Ergonomics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/terg20

User experience while viewing stereoscopic 3D television

Jenny C.A. Read[a] & Iwo Bohr[b]
[a] Institute of Neuroscience, Newcastle University, Newcastle upon Tyne NE2 4HH, UK
[b] Campus for Ageing and Vitality, Newcastle University, Newcastle upon Tyne NE4 5PL, UK
Published online: 30 May 2014.

To cite this article: Jenny C.A. Read & Iwo Bohr (2014): User experience while viewing stereoscopic 3D television, Ergonomics, DOI: 10.1080/00140139.2014.914581

To link to this article: http://dx.doi.org/10.1080/00140139.2014.914581

Ergonomics, 2014
http://dx.doi.org/10.1080/00140139.2014.914581

Taylor & Francis
Taylor & Francis Group

User experience while viewing stereoscopic 3D television

Jenny C.A. Read[a]* and Iwo Bohr[b]

[a]*Institute of Neuroscience, Newcastle University, Newcastle upon Tyne NE2 4HH, UK;* [b]*Campus for Ageing and Vitality, Newcastle University, Newcastle upon Tyne NE4 5PL, UK*

(*Received 15 January 2014; accepted 27 March 2014*)

3D display technologies have been linked to visual discomfort and fatigue. In a lab-based study with a between-subjects design, 433 viewers aged from 4 to 82 years watched the same movie in either 2D or stereo 3D (S3D), and subjectively reported on a range of aspects of their viewing experience. Our results suggest that a minority of viewers, around 14%, experience adverse effects due to viewing S3D, mainly headache and eyestrain. A control experiment where participants viewed 2D content through 3D glasses suggests that around 8% may report adverse effects which are not due directly to viewing S3D, but instead are due to the glasses or to negative preconceptions about S3D (the 'nocebo effect'). Women were slightly more likely than men to report adverse effects with S3D. We could not detect any link between pre-existing eye conditions or low stereoacuity and the likelihood of experiencing adverse effects with S3D.

Practitioner Summary: Stereoscopic 3D (S3D) has been linked to visual discomfort and fatigue. Viewers watched the same movie in either 2D or stereo 3D (between-subjects design). Around 14% reported effects such as headache and eyestrain linked to S3D itself, while 8% report adverse effects attributable to 3D glasses or negative expectations.

Keywords: stereoscopic displays; 3D television; stereo vision; binocular vision; eyestrain; visual fatigue

Introduction

The last decade has seen a dramatic expansion of the use of stereo 3D (S3D) technology in entertainment and communication, including cinema, television, game consoles and mobile phones. However, anecdotal evidence and the manufacturers' own safety information has suggested for some time that S3D may have negative impacts on viewers, with symptoms such as headache, eye strain, dizziness and impaired motor coordination (Samsung Electronics, Seoul; 'Viewing TV using the 3D function'). Over the last few years, this topic has started attracting the attention of scientific researchers, with some evidence confirming that moderate adverse effects can be associated with S3D TV viewing (Yang et al. 2012; Yang and Sheedy 2011; Shibata et al. 2011; Lambooij et al. 2009).

There are several possible causes for these adverse effects. There is considerable evidence that visual symptoms such as eyestrain or blurred vision can be caused by the disruption of the natural relationship between binocular convergence and accommodation (Howarth 2011; Shibata et al. 2011; Yang and Sheedy 2011). This occurs because current S3D displays require viewers to maintain accommodation on the screen plane while verging in front of or behind it. Motion sickness can occur when video content suggests that the viewer is moving, while their vestibular system indicates they are not. Because S3D appears more real and immersive, such cue conflicts may be particularly troubling. Depending on the particular S3D protocol, unnatural timing between left and right eyes may produce depth artefacts or a perception of motion blur or judder (Hoffman, Karasev, and Banks 2010). Finally, S3D displays rarely depict the true horizontal and vertical disparities which a real object would produce. Over time, these subtle distortions might contribute to viewer discomfort (Banks et al. 2012).

We are not aware of any published work addressing which viewers are most likely to experience adverse effects while viewing S3D content. The American Academy of Ophthalmology (AAO) has suggested that adverse effects experienced while viewing S3D may reflect pre-existing visual disorders: 'If a healthy child consistently develops headaches or tired eyes or cannot clearly see the images when using 3-D digital products, this may indicate a vision or eye disorder' (AAO website http://www.aao.org/newsroom/release/20110118.cfm, retrieved 16 Nov 2012). The American Optometric Association also implies that problems with 3D may indicate visual disorders:

> The AOA recommends seeing a doctor of optometry for further evaluation if consumers answer yes to any of the following questions:
>
> - Do you experience eyestrain or headaches during or after viewing?
> - Do you feel nauseated or dizzy during or after viewing?

*Corresponding author. Email: jenny.read@ncl.ac.uk

- Are you more comfortable viewing 2D TV or movies instead of 3D TV/movies?
- Is it difficult for your eyes to adjust back to normal after watching 3D TV/movies?
- Do other people seem to be enjoying the 3D viewing experience more than you? (http://www.3deyehealth.org/, retrieved 19 March 2014)

The same website states that treatment for 'vision problems that interfere with viewing 3D content' 'often consists of wearing regular glasses, therapy glasses (with prism and multifocal lenses) and/or, Optometric Vision Therapy' (http://www.3deyehealth.org/faq.html#, retrieved 19 March 2014).

However, we are not aware of any published data suggesting that people with eye disorders experience more problems viewing S3D than conventional 2D content. Of course people with disorders of binocular vision may not be able to experience the S3D depth percept, but this would not in itself predict that they should be especially visually fatigued by S3D content as opposed to conventional 2D content or indeed viewing real scenes.

Similarly, few data are available on the prevalence of adverse effects with S3D. The above-mentioned studies were lab-based experiments, often using stimuli designed to cause discomfort, so may not necessarily apply to S3D content viewed for pleasure. An additional problem is that negative effects with S3D have been widely reported in the media, so many viewers may come to S3D with negative expectations. Given the subjective nature of many adverse effects, this could therefore become something of a self-fulfilling prophecy.

As part of a wider study, we acquired a range of data from 433 subjects, aged from 4 to 82 years old, which enables us to address the issues raised earlier. In this article, we report how often participants reported subjective adverse effects after viewing 3D TV, in a laboratory environment designed to resemble home viewing. We examined whether the likelihood of experiencing adverse effects relates to pre-existing eye conditions. In a between-subjects design, we compared active and passive S3D display technologies and attempted to disentangle genuine adverse effects from those caused by negative expectations.

Methods

In brief, 433 participants visited Newcastle University's Institute of Neuroscience, watched a film in either S3D or 2D (between-subjects design) and then reported a range of subjective judgements on the visual appearance of the film and whether they had noticed any adverse effects such as headache. On a separate occasion, most participants also visited a local optometry practice where they underwent a set of optometric and orthoptic tests listed in Table 3. Optometry relates to the general health of the eyes and quality of vision; orthoptics relates to the control of eye movements, and specifically binocular coordination. This enabled us to examine whether people with particular eye conditions were more or less likely to experience adverse effects. These experimental procedures are described in more detail later. The participants also performed tests of balance and coordination.

Ethics

The study was approved by the Newcastle University Faculty of Medical Sciences Ethics Committee (approval number 00431) and adhered to the tenets of the Declaration of Helsinki. All participants, or in case of children, adults with parental responsibility, gave written informed consent. Year of birth and gender were reported by the participants.

TV viewing conditions

Participants watched the animated film 'Toy Story' (1995, produced by Pixar Animation Studios, duration 80 minutes) in groups of up to five, grouped in family or friendship groups where possible. Although the TV viewing took place in a

Figure 1. TV viewing room. (A) location of the TV set; note: during the TV viewing, the ceiling lights were turned off so the only illumination comes from the bias lighting behind the screen. (B) location of the viewing space, including seating.

laboratory setting, efforts were made to approximate the experience of viewing at home (Figure 1). The room was 3.7 m × 2.9 m in size, and was furnished with a sofa, bean-bags, rug and pictures, and refreshments such as juice, popcorn and crisps were available during viewing. The viewing distance was ~2.5 m for participants seated on the sofa. The bottom edge of the TV screen was 85 cm above the ground. The TV was positioned in front of a light grey fabric background (2.8 m wide × 2.9 m high) forming the rear wall of the room (Figure 1(A)). White LED bias lights, screened from sight of the viewer, surrounded the edges of this background and illuminated it. During TV viewing, these LEDs and the TV itself provided the sole illumination. The luminance of the fabric background varied from around 25 candela per square meter (cd/m^2) at the edges, near the bias lights, to around 2 cd/m^2 at the centre, near the TV.

Subjective reports

After viewing the film, participants were taken to a separate reporting room and asked, 'How would you rate the visual appearance of what you watched today?'. The available answers were 1 ('dreadful'), 2 ('poor'), 3 ('not very good'), 4 ('acceptable'), 5 ('good'), 6 ('very good') and 7 ('fantastic'). For the benefit especially of child participants, a paper chart was available showing these options with appropriate cartoon faces (smiling/neutral/frowning). In each case, the answer was entered into the computer interface by the research assistant. Most participants were also asked 'Specifically, how realistic did you find the 3D depth?', with the same available answers. Any other comments spontaneously volunteered by the participant were also recorded. Participants were then asked 'Did you experience any unpleasant effects or sensations?'. If they answered yes, they were then given the option of choosing from the following list:

- blurred vision
- difficulty focusing eyes
- discomfort in nose/face/ears
- cramps
- double vision
- eyestrain
- faintness
- fatigue
- fever
- headache
- headache behind eyes
- impaired coordination
- impaired balance
- itching
- joint pain
- muscle pain
- nausea
- skin rash
- stomach ache
- tooth ache
- other.

Participants who chose 'other' were then asked to specify this, and the answer was entered onto the computer. This list was chosen to include items which have been associated with S3D, either in the scientific literature, e.g. eyestrain (Lambooij et al. 2009), or elsewhere, e.g. cramps (Samsung, Inc.; 'Viewing TV using the 3D function') and those where no such link is expected, e.g. toothache.

Television sets

The TV sets used were manufactured by LG Electronics (www.lg.com): model 47LX6900 (active S3D) and model 47LD920-2A (passive S3D). Details of these TVs are given in Table 1. They use different technologies to display S3D images when used in 3D mode. The active 3D TV displays left and right images temporally interleaved, with each eye's image being refreshed at 60 Hz. To view this type of S3D content, viewers must wear active 3D glasses, which are powered by a battery and are therefore somewhat bulky. The manufacturer did not make child-size active glasses, so child participants wore adult-sized glasses to view active S3D.

The passive 3D TV uses a patterned retarder, which means that left and right images are displayed spatially interleaved on alternate pixel rows of the display. This halves the vertical resolution of the display in each eye. Viewers wear passive 3D

Table 1. Specifications for the two television sets used in the study.

	Television set	
TV manufacturer	LG	LG
TV model	47LX6900	47LD920
3D technology	Active shutter (temporally interleaved)	Passive polarised (patterned-retarder)
Viewed by groups	A and C–E	B
Screen size (inches along the diagonal)	47	47
Display type	LED	LCD
Resolution (width × height, pixels)	1920 × 1080 (full HD)	1920 × 1080 (full HD)
Contrast ratio	8,000,000:1	150,000:1
Audio output	10W + 10W	10W + 10W
Dimensions of set (without stand), width × height × depth (mm)	1127 × 692 × 29.3	1173.4 × 723.4 × 100.8

Note: The last six rows use information provided by the manufacturer, LG Electronics.

glasses, whose lenses are circular-polarising filters. Because passive glasses are not powered, they are much lighter than active glasses. Smaller passive glasses were available for children.

Experimental groups

Ideally, we would have used a double-blind design, in which neither the experimenter nor the participant was aware whether they were watching 2D or S3D content. This was not practical in our study. It was impossible to 'blind' the experimenters, as they were responsible for setting up the correct content. We expected that it would also be essentially impossible to 'blind' the participants, since we thought it would be obvious to them whether they were viewing 2D or S3D content. Thus, initially, participants were assigned in alternation to one of three TV groups: A, B and C (participants from the same family or friendship group were all assigned to the same TV group, so they could do the experiment together). The A group viewed 'Toy Story 3D' on the active 3D TV 47LX6900, wearing active 3D shutter glasses. The B group viewed 'Toy Story 3D' on the passive 3D TV 47LD920-2A, wearing passive 3D glasses. The C group viewed 'Toy Story' in 2D on the active 3D TV 47LX6900, operated in its 2D mode. The C group did not wear any 3D glasses.

However, when initial results indicated that adverse effects were substantially higher in the S3D group, we became concerned about a possible nocebo effect. A nocebo effect is the opposite of a placebo effect, when an intrinsically harmless substance or procedure causes adverse effects due to negative expectations. We therefore experimented with 'fake 3D', where people were shown 2D content while wearing 3D glasses. We had expected that this would be impractical, because participants would realise that the 3D was 'not working' and break off watching to complain. To our surprise, participants accepted this manipulation, apparently assuming they were watching 3D content, although they were not told this. As a result, we were able to collect data from two smaller control groups, D and E. Both groups viewed 2D content on the active 3D TV 47LX6900 operating in 2D mode. The D group wore active shutter glasses, although since the shutter function requires an infrared signal which the TV set broadcasts only in 3D mode, the glasses were not shuttering during viewing. The E group wore passive 3D glasses. After viewing, these groups were asked about visual appearance and depth realism as described above. Initially C group participants were not asked about depth realism, but when we introduced the D and E groups, we began asking all participants this question.

Because the D and E groups were recruited later than the other groups, their demographics differed: they were predominantly university students, whereas the A–C groups also included many more people from the wider population. The five groups are summarised in Table 2.

Table 2. Demographic data of all participants and broken down into three major categories (SD, 2D, 'fake 3D') as well as particular experimental groups.

	Content viewed	Glasses worn	Number of participants	Age in years: mean, median, interquartile range	Gender breakdown: male, female, not recorded
A	Active S3D	Active 3D, shuttering	115	27.6, 23.0, 17.0–38.8	46M, 63F, 6NR
B	Passive S3D	Passive 3D	131	28.4, 24.0, 21.0–34.0	54M, 69F, 8NR
C	2D	None	122	26.3, 23.0, 19.0–33.0	48M, 69F, 5NR
D	2D	Active 3D, not shuttering	33	25.8, 23.0, 22.0–30.3	8M, 22F, 3NR
E	2D	Passive 3D	32	23.8, 23.0, 22.0–25.0	9M, 20F, 3NR

Notes: Groups within the same category are depicted in the same colour. To aid with anonymisation, we only recorded the year of birth, so 'age in years' actually means the difference between year of birth and the year of the study (2011).

Recruitment

In total, 433 participants took part in the study (see Table 2 for demographic data). A total of 246 participants watched 3D TV (A and B groups) and 187 watched 2D TV (C–E groups).

Participants were recruited via the Newcastle University Institute of Neuroscience Research Volunteer Database, an email list of local people who are interested in research and willing to participate in experiments. Further participants were recruited by an advertisement in a local newspaper, by word of mouth and by snowball recruitment. Some participants were recruited via a mailshot to 1000 BSkyB customers in the Newcastle area.

Initially, participants were assigned in alternation to the A–C groups, in the order in which they contacted us. This ensured that sampling was similar in these three groups. After the decision to introduce the additional D and E control groups, we wanted to recruit participants into these groups rapidly, and so later participants were assigned to D and E in alternation. Thus, the D and E participants were drawn from the sample of late-entrants only, and have a somewhat different demographic profile.

People who reported having photosensitive epilepsy were excluded from participating on safety grounds, even though there is no evidence that S3D presents a specific risk to this group (Prasad et al. 2011).

Eye tests

Eye tests were carried out at local optometry practice C4 Sightcare (www.C4sightcare.com), at either their Newcastle or their Morpeth site. All tests were run by qualified optometrists or orthoptists as appropriate. Any existing optical correction was recorded along with any medication the participant reported taking. The orthoptic examination probed aspects of binocular function which are particularly relevant for S3D content. Many of these eye tests are performed at two viewing distances: a short distance (30–80 cm in our case), corresponding to typical reading distance, and a long distance: 6 m. For the human visual system 6 m is effectively infinity, as the convergence and accommodation at 6 m do not differ significantly from those at infinity. Table 3 lists the optometric and orthoptic tests carried out.

In some cases, the eye examination revealed possible areas of concern, ranging from the need for a new glasses prescription to possible undiagnosed eye disease. In these cases, the participant was referred on to the appropriate health care provider. All participants were made aware of this potential outcome in the information sheets provided. Not all participants chose to complete the study by visiting C4 Sightcare for their eye tests. Out of the 433 participants, we have optometric data for 339 (78%) and orthoptic data for 333 (77%).

Statistical analysis

Analysis was carried out in the Matlab programming environment (Matlab R2012a; Mathworks, Inc., Sherborn, MA, USA), using custom scripts. Our data were not normally distributed, so we used non-parametric tests for significance including the Kruskal–Wallis test and Mann–Whitney *U*. Whether or not a participant reported adverse effects is a binomial variable, so to compare the rate of adverse effects between two groups of participants, e.g. 2D versus S3D, we used binary logistic regression with group as a factor.

Results

Participant demographics and TV viewing habits

Table 2 reports the number of participants in each of the five groups. There are over 100 participants in each of the three main groups (A–C). The second two control groups, D and E, where participants wore 3D glasses while watching 2D TV, were introduced only later on in the study and therefore have fewer participants.

Not all participants completed the more detailed recruitment questionnaire or went for the requested eye examinations. We have recruitment questionnaires for 342/430 participants: orthoptic data for 287/430 and optometric data for 277/430. In each figure, we report results for the subset of participants for whom that information is available.

Figure 2 shows the distribution of participant ages for the 433 participants, broken down by TV group. To reduce the amount of identifying information stored, we did not record participants' day or month of birth, so 'age' was estimated by subtracting the year of birth from 2011, the year the study took place. The horizontal axis shows age in years; the height of each bar shows the fraction of participants in that group aged within 5 years of the age indicated on the horizontal axis. Nearly half our participants were born within 5 years of 1990, so people in their 20s are over-represented in our sample. The late-recruited D and E groups contain fewer children and old people. They were also more highly educated on average, being recruited largely from university students. Thus, unfortunately, comparisons between the A + B + C groups and the

Table 3. Optometric and orthoptic tests carried out for the study.

Eye test	Brief description and explanation
	Orthoptic tests (relating to control of eye movements and binocular vision)
Frisby stereotest at 30–80 cm Frisby-Davis stereotest at 6 m (FD2)	Estimates the smallest binocular disparity between two objects which the person is able to distinguish. Measures the quality of 3D stereo vision.
Abnormal head posture	To detect abnormal head posture which could indicate a problem with vision, e.g. chin elevation resulting from ptosis.
Ocular motility	Records any obvious problems with eye movements.
Cover test at 33 cm and 6 m	Performed at a viewing distance of 33 cm and 6 m to detect any abnormalities of binocular control. The participant is asked to fixate an object at the desired distance, and the orthoptist then covers and uncovers each eye in turn. If the participant has good binocular control, no movement of the eyes is visible as they are each covered and uncovered. If the participant has a tropia (a manifest squint), she/he will not be able to fixate the object with both eyes. In this case, when the fixating eye is covered, the other eye visibly moves in order to take up fixation. If the participant has a phoria (a latent squint), correct fixations occur with both eyes, but when one eye is covered, it will drift into its preferred position. If the cover test revealed tropia or phoria, the orthoptist then used a prism bar to quantify the extent of the deviation in prism-dioptres, both horizontally and vertically.
Near point of convergence	Measures the eye muscles' ability to converge the eyes, using an RAF rule.
	Optometric tests (relating to general eye health and vision)
Refractive error at 0.4 m and 6 m	Optometrist measures the refractive error of each eye at a viewing distance of 0.4 m and again at 6 m.
Monocular and binocular visual acuity at 0.4 m and 6 m	For participants aged 8 years and over, visual acuity with the left eye, right eye and both eyes were measured at 0.4 m and at 6 m, in each case using the best optical correction for that participant at that distance, as determined in the measurement of refractive error. At 6 m, visual acuity was measured again with the participant wearing their habitual optical correction (i.e. their usual glasses or contact lenses, or without glasses/lenses if they do not usually wear any). Visual acuity was measured in logMAR units; at 0.4 m, using the printed Sussex logMAR test; at 6 m, using the Thomson logMAR test administered on a computer. This was a total of nine acuity measurements, which was too demanding for young participants. For participants aged 7 years and under, monocular and binocular visual acuity was measured at 3 m using the Keeler logMAR test and the participant's habitual optical correction.
Intra-ocular pressure	If elevated, this can indicate eye disease such as glaucoma.
Fundus exam and photograph	Includes examination of the fundus, the interior surface of the eye including the retina, optic disc, macula and fovea. A photograph of the fundus was taken and a note of any abnormalities made.

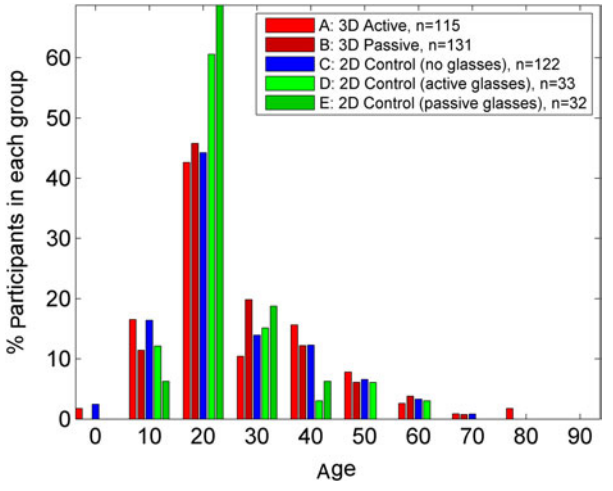

Figure 2. Distribution of ages, separated out by TV group. The bars show the percentage of participants in that group aged within 5 years of the age shown on the horizontal axis. Participants in the '0' bin were aged 5 or under. All the bars of a given colour sum to 100%.

D + E groups are complicated by sample differences. All groups contained greater numbers of female than male participants (overall 243 female to 165 males; gender information was not recorded for a further 25 participants).

The recruitment questionnaire asked participants about their typical viewing habits. Our randomisation procedure was intended to ensure that the five groups are comparable. Figure 3 shows self-reported average daily TV viewing. TV viewing time was self-reported on a five-point scale, from 'less than 60 minutes' to 'more than 5 hours'. We see that the three main groups report similar amounts of time spent watching television ($p = 0.07$, Kruskal–Wallis test on 368 A–C participants only, with TV group as a factor). However, there are significant differences between the participants of A + B + C groups and the late-recruited D + E groups ($p < 10^{-5}$, Kruskal–Wallis test on all five TV groups as a factor). The D and E participants watch less TV, typically under an hour a day. This may be related to their higher educational level, as we also found an inverse correlation between highest educational qualification and amount of time spent watching TV.

Clearly, for this study it was also critical to ask how often participants usually view S3D displays. Figure 4 shows this information, in the same format as Figure 3. Most participants view S3D content only a few times a year. The A and E groups watched S3D content slightly more often at recruitment (median = 'a few times a year' for A and E; 'less than once a year' for B–D; $p = 0.005$, Kruskal–Wallis test with TV group as the factor).

Visual appearance

After viewing the movie, each participant was asked to rate the visual appearance of the TV on a seven-point Likert scale ranging from '1 = dreadful' to '7 = fantastic'. Figure 5 shows these responses for the five groups specified in Table 2. Note that all groups other than B were watching the same physical TV set. For the A group, it was set to its 3D mode; for the C–E groups it was displaying 2D.

The responses are quite similar for all five groups, but it is clear that the 2D C group gave more of the very highest ratings whereas the 3D A group gives more of the very lowest ratings. We found a highly significant effect of TV group ($p < 10^{-6}$, Kruskal–Wallis test). The C group, which viewed 2D, differed significantly from both 3D groups ($p < 10^{-6}$, A + B vs. C, Mann–Whitney). This is particularly striking as the active 3D (A) and 2D control groups were viewing the same physical TV set; the only difference was whether it was set to 3D or 2D mode. There was also a highly significant difference between active and passive S3D ($p < 10^{-4}$, A vs. B, Mann–Whitney), with the passive 3D TV appearing slightly better.

One possible interpretation of these results is that 3D glasses reduce the quality of the visual appearance, probably by reducing the luminance. The fact that passive 3D was rated the same as 2D content viewed through passive glasses ($p = 0.95$, B vs. E, Mann–Whitney) suggests that it is the glasses that are responsible for the generally lower ratings given to S3D content, rather than the S3D itself. Consistent with this interpretation, the visual appearance was rated slightly lower by the 2D groups wearing 3D glasses than by the 2D group without glasses ($p < 0.01$, D + E vs. C, Mann–Whitney). The higher rating given to passive 3D as compared to active 3D may be due to flicker introduced by active 3D shutter glasses. Consistent with this interpretation, the A group (viewing active 3D with shutter glasses switched on) reported poorer visual appearance than the D group (viewing 2D content on the same TV, with shutter glasses switched off; $p = 0.003$, A vs. D, Mann–Whitney).

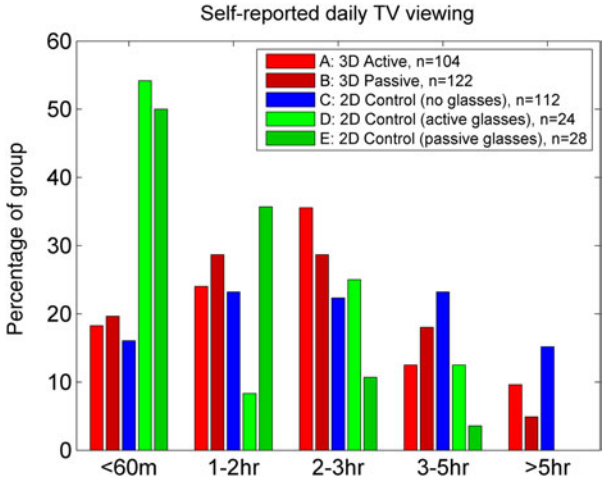

Figure 3. How much time participants in the five different groups reported watching TV. In the recruitment questionnaire, typical daily TV viewing time was self-reported on a five-point scale, from 'less than 60 minutes' to 'more than 5 hours'.

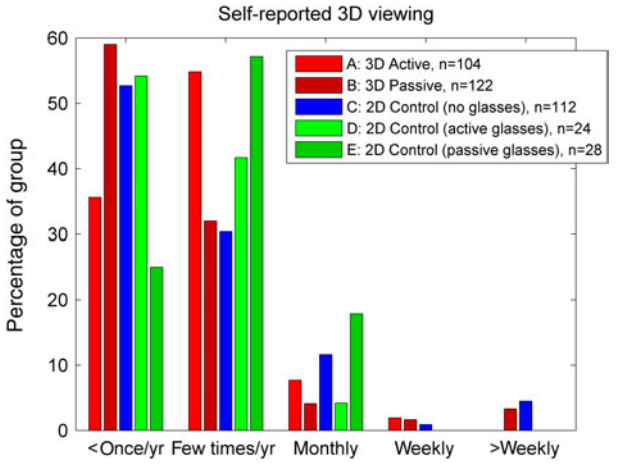

Figure 4. Frequency of exposure to 3D displays. Frequency was self-reported on a five-point scale, from 'less than once a year' to 'more than once a week'.

Depth realism

Participants were asked to rate the realism of the 3D depth on the same seven-point Likert scale. Initially, when there were only three groups, only 3D participants (groups A and B) were asked this question, as it seemed to be meaningless for the participants who viewed 2D content (control group C). When the additional 'fake 3D' control groups were added (D and E), we started asking this question of all participants. This is why there are fewer responses available for group C.

Figure 6 shows the judgements made regarding depth. Here, we again observe a highly significant effect of TV group ($p < 10^{-8}$, Kruskal–Wallis test). Unsurprisingly, this is driven by differences between the 3D and 2D groups. The 3D groups are most often rated 'good' or 'very good', but the 2D groups are mainly rated 'acceptable'. There was no significant difference between the two 3D groups ($p = 0.52$, A vs. B, Mann–Whitney), but there was a very highly significant difference between the 3D and 2D groups ($p < 10^{-6}$, A + B vs. C + D + E, Mann–Whitney). This is reassuring, as it confirms that S3D does produce a substantial improvement in depth realism, even when viewers are not aware whether they are watching 2D or S3D.

Subjective adverse effects

We also asked participants to report any adverse effects they experienced while viewing the TV. Figure 7 shows the percentage of participants who reported one or more adverse effects. In the three 2D groups (C–E; blue and green bars), 4%

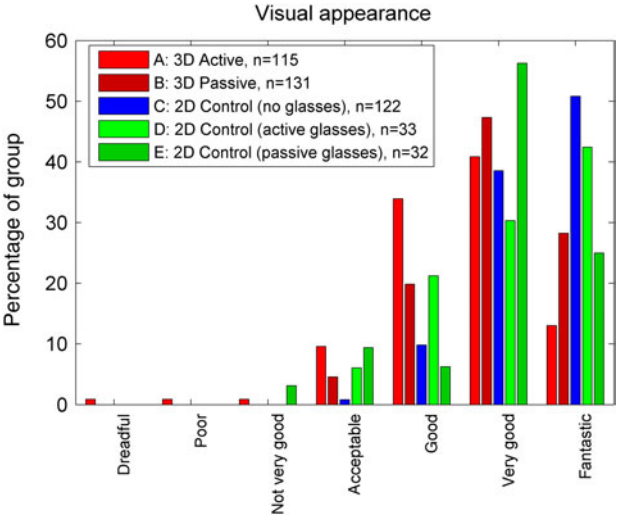

Figure 5. Judgements made regarding visual appearance, for the five TV groups. Ratings were made on a seven-point Likert scale; the description given to each point on the scale is shown on the right.

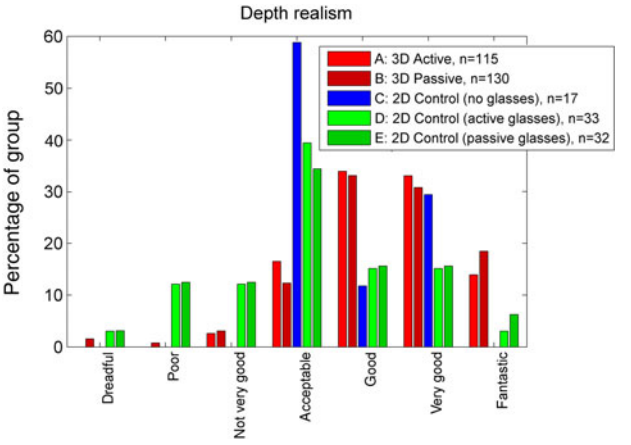

Figure 6. Judgements made regarding depth realism, for the five TV groups. Details as for Figure 5. Note the low number of C group participants for whom these data were recorded.

(8 out of 187) of participants reported experiencing adverse effects. In the two S3D groups (A and B; red bars), this rose to 24% (58 out of 246 participants). There was no significant difference in the rate of adverse effects between the active and passive 3D groups ($p = 0.39$, binary logistic regression), but the difference between both 3D groups together and the 2D C group was highly significant ($p < 10^{-5}$, A + B vs. C + D + E, binary logistic regression).

A nocebo effect contributes but is not solely responsible

We wondered whether the high rate of reported adverse effects of S3D could be due, at least partially, to expectations. There have been widespread media reports linking S3D to a range of adverse effects, so perhaps people expect to feel adverse effects when they view 3D, and this leads them to report more adverse effects. This would be an example of a nocebo effect. As noted in the 'Methods' section, we had originally regarded even a single-blind design as impractical. However, when initial results indicated that adverse effects were so much higher in the S3D groups, we decided to introduce the two 'fake 3D' control groups. As described earlier, in groups D and E, participants viewed 2D TV while wearing active or passive 3D glasses. They were not told they were viewing 3D TV, but from comments made to the experimenters, many of them apparently assumed that they were. There was no significant difference between the rate of adverse effects in the two 'fake 3D' groups.

The addition of these groups enables us to estimate the contribution of any nocebo effect to the complaints of adverse side effects of 3D viewing. Grouping together all 65 participants in the D + E 'fake 3D' groups, the adverse effect rate was $A_{fake3D} = 6/65 = 9.2\%$. This is significantly higher than the rate of adverse effects in the 'known 2D' control group C:

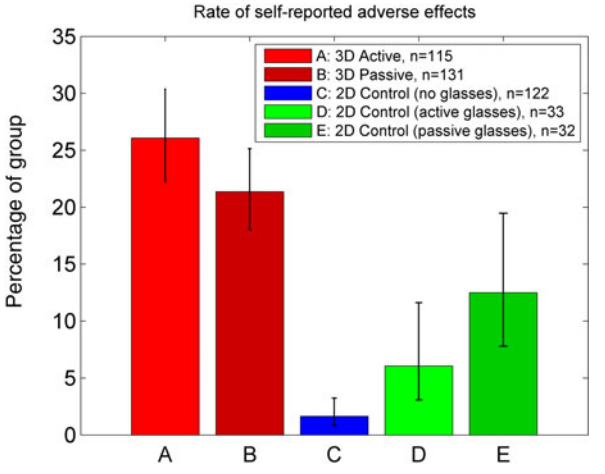

Figure 7. Frequency of adverse effects. Bars show percentage of participants who reported experiencing one or more adverse effects, for the five groups specified in Table 2. Error bars show the 68% confidence interval assuming simple binomial statistics.

$A_{2D} = 2/122 = 1.6\%$ ($p = 0.03$, C vs. D + E, binary logistic regression). There are two potential explanations for this. One is the nocebo effect previously mentioned. The other possibility is that simply wearing 3D glasses caused some adverse effects, independent of the 3D content. For example, this could be due to the lower luminance.

However, even wearing glasses, only 9% of the 'fake 3D' groups D + E reported adverse effects. This is very significantly less than the real 3D groups A + B, where $A_{real3D} = 58/246 = 23.6\%$ ($p = 0.01$, A + B vs. D + E, binary logistic regression). As noted earlier, one reason for this may be the different composition of the D + E groups as compared to the A + B + C groups. However, it seems possible that viewing S3D content is associated with adverse effects, over and above any effect simply of wearing the glasses or of negative expectations of 3D. We carried out a binary logistic regression with two categorical factors: whether or not participants viewed S3D (set to 1 for groups A + B and 0 otherwise) and whether or not they believed they were viewing S3D (set to 1 for groups A + B + D + E and 0 for group C). Both factors were significant, with $p = 0.01$ and $p = 0.03$, respectively. This indicates a significant effect both of S3D itself, and of a nocebo effect.

Ignoring sample differences between the A + B + C and D + E groups, we can take A_{2D} as an estimate of the baseline rate of reporting adverse effects in an experimental setting like ours; we can take ($A_{fake3D} - A_{2D}$) as an estimate of the nocebo effects produced merely by the belief that one is viewing S3D, and we can take ($A_{real3D} - A_{fake3D}$) as an estimate of the effects actually due to S3D. This produces the following estimates:

(1) Around 2% (A_{2D}) of people report adverse effects after viewing 2D TV. This includes any effects specifically due to viewing television or sitting in a dark room for over an hour, plus simply people who happen to have a headache that day, etc.
(2) An additional 8% ($A_{fake3D} - A_{2D}$) of people report adverse effects after watching 2D TV with 3D glasses while believing it to be 3D. This could be due to negative preconceptions regarding S3D, or to some factor associated with the glasses, e.g. the reduction in luminance.
(3) An additional 14% ($A_{real3D} - A_{fake3D}$) of people report adverse effects if they *have* actually viewed 3D TV.

This suggests that around 14% of a typical population experience some form of adverse effect due specifically to S3D content. However, given the sample differences between our groups, e.g. in education level and frequency of TV viewing, this conclusion must be regarded as tentative.

Headaches and eyestrain are the most common adverse effects

We next examine the types of adverse effects reported by our participants. Figure 8 shows the probability of reporting different types of adverse effects, i.e. the number of participants in a group who reported each type of adverse effect, divided by the number of participants in the group. As described in the 'Methods' section, participants were offered an array of possible descriptions to choose from, or could supply their own. To make this more manageable, in Figure 8 we have combined similar complaints and active/passive TV groups. The label 'headaches' describes people who selected either 'headache' or 'headache behind eyes'. The label 'eyes' covers 'blurred vision', 'difficulty focusing eyes' and 'eyestrain'. The label 'glasses' covers 'discomfort in nose/face/ears', which was included based on pilot studies where some

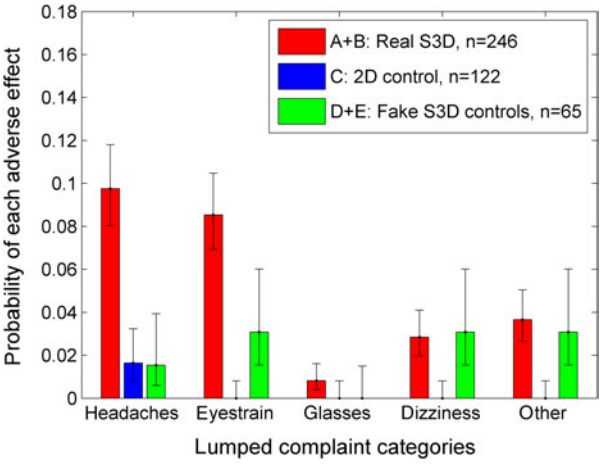

Figure 8. Probability of experiencing different classes of adverse effects. See text for details concerning the categorisation of adverse effects. The error bars represent 68% confidence intervals, computed with the simple binomial test.

participants complained that the active 3D shutter glasses were uncomfortable to wear. The label 'dizziness' covers 'impaired balance', 'impaired coordination', 'faintness' as well as 'dizziness'. 'Dizziness' was not included as an option in the list of possible effects participants were shown, but a few participants gave it under 'other'. Finally, 'other' in Figure 8 covers 'nausea' (reported five times), 'cramps' (reported once), 'fatigue' (four times) and 'fell asleep during movie' (once).

Participants were free to report as many adverse effects as they liked. The maximum number of adverse effects reported by any one participant was four (participant L2E001 in the 'fake 3D' group, who reported fatigue, faintness and impaired balance and coordination). In generating Figure 8, multiple descriptions of the same type of adverse effect were counted only once. For example, participant L2E001's reports of faintness, impaired balance and impaired coordination added one increment to the 'dizziness' category. In practice, such decisions make little difference given that out of 66 participants who reported any adverse effects, 60 (91%) reported only one adverse effect.

The most frequent types of complaints in the S3D groups (A + B) were headaches and eyestrain. These symptoms were reported much more often in the S3D groups than in any of the 2D groups, including those where people wrongly believed they were watching S3D. In these C + D + E groups, the probability of reporting a headache was around 2% (3 out of 187); in the A + B groups, it was around 10% (24 out of 246). This is a significant difference ($p = 0.02$, simple binomial statistics; under the null hypothesis that the probability of headache is the same in all groups, at 27/433, the probability that $>23/246$ AB participants would report headache is $p = 0.02$). This finding is consistent with previous literature suggesting that eyestrain and visual fatigue can be caused by S3D content, perhaps due to the violation of the natural relationship between accommodation and vergence (Howarth 2011; Shibata et al. 2011; Yang and Sheedy 2011; Hoffman et al. 2008; Lambooij et al. 2009; Shibata et al. 2011). In contrast, in the 'fake 3D' groups (D + E), dizziness and other effects such as nausea were reported as often as headache and eyestrain. These were not reported so often by either the true S3D groups, or by 2D viewers who knew they were watching 2D. Faintness and dizziness have not to our knowledge been linked with S3D in the scientific literature, but have often been linked to S3D in the media and by manufacturers. For example, a report in the British newspaper The Telegraph in 2010 suggested that the film *Avatar* could cause 'extreme dizziness' (http://www.telegraph.co.uk/health/6952352/Do-3D-films-make-you-sick.html, retrieved April 22 2013). We speculate that at least some of these reported symptoms may represent the nocebo effect discussed earlier: participants may have expected these symptoms based on what they had previously read about S3D.

Effect of gender

Combining all TV groups, there was no gender difference in the reporting of adverse effects ($p = 0.07$, binary logistic regression with gender as the factor). However, when we analyse only the S3D groups, women were slightly more likely to report adverse effects with S3D. Our data-set contains 232 participants who viewed S3D TV and for whom gender information was recorded: 132 female and 100 male. 30% of the females reported adverse effects, compared with 17% of males (Table 4). This was marginally significant ($p = 0.03$; binary logistic regression with gender as the factor). However, when we compute a binary logistic regression with gender and 'S3D viewing' (whether the person was in groups A + B or C + D + E) as categorical factors, the main effect of both gender and S3D were significant, but the interaction between them was not ($p < 10^{-4}$ for S3D, $p = 0.03$ for gender, $p = 0.73$ for S3D*gender).

Adverse effects are not predicted by eye or vision problems

For the groups who viewed S3D content, we examined the results of the eye examinations to see if we could detect any relationship between pre-existing eye problems and the likelihood of reporting adverse effects with S3D. We pooled all S3D participants (A + B) and then grouped them into 'adverse' and 'none' subgroups: those who reported adverse effects and those who did not. We looked for significant differences between the results of the eye tests for 'adverse' and 'none'. We also approached the data from the other direction; that is, for each eye test we grouped people into those who 'passed' and

Table 4. Gender differences in adverse effects with S3D.

Percentage who reported adverse effects	Female	Male	Significance of gender difference, p
S3D (A + B groups)	30% (39/132)	17% (17/100)	0.03*
2D (C–E groups)	5% (5/111)	3% (2/65)	0.64
Fake 3D (D + E groups)	10% (4/42)	12% (2/17)	0.84
All groups (A–E)	18% (44/243)	12% (19/165)	0.07

Note: We assessed the significance using binary logistic regression with gender as the factor.
*Indicates significance $p < 0.05$.

those who could be considered as having some problem. For example, we grouped people according to whether they had 'normal' or 'poor' visual acuity, and looked to see if these two groups differed in their likelihood of reporting adverse effects. We did this for many different visual tests (e.g. normal vs. poor stereoacuity, no phoria/tropia vs. those with phoria/tropia) and for many different definitions of 'normal' versus 'poor'. Similarly, there was no significant difference between the ages of people reporting adverse effects versus those who did not ($p = 0.98$, Mann–Whitney U test). With the exception of gender, discussed in the 'Effect of gender' section, we could not identify any significant relationships which would enable us to predict in advance which participants would experience adverse effects.

Discussion

Our results confirm previous reports (Yang et al. 2012) that a small number of viewers may experience minor adverse effects after viewing around an hour of S3D TV. Our work differs from previous studies in that it attempts to control for negative expectations regarding 3D. In addition, it may have more ecological validity than other lab studies, since it was carried out in a relatively natural setting, watching a real 3D movie such as people view at home.

We find that around 14% of viewers report adverse effects which appear to be directly related to 3D. In agreement with previous work (Bando, Iijima, and Yano 2012; Hiruma and Fukuda 1993; Hoffman et al. 2008; Howarth 2011; Lambooij et al. 2009; Nojiri et al. 2004; Shibata et al. 2011; Solimini et al. 2012; Yang and Sheedy 2011; Yano, Emoto, and Mitsuhashi 2004; Yano et al. 2002), we report that the symptoms most commonly associated with S3D were headache and eyestrain.

We did not find any evidence to support previous suggestions that adverse effects with S3D may indicate problems with the eyes or with binocular vision. This is not surprising. Adverse effects with S3D appear to stem from cue conflicts between the depth information provided by binocular disparity and other cues, for example accommodation, motion parallax and vestibular input. Thus, individuals' differences in sensitivity to S3D would be expected to reflect factors such as their tolerance for cue conflict, rather than low-level visual abilities. Indeed, there are good reasons to expect visual pathology to reduce the probability that an individual would experience problems with S3D, rather than to increase it. In the extreme case, someone who is blind in one eye could experience no problems due to S3D itself (they could of course experience problems caused by 2D content, or by the glasses, e.g. flicker). People with binocular eye disorders such as strabismus are much more likely to experience inappropriate disparities in their everyday life, and their visual systems have developed mechanisms to compensate for this, e.g. suppression of one eye's input (Serrano-Pedraza, Clarke, and Read 2011; Jampolsky 1955; Von Noorden and Campos 2002). More generally, the fact that cue combination is generally close to statistically optimal would suggest that the less reliable the visual input, the more cue conflict should be tolerated. This would imply that people with visual problems should be less, rather than more, likely to experience adverse effects with S3D. In fact, our data revealed no effect either way.

We also could not detect an effect of age, in contrast to a recent study. Yang et al. (2012) found that participants aged over 45 reported more dizziness and nausea after viewing 2D as compared with S3D, whereas younger participants reported more blurred and double vision, dizziness, and nausea after viewing S3D as compared with 2D. As those authors point out, there are theoretical reasons for expecting older individuals to experience fewer problems with S3D content. The vergence/accommodation conflict has been identified as a key reason for discomfort in S3D displays (Yang and Sheedy 2011; Shibata et al. 2011; Yano and Emoto 2002; Hoffman et al. 2008; Emoto, Nojiri, and Okano 2004). People older than 45 or so are presbyopic, i.e. have lost the ability to accommodate. They therefore routinely experience a mismatch between vergence and accommodation in everyday life. One would imagine, therefore, that they should be less troubled by vergence/accommodation conflict in S3D displays. As Figure 2 shows, our participants were disproportionately in their 20s. Thus, our failure to detect an effect of age may reflect a lack of power. However, the absolute number of older participants was comparable to Yang et al. (we had 45 participants aged 46 years or over, they had 50).

We did find a small effect of gender, with females being more likely to report adverse effects after watching S3D. Yang et al. (2012) also reported a significant effect of gender, with women reporting worse adverse effects than men. However, a binary logistic regression indicated that our data are consistent with the possibility that women are slightly more likely to report adverse effects in all conditions, and that both men and women are more likely to report adverse effects after viewing S3D than after 2D, with no gender difference relating specifically to S3D. One factor to take into account is that the average female inter-pupillary distance is about 0.96 that of males (Dodgson 2004). In principle, this could affect women's experience of S3D. Disparities encountered in natural viewing scale with inter-pupillary distance, so are generally smaller for women. Disparities in S3D content are controlled by the camera parameters, and so on average would be slightly larger, relative to natural disparities, for women as compared with men. Conceivably, this could contribute to making adverse effects more likely in female viewers.

The reason for the discomfort some viewers experience with S3D is not clear. As discussed, vergence/accommodation conflict has been identified as one potential reason, but creators of S3D content are well aware of this issue and work hard to

keep disparities small. In our experiment, the viewing distance was about 2.5 m or a focal distance of 0.4 D. According to Figure 26 and Equation (7) of Shibata et al. (2011), we would expect no significant discomfort for content presented behind the screen, and discomfort for near content only when this is presented closer than 1 m, i.e. 1.5 m in front of the screen. This would require a negative screen parallax of nearly 10 cm or − 8% of the screen width. This flouts industry guidelines; for example, Sky's recommended depth budget is 3% (comprising positive parallax of +2% and negative parallax of − 1%). Thus, most 3D content is well within bounds where lab studies suggest the vergence/accommodation conflict should not be causing discomfort. Other sources of discomfort likely also contribute. For example, S3D may present stronger cues to scene structure and motion, which then provide a stronger conflict with vestibular information (Howarth 2011).

Disclosure statement

The study was funded by BSkyB. JR was also funded by Royal Society University Research Fellowship UF041260 during the course of this work. BSkyB also part-funds an industrial CASE Ph.D. student in JR's lab. The authors have no financial interest or benefit arising from applications of this research.

Acknowledgements

We thank Clare Associates (http://www.clareassoc.co.uk/) for implementing the recruitment questionnaire and Suzanne Pinkney for running the Research Volunteer Database. We thank Andrew Baron, Stephanie Clutterbuck, Laura Gray, Yonggang He, Eva Karyka, Ahmad Khundakar, Emma Kirkpatrick, Emma Malcolm, Carmen Martin Ruiz, Danielle McCutcheon, Richard Morris, Bahaa Omran, Preeti Singh and Kun Wang for collecting the lab data. We thank Paul Boyle and the team at C4 Sightcare for hosting the eye tests. We thank Roseanne Robinson for drawing up the standard operating procedures for the orthoptists; Roseanne Robinson, Caroline Penrose, Laura Crawford, Michelle Dent for carrying out the orthoptic examinations; and Paul Boyle, Paul Garvey, Niall Armstrong, Jonathan Bone and Priya Jivani for carrying out the optometric examinations. We thank Suzanne Englebright for excellent administrative support throughout.

Funding

This work was supported by British Sky Broadcasting Limited, Grant Way Isleworth, TW7 5QD, UK (BSkyB, http://corporate.sky.com/). JCAR was supported by Royal Society University Research Fellowship UF041260.

References

Bando, T., A. Iijima, and S. Yano. 2012. "Visual Fatigue Caused by Stereoscopic Images and the Search for the Requirement to Prevent Them: A Review." *Displays* 33 (2): 76–83. doi:10.1016/j.displa.2011.09.001.

Banks, M. S., J. C. A. Read, R. S. Allison, and S. J. Watt. 2012. "Stereoscopy and the Human Visual System." *Motion Imaging* 121 (4): 24–43.

Dodgson, Neil. 2004. "Variation and Extrema of Human Interpupillary Distance." Paper read at proceeding of the SPIE 5291, San Jose, CA.

Emoto, M., Y. Nojiri, and F. Okano. 2004. "Changes in Fusional Vergence Limit and Its Hysteresis After Viewing Stereoscopic TV." *Displays* 25 (2–3): 67–76. doi:10.1016/j.displa.2004.07.001.

Hiruma, N., and T. Fukuda. 1993. "Accommodation Response to Binocular Stereoscopic TV Images and Their Viewing Conditions." *SMPTE Journal* 102 (12): 1137–1140.

Hoffman, D. M., A. R. Girshick, K. Akeley, and M. S. Banks. 2008. "Vergence-Accommodation Conflicts Hinder Visual Performance and Cause Visual Fatigue." *Journal of Vision* 8 (3), 33, 1–30.

Hoffman, D. M., Vasiliy I. Karasev, and M. S. Banks. 2010. "Temporal Presentation Protocols in Stereo Displays: Flicker Visibility, Perceived Motion, and Perceived Depth." *Journal of the Society for Information Display* 19 (3): 271–297.

Howarth, P. A. 2011. "Potential Hazards of Viewing 3-D Stereoscopic Television, Cinema and Computer Games: A Review." *Ophthalmic Physiological Optics* 31 (2): 111–122. doi:10.1111/j.1475-1313.2011.00822.x.

Jampolsky, A. 1955. "Characteristics of Suppression in Strabismus." *Archives of Ophthalmology* 54 (5): 683–696.

Lambooij, M., W. Ijsselsteijn, M. Fortuin, and I. Heynderickx. 2009. "Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review." *Journal of Imaging Science and Technology* 53 (3). doi:10.2352/J.ImagingSci.Technol.2009.53.3.030201.

Nojiri, Y., H. Yamanoue, A. Hanazato, M. Emoto, and F. Okano. 2004. "Visual Comfort/Discomfort and Visual Fatigue Caused by Stereoscopic HDTV Viewing." *Stereoscopic Displays and Virtual Reality Systems XI* 5291: 303–313.

Prasad, M., M. Arora, I. Abu-Arafeh, and G. Harding. 2011. "3D Movies and Risk of Seizures in Patients with Photosensitive Epilepsy." *Seizure*. doi:10.1016/j.seizure.2011.08.012.

Serrano-Pedraza, I., M. P. Clarke, and J. C. Read. 2011. "Single Vision During Ocular Deviation in Intermittent Exotropia." *Ophthalmic Physiological Optics* 31 (1): 45–55. doi:10.1111/j.1475-1313.2010.00805.x.

Shibata, T., J. Kim, D. M. Hoffman, and M. S. Banks. 2011. "The Zone of Comfort: Predicting Visual Discomfort with Stereo Displays." *Journal of Vision* 11 (8). doi:10.1167/11.8.11.

Solimini, A. G., A. Mannocci, D. Di Thiene, and G. La Torre. 2012. "A Survey of Visually Induced Symptoms and Associated Factors in Spectators of Three Dimensional Stereoscopic Movies." *BMC Public Health* 12 (1): 779, doi:10.1186/1471-2458-12-779.

Von Noorden, G., and E. C. Campos. 2002. *Binocular Vision and Ocular Motility: Theory and Management of Strabismus*. 6th ed. St Louis, MO: Mosby-Year Book.

Yang, S. N., T. Schlieski, B. Selmins, S. C. Cooper, R. A. Doherty, P. J. Corriveau, and J. E. Sheedy. 2012. "Stereoscopic Viewing and Reported Perceived Immersion and Symptoms." *Optometry and Vision Science* 89 (7): 1068–1080. doi:10.1097/OPX.0b013e31825da430.

Yang, S. N., and J. E. Sheedy. 2011. "Effects of Vergence and Accommodative Responses on Viewer's Comfort in Viewing 3D Stimuli." *Stereoscopic Displays and Applications XXII*, edited by Andrew J. Woods, Nicolas S. Holliman, and Neil A. Dodgson. Proceedings of the SPIE, Vol. 7863, article id. 78630Q, 13 pp.

Yano, S., and M. Emoto. 2002. "Two Factors in Visual Fatigue Caused by Stereoscopic HDTV Images." *Three-Dimensional TV, Video and Display* 4864: 157–166.

Yano, S., M. Emoto, and T. Mitsuhashi. 2004. "Two Factors in Visual Fatigue Caused by Stereoscopic HDTV Images." *Displays* 25 (4): 141–150. doi:10.1016/j.displa.2004.09.002.

Yano, S., S. Ide, T. Mitsuhashi, and H. Thwaites. 2002. "A Study of Visual Fatigue and Visual Comfort for 3D HDTV/HDTV Images." *Displays* 23 (4): 191–201. doi:10.1016/S0141-9382(02)00038-0.