the post-VERT Total scores were higher in 35 participants, unchanged in 23 and reduced in 4. Amongst those who reported higher post-VERT scores, the median increase in Total score was 3 points.

Forty-three participants completed pre- and post-VERT questionnaires after a second VERT session. This time we found that there was no statistically significant difference between pre- and post-VERT Total, Eye and General scores. Given this result we wished to compare the two sessions. When we compared the pre-VERT scores (i.e. the baseline scores) from the two sessions, we found that there was considerable variability and a statistically significant difference between them. Thus the median pre-VERT Total score was 1 (IQR 3) in session 1 and 2 (IQR 8) in session 2, a difference that was statistically significant (Wilcoxon Signed Ranks test, $p = 0.0016$). The differences between pre-VERT Eye and General sub-scores were also statistically significant ($p = 0.012$ and $p = 0.01$, respectively). Therefore in order to compare the two sessions, given this shifting baseline, we computed a difference score for each session separately by subtracting the pre-VERT score from the post-VERT score for each item. Thus no change in symptoms would be represented by 0, and any increase by a positive number.

The distributions of difference in scores for the two sessions are shown in Fig. 3. The proportion of participants reporting an increase in scores was consistent over the two sessions. For the Total scores, in both sessions 49% of the participants had positive difference scores, i.e. they reported higher post-VERT scores. For the Eye sub-score, 42% of participants had positive difference scores in session 1 and 40% in session 2. The equivalent figures for the General scores were 30% and 37%, respectively. The main difference between the sessions was that in session 2 the proportion of participants with negative difference scores was higher. Thus for the Total score, while only 7% of participants had negative differences in session 1, this rose to 26% in session 2. The equivalent figures for the Eye sub-score were 7% rising to 19%, and for the General sub-score, 19% rising to 30%. The overall effect of these patterns of response was a much more symmetrical distribution, reflecting the lack of a significant difference between median pre- and post-VERT scores for session 2 (Fig. 3).

We used the frequency of endorsement (FoE) to examine the pattern of response across the two VERT sessions (Fig. 4). In the General domain, the three symptoms with the highest FoE were the same in both sessions: general discomfort, fatigue and headache. In the Eye domain the three symptoms with the highest FoE were tired eyes, sore eyes and eye strain, with the same order in both sessions. There was a statistically significant correlation between the FoE in the two sessions (Spearman's $Rho = 0.92$; $p < 0.0001$). The four symptoms with the highest FoE averaged across the two sessions were general fatigue (0.45), tired eyes (0.4), general discomfort (0.36) and sore eyes (0.34).

Finally, for the 33 participants for whom we had vision and orthoptic data, we looked for correlations

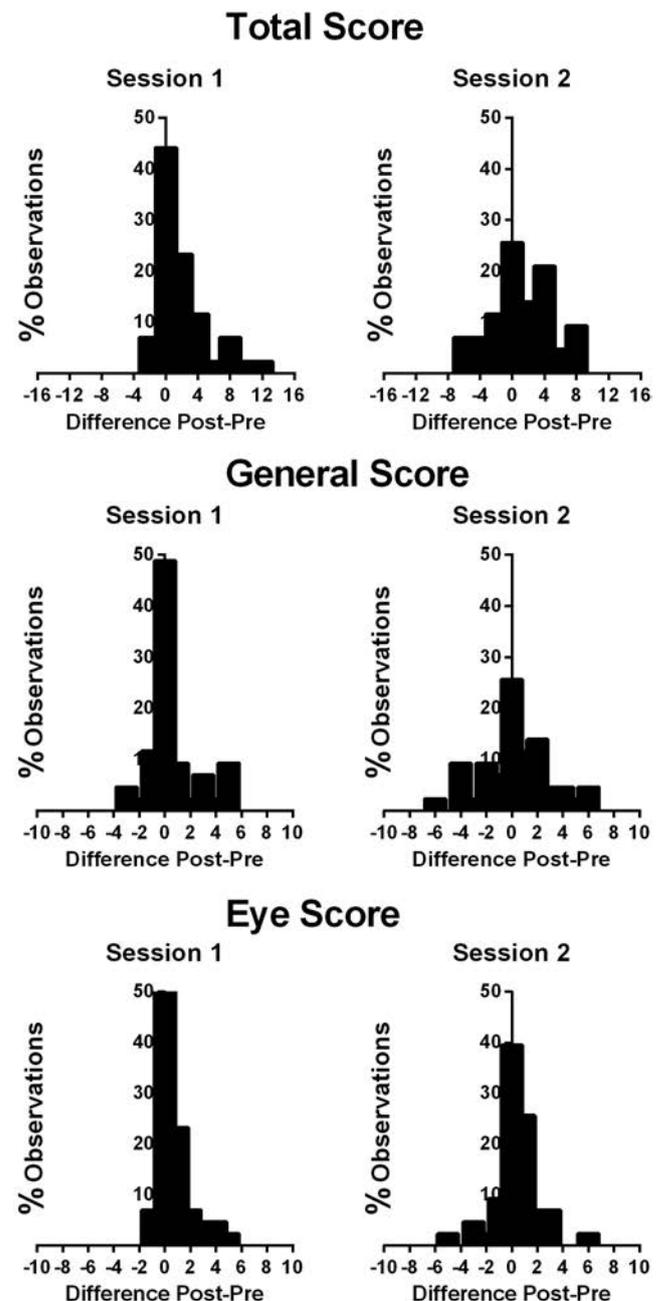


Fig. 3. Comparison of distributions of the difference in scores from two VERT sessions for 43 participants. The difference was calculated by subtracting the pre-VERT score from the post-VERT score for each subject. Total, General and Eye symptom scores are shown.

between these and their questionnaire scores. The only measurement for which a statistically significant correlation emerged was distance prism fusion amplitude (PFA; Fig. 5). A negative correlation (Spearman's $Rho = -0.59$, $p = 0.001$) was evident for the Total score; both the Eye and General sub-scores were statistically significantly correlated with PFA.

As mentioned previously, 2 of these subjects had suppression. However, their symptom scores from pre- to post-exposure showed no changes for both participants in the first session and in the second session. One subject reported a very slight increase in symptoms with the other showing a decrease.

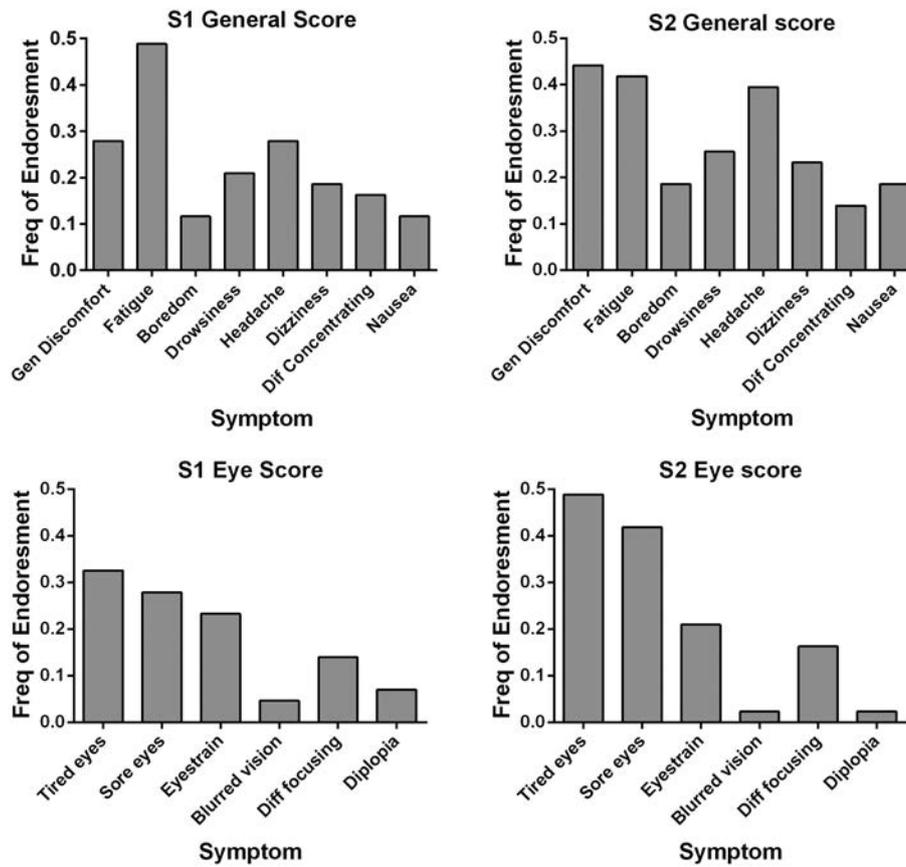


Fig. 4. Frequency of endorsement for each symptom in post-VERT questionnaires for 43 participants who completed two VERT sessions. General and Eye sub-group scores are shown separately. S1, first VERT session; S2, second VERT session.

Discussion

Given public concern about the generation of unwelcome symptoms when viewing 3D displays as reflected in the media, and reports in the literature, our purpose was to investigate symptoms in a group of undergraduates exposed to a VR radiotherapy training environment. The advantage of using such systems is that they provide a safe means of teaching basic skills, and potentially relieve hard-pressed clinical services of basic training duties. However, it is important to establish that these advantages are not outweighed by inducing an unacceptable range or intensity of unpleasant symptoms in users.

We found that in the first VERT session there was a statistically significant increase in symptom scores. However, none of our participants reported symptoms that required them to discontinue exposure. This result is similar to that reported by Flinton and White,¹⁰ who used the Simulator Sickness Questionnaire (SSQ)¹¹ to examine the effects of the same system (although running different simulations) on a group of staff and students. A majority of their participants (53 of 75) reported an increase in symptoms with exposure, with the four most frequently reported being ‘eyestrain’, ‘difficulty focusing’, ‘headache’ and ‘general discomfort’. These were generally similar to the four most

frequently endorsed symptoms in our sample. However, again in their sample, symptoms were not of a severity that required participants to discontinue exposure. Note that Flinton and White excluded from their experiment participants who when screened were ‘not in their usual state of fitness’. We made no attempt to use the baseline VRSQ data, or any other measurement, to select participants. In this context it is interesting to note that we found general symptoms were as prominent as specific eye symptoms, as might be expected from a heterogeneous group of undergraduates in an actual teaching session.

Although the pattern of symptoms generated across the two VERT sessions in our study was similar, in the second session VERT exposure did not significantly increase symptom scores. In part this was because the baseline scores for the second session were higher than in the first. This demonstrates the importance of having baseline measures. Because of this relatively high baseline in the second session, there was an increase in the number of participants who reported a reduction in scores post-VERT. There was no evidence of a build-up of symptoms over the two sessions.

A number of hypotheses would be consistent with our results. The first is that the rise in scores in the first VERT session was due to the novelty of the experience, but that even a relatively short exposure allows users to

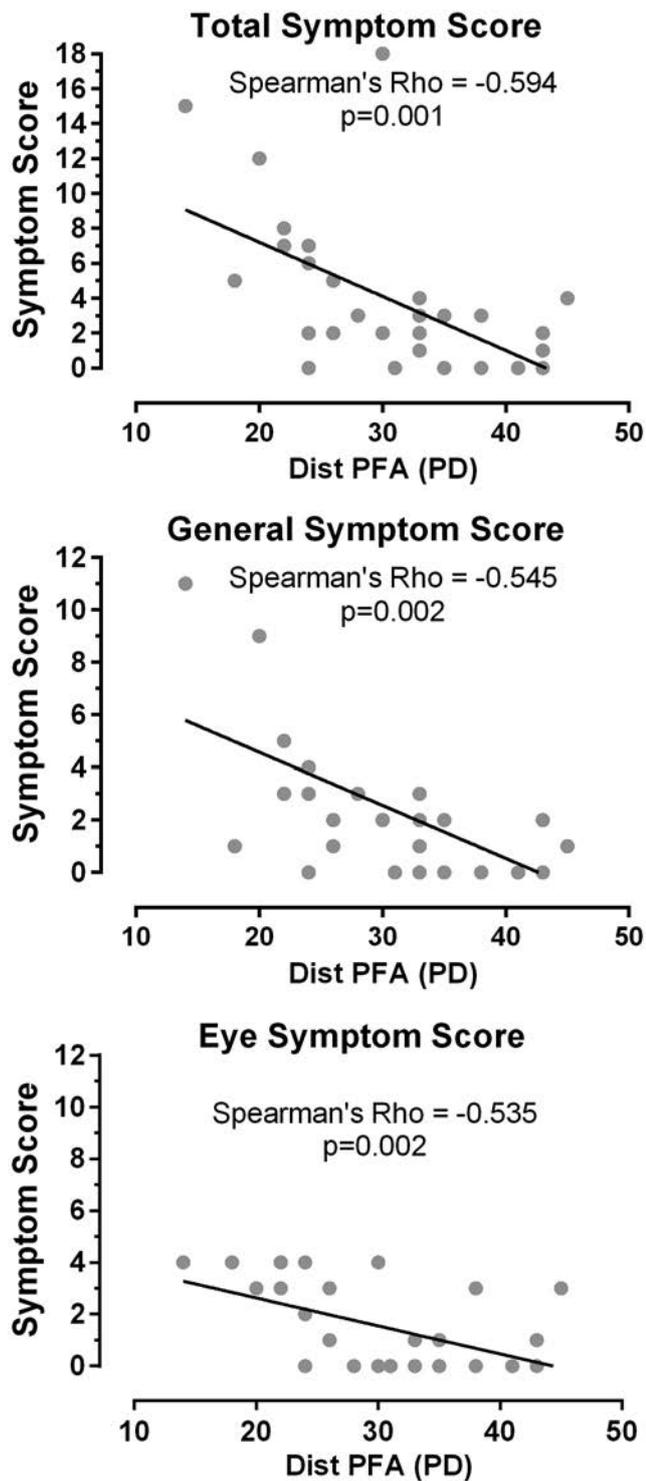


Fig. 5. Correlation between Total, General and Eye symptom scores and distance prism fusion amplitude (Dist PFA; PD). Data from the first session for 31 participants who had a vision assessment prior to the VERT session. The line is the least-squares linear regression line.

adapt. Subsequent effects are reduced. Secondly, in a situation in which the visual and eye impacts of using a system like the VERT are minimal, general factors might be expected to dominate. Given that we studied undergraduates in actual teaching sessions, the content of the

simulations, the style and standard of teaching, and other factors unrelated to the generation of visual symptoms, may have had a large influence on what was captured by the questionnaire.

Although the VRSQ was originally developed for studying the effects of VR systems using head-mounted displays (HMDs), we used it for three reasons. It was short and could be completed quickly and it addressed all the issues we thought relevant, while excluding others that feature in other simulator sickness questionnaires but we thought less relevant to the system we were examining. Thirdly, it has been shown previously that using it at baseline does not prime post-exposure responses.⁹ As we were interested in examining the change in symptoms pre-exposure to post-exposure, this was an important consideration. What it did not provide was information about the subjective feel of being in the VR environment provided by the VERT, sometimes called 'presence' in the VR literature. However, given the purpose of this particular system, this was of less concern to us in this study, and has been partly addressed previously.⁹

Our general result (minor symptoms resulting from exposure to the VERT) is consistent with other studies on less immersive VR systems, and those which do not employ HMDs. A direct comparison of HMD, desktop and projection systems using the SSQ demonstrated that the HMD system generated a higher level of symptoms than the others.⁶ Projected 3D environments in general seem to evoke mild, though measurable, symptoms.⁸ There is also evidence that prior experience of 3D projected environments can reduce visual symptoms,⁷ similar to our observation that in the second VERT session there was no statistically significant increase in post-VERT scores. This could be due either to habituation (in which participants develop a tolerance to stimuli with no change in their underlying function) or adaptation (in which there are changes in underlying functions). While adaptive changes in some aspects of binocular function have been reported previously,^{12,13} given the relatively brief nature of exposure in our study, and the pattern of questionnaire responses, it seems more likely that what we have observed is a habituation to the general features of the VERT.

Previous studies have reported that participants with weaker binocular status (i.e. presence of a phoria, reduced prism fusion amplitude) are more likely to experience symptoms of discomfort and fatigue while viewing stereoscopic displays because of the induced accommodation-vergence conflicts inherent in such contexts.^{2,13} It is thus interesting to note that the only clinical measure of function which generated a statistically significant correlation with symptom scores in our study was the distance prism fusion amplitude (dPFA).

While a variety of values for the 'normal' dPFA are quoted in the literature, a value of 22 (16 BO to 6 BI) for a participant group similar to the undergraduates we recruited was recently reported.¹⁴ The negative correlation between scores and dPFA shows that individuals with a reduced dPFA were those who experienced higher levels of symptoms. Individuals with a low dPFA might not experience symptoms in normal viewing environments. However, in a 3D simulated environment

participants must accommodate to a distance different to the vergence distance in order to achieve stereopsis (3D environments maintain constant accommodation while the vergence distance varies depending on the image contents), thus creating a vergence–accommodation conflict.² Because of the neural coupling between vergence and accommodation, this could result in visual discomfort generated by the attempt to resolve this conflict.^{12,13,15} Thus, we might expect participants with reduced dPFA (i.e. reduced zone of clear single binocular vision) to be more likely to be symptomatic.

It has been assumed that one of the main factors causing visual discomfort and fatigue while viewing 3D content may be due to the motor responses, i.e. trying to keep images clear (with accommodation) and single (with vergence).² Therefore if there is no attempt to make a motor response when there is a vergence–accommodation conflict, no discomfort will occur.² This is likely to be the reason why the participants with suppression in our study were not symptomatic.

Given that our participants were in effect a highly selected group (albeit selected on academic rather than functional visual criteria) care needs to be taken in extending our results to the population as a whole. This would require research on appropriately representative samples. However, as we have discussed above, projected environments in which users are relatively passive appear to cause little in the way of distressing symptoms. Research on larger, unselected groups would be useful, particularly if concerns about symptom generation by 3D cinema and TV persist. The other issue that such research could address would be the effects of longer sessions of exposure (as might be experienced from an evening spent in front of a 3D TV) and longer term exposure (i.e. over multiple sessions beyond the two we investigated).

Conclusion

A projected VR environment, in which users are relatively passive observers, is unlikely to induce symptoms that are problematic; this has been reported previously with a similar VR system.¹⁰ However, it should be assumed that in large, heterogeneous groups of participants, there will be a proportion who will derive no benefit from the VR display.

This work was supported by a UK College of Radiographers Industrial Partnership Scheme Research Award. The cooperation of the staff and students of the School of Health Sciences, University of Liverpool, in performing this study is acknowledged.

All authors warrant no commercial associations that might be perceived as a conflict of interest in connection with this paper.

This paper was presented as a poster at the annual meeting of the Association for Research in Vision and Ophthalmology (ARVO), May 2010, Fort Lauderdale, Florida.

References

1. The Daily Telegraph. Do 3D films make you sick? 2010. Available at: <http://www.telegraph.co.uk/health/6952352/Do-3D-films-make-you-sick.html> [last accessed 28 Oct 2013]
2. Shibata T, Kim J, Hoffman D, *et al.* The zone of comfort: predicting visual discomfort with stereo displays. *J. Vision* 2011; **11**: 1–29.
3. Regan E, Price K. The frequency of occurrence and severity of side-effects of immersion virtual-reality. *Aviat Space Env Med* 1994; **65**: 527–530.
4. Stanney KM, Mourant RR, Kennedy RS. Human factors issues in virtual environments: a review of the literature. *Presence: Teleoperators and Virtual Environments* 1998; **7**: 327–351.
5. Cobb SVG, Nichols SC, Ramsey AR, *et al.* Virtual reality-induced symptoms and effects (VRISE). *Presence: Teleoperators and Virtual Environments* 1999; **8**: 169–186.
6. Sharples S, Cobb S, Moody A, *et al.* Virtual reality induced symptoms and effects (VRISE): comparison of head mounted display (HMD), desktop and projection display systems. *Displays* 2008; **29**: 58–69.
7. Howarth PA. Potential hazards of viewing 3D stereoscopic television, cinema and computer games: a review. *Ophthalmic Physiol Opt* 2011; **31**: 111–122.
8. Pölonen M, Salmimaa M, Takatalo J, *et al.* Subjective experiences of watching stereoscopic Avatar and U2 3D in a cinema. *J Elect Imaging* 2010; **21**: 1–8.
9. Ames S, Wolffsohn J, McBrien N. The development of a symptom questionnaire for assessing virtual reality viewing using a head mounted display. *Opt Vis Sci* 2005; **82**: 168–176.
10. Flinton DM, White N. Preliminary findings on the Virtual Environment for Radiotherapy Training (VERT) system: simulator sickness and presence. *J Radiother Pract* 2009; **8**: 169–176.
11. Kennedy RS, Lane NE, Berbaum KS, *et al.* Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. *Int J Aviat Psychol* 1993; **3**: 203–220.
12. Emoto M, Niida T, Okano F. Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television. *J Display Tech* 2005; **1**: 328–340.
13. Lambooij M, Fortuin M, Ijsselstein W, *et al.* Measuring visual fatigue and visual discomfort associated with 3D displays. *J Soc Inf Disp* 2010; **18**: 931–943.
14. Rowe FJ. Fusional vergence measures and their significance in clinical assessment. *Strabismus* 2010; **18**: 48–57.
15. Hoffman DM, Girshick AR, Akeley K, *et al.* Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *J Vision* 2008; **8**: 1–30.