Viewer experience with stereoscopic 3D television in the home

Jenny C.A. Read *

Institute of Neuroscience, Newcastle University, Newcastle Upon Tyne, NE2 4HH, UK

ABSTRACT

Stereoscopic 3D television (S3D TV) is now available in the home. However, little published information is available on viewer use or experience. In this study, 120 people from 29 households were given a new TV (active or passive stereoscopic 3D, or conventional 2D) and reported on their television viewing and other screen use on a near-daily basis over 8 weeks. People reported enjoying S3D TV and cinema more than TV and cinema in general, but enjoying S3D video games less than video games in general. S3D TV and video games were both associated with an increased, though still low (~10%) level of adverse effects, such as headache and eyestrain. I speculate that this may be because video games present a particularly strong conflict between vergence and accommodative demand.

© 2014 Published by Elsevier B.V.

1. Introduction

In recent years, many studies have examined viewers' experience with stereoscopic 3D (S3D) displays. Most of these have been lab studies [1–6,7]; a few have been observational or survey studies examining reported experience with S3D movies viewed in a cinema [8,9]. To my knowledge, there is so far no published study examining viewer experience in the home. This is a particularly salient omission, given that S3D television in the home is now a reality. Since 2010, several television channels across the world have been broadcasting S3D content.

In this study, 30 households were given either an active S3D, a passive S3D or a conventional 2D high-definition television set. They were asked to fill in daily online reports about their screen use. These reports covered not only television, but also computer use, gaming and cinema. They related both to screen use in general, and S3D displays in particular.

This paper analyses the data provided by 120 participants over a period of 8 weeks each. Its aim is to quantify how ordinary viewers use and experience S3D displays in daily life. This may provide insight into the relatively slow uptake of S3D among the viewing public.

2. Material and methods

In summary, participating households were given a new HD TV to keep at home. They were asked to complete baseline questionnaires before participating, and also to complete brief daily online reports on their screen use. 10 households were given an active S3D TV using shutter glasses; these are referred to as the "A group". 10 "B-group" households were given a passive S3D TV using circularly polarizing glasses. As a control group, a further 10 "C-group" households were given a 2D TV. Details on the three different TV models are given in Table 1.

2.1. Recruitment

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.

Households were recruited by sending out information sheets to 1000 Sky customers in the Newcastle area, explaining the study and inviting them to participate. The information described the study as "Evaluating how watching different types of programmes on different TVs affects the overall enjoyment of TV viewing", but did not explain that we were specifically interested in S3D television, and did not explain that the TV set provided might have S3D functionality. Otherwise, households who were assigned to the C group might have felt that they were not "really" in the study, possibly leading to differences in behaviour which could confound our results. The ethics committee approved this omission.

Households who were interested in participating completed an online recruitment questionnaire, giving details of the members of their household and the television sets currently in the household.

* Tel.: +44 191 208 7559.
E-mail address: jenny.read@ncl.ac.uk

The computer assigned households alternately to the three different groups, in the order they logged on and completed the online recruitment questionnaire. One person in the household was asked to report data for all participating members of the household.

From the information supplied, 30 suitable households were selected. Preference was given to households consisting of 4 or more members who were willing to participate in the study, so as to maximise the number of participants. Households which already owned a S3D TV were excluded, in order to ensure that all households in the study had had the same exposure to home S3D TV, i.e. 8 weeks of ownership. Households where someone worked in the TV, PR or media industries were excluded in order to minimise the chance of breaching study confidentiality. Households where someone had photosensitive epilepsy were excluded on safety grounds, despite the vanishingly small risk of problems [10].

Selected households were then invited to come to Newcastle University, where they had the opportunity to ask questions before giving written consent. On this occasion they also participated in a lab-based TV viewing experiment [11].

Separately, BSkyB engineers made an appointment with the 30 selected households to come in and install a new television set of the type specified by the code assigned to that household (Table 1). In 29/30 cases, this was the code initially assigned by the computer. However, in order to obtain 10 households in each of the three groups, one household was reassigned from "2D control" to "S3D passive" before their TV was installed. One C-group household (H152C) dropped out of the study without giving a reason, shortly after completing the initial lab and eye tests. Thus, 29 households, containing 120 participants, completed the study (see Table 2). A further C-group household (H83C) was excluded from some analysis due to an unusual pattern of results, as described below.

For 8 weeks after receiving their new television set, all participants were asked to log on to a study website every day, using their participant code and a unique password. There, they were asked a

brief series of questions about their screen time that day. Compliance was generally good, although inevitably there were times when households went on holiday or were otherwise unable to complete their questionnaires. Household H112B filled in questionnaires only once a week or so due to technical problems.

In total, we collected 6503 responses from the 120 participants (2305 from A-group households, 2273 from B, 1925 from C). We had aimed to collect 56 responses from every participant (every day for 8 weeks) and our mean was very close to this: 54 responses averaged over all 120 participants who completed the study. We only collected 9 responses from H112B, due to their technical problems accessing the website. Amongst the other households, we collected at least 31 responses from each household.

2.2. Daily questionnaires

For 8 weeks after receiving their new television set, all participants were asked to log on to a study website every day, using their participant code and a unique password. There, they were asked a brief series of questions about their screen time that day. They were asked about 4 different forms of screen time: non-gaming computer use, computer or video games, television and cinema.

The first question was, "How long did you spend using a computer, but not for games?" They chose from 6 options: "0 min, less than 60 min, 1–2 h, 2–3 h, 3–5 h, >5 h". If they selected "0 min", the program proceeded to the next type of screen use. Otherwise, they were asked "How much did you enjoy this? (1 = dreadful, 7 = fantastic)", and "Were there any adverse effects?" (yes/no). If they answered "yes " to adverse effects, the program then offered them a long list of possible adverse effects (blurred vision, difficulty focusing eyes, cramps, double vision, eyestrain, faintness, fatigue, fever, headache, impaired coordination, impaired balance, itching, joint pain, muscle pain, nausea, skin rash, stomach ache, tooth ache, tiredness). They were asked to tick any that applied, and/or to select "other". In the Results section, "headache" is

Table 1

Specifications for the two television sets used in the study. The last 6 rows use information taken from the website of the manufacturer, LG Electronics, at www.lg.com.

	Television set		
TV manufacturer	LG	LG	LG
TV model	47LX6900	47LD920	47LD450
S3D technology	Active shutter (temporally interleaved)	Passive polarized (patterned-retarder)	None (2D TV)
Provided to group	A	B	C
Screen size (inches along the diagonal)	47	47	47
Display type	LED	LCD	LCD
Resolution (width × height, pixels)	1920 × 1080 (Full HD)	1920 × 1080 (Full HD)	1920 × 1080 (Full HD)
Contrast Ratio	8,000,000:1	150,000:1	60,000:1
Audio Output	10 W + 10 W	10 W + 10 W	10 W + 10 W
Dimensions of set (without stand), width × height × depth (mm)	1127 × 692 × 29.3	1173.4 × 723.4 × 100.8	1136 × 698 × 76.5

Table 2

Details of participants in the three different study groups.

Code	A	B	C
Group	Active S3D	Passive S3D	2D control
Model TV set provided	LG 47LX6900 with 4 pairs of AG-S100 active S3D shutter glasses	LG 47LD 920 with 6 pairs of adult and 6 pairs of children's passive S3D glasses	LG 47LD450 with no S3D function
Number of households	10	10	9 (after 1 household dropped out of the study)
Number of participants (male/female)	41 (22/19)	42 (21/21)	37 (19/18)
Birthyear (mean/median, SD)	1984/1990, 17	1981/1987, 16	1984/1989, 15

represented unchanged but the other possible adverse effects are grouped together for reporting as follows. "Eyes" includes blurred vision, difficulty focusing eyes, double vision and/or eyestrain. "Balance" includes nausea, faintness, impaired coordination, impaired balance. "Fatigue" includes fatigue and/or tiredness. Cramps, fever, itching, joint pain, muscle pain, skin rash, stomach ache, tooth ache, other were all classed as "other".

The program then proceeded to the next form of screen use, and asked "How long did you spend playing computer or video games?" The options then unfolded as previously, except that if participants selected more than "0 min", then after asking about enjoyment and adverse effects, the program then proceeded to ask "How long did you spend playing S3D computer or video games?", with the same 6 options. If the participant responded more than "0 min", they were again asked about enjoyment and adverse effects.

Next, participants were asked "How long did you spend watching TV?" and then "How long did you spend watching movies at the cinema?", with the same follow-up questions as for the games section. Note that participants were first asked about screen use, without specifying 2D or S3D, and subsequently, if they reported any screen time, asked specifically about S3D use.

The daily questionnaire was hosted on a secure BSkyB server using code written by web design company Clare Associates Ltd.

3. Results

3.1. Demographics

Many existing studies of S3D displays have recruited primarily young, educated participants, e.g. university students. A strength of this study is that we have recruited a more representative sample of the UK population. Our study population was aged 4–67 years, concentrated on teens and middle-aged, and with roughly equal numbers of males and females. Most people reported viewing S3D content only a few times a year before entering the study. It is also important to confirm that our randomisation procedure did succeed in avoiding substantial differences between the three TV groups. Of course, the sample size (10 households in each group) is not large enough to average out all differences, but the results presented in this section confirm that there are no overwhelming differences and documents the variation that does exist.

3.1.1. Age

Fig. 1 shows the distribution of age for the 3 groups (see also Table 2). Note that to minimise the use of identifying information, we recorded only participants' year of birth. For the purposes of this paper, we defined "age" as the year of study, 2011, minus year of birth, which could be up to 1 year different from age. Because most of our households were families, this is bimodal, with a peak in the under-20 s for children and a peak in the 40–50 s for parents. This is a good range of ages compared to lab-based studies, which often recruit mainly from student populations in their twenties. There was no significant difference in age between the three TV groups ($p = 0.39$, Kruskal–Wallis test).

3.1.2. Gender

We had roughly equal numbers of male and female participants in each group and age-range. The greatest imbalance was in the age-group 9–18, where we had 20 males and 13 females, but this is consistent with equal sampling (such differences occur 30% of the time when 33 people are picked with equal probability of either gender).
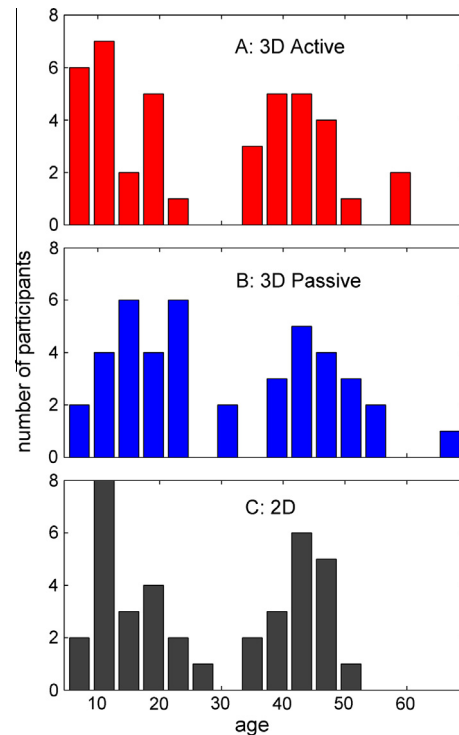

Fig. 1. Distribution of age of participants in the three groups. Bin width is 4 years. Only year of birth was recorded in the study, so "age" refers to the difference between the year of the study, 2011, and the year of birth.
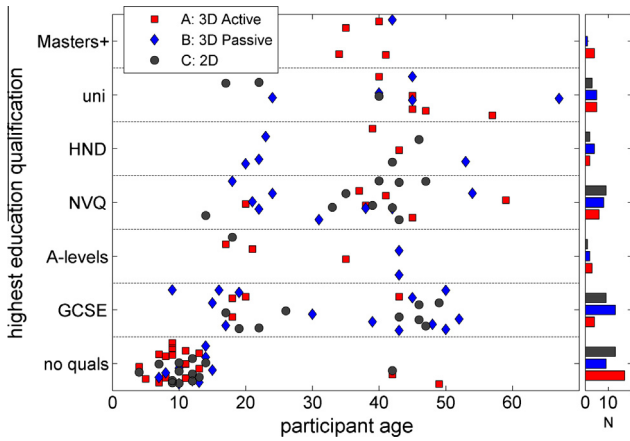
3.1.3. Education

Fig. 2 shows the highest reported education level as a function of age. To avoid points lying on top of one another, points are jittered vertically within each band. The bar-graph on the right shows how many participants in each TV group had which qualification. Out of 76 participants born before 1994, 17 (22%) had a degree-level university qualification. There is internal evidence that these data were not always accurately reported. For example, one participant, aged 9, was reported as having GCSE-level qualifications; another, aged 17, was reported as having a university undergraduate degree, but no GCSEs or A-levels. Importantly, there was no significant difference in education level between the three TV groups ($p = 0.58$, Kruskal–Wallis test).
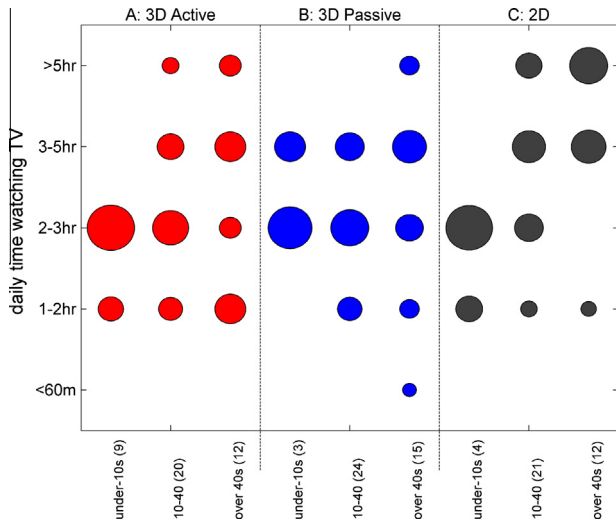
3.1.4. Viewing habits at recruitment

At recruitment, participants were also asked about their typical viewing habits. In this section, we check whether the 3 groups are comparable. Fig. 3 shows these data, grouped by TV group and further sub-divided by age (under-11 s, 11–24 s, 30–40 s, over-40 s). A dot is drawn at each of the possible answers. The area of each dot represents the fraction of participants in that TV and age group who gave that answer. The total area of the dots is the same in each column.

Participants in the A and B groups report very similar viewing habits, with a modal response of 2–3 h viewing TV per day. The C group has slightly more teens and adults who watch more than 5 h TV per day. The median report was "2–3 h" for the A and B groups, and "3–5 h" for the C group; this reflected a significant difference between the three groups ($p = 0.025$, Kruskal–Wallis test) which should be taken into account when interpreting the results.

Clearly, for this study it was also critical to ask how often participants usually view S3D displays. Fig. 4 represents responses to the question "How often do you watch films or TV programmes in 3D?", in the same format as Fig. 3. Most participants view S3D
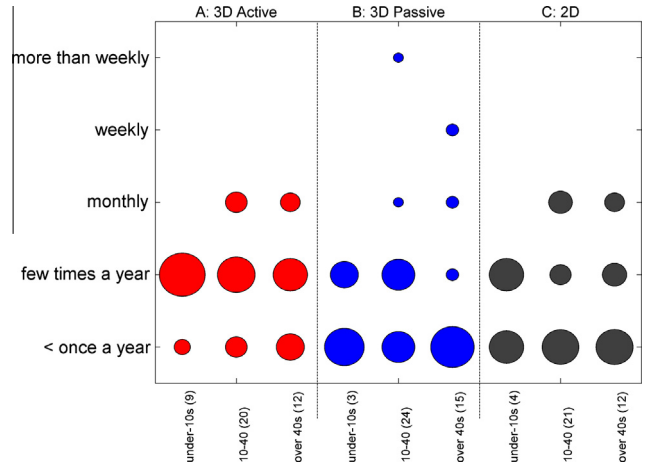
Fig. 2. Education level of participants. The highest education level was defined as a number on a 7-point scale, where 0 = no qualifications and 6 = higher degree. The highest qualification reported for each participant is plotted against their age; color shows TV group. Histograms on the right show total numbers at each level within each TV group.

Fig. 4. Self-reported frequency of viewing S3D content at study onset, in different TV groups and age-groups. Format as in Fig. 3.

Fig. 3. Self-reported time typically spent watching TV at study onset, in different TV groups and age-groups. The total area of the dots in each column is the same. Their relative area indicates the proportion of participants in that column who gave the response indicated on the vertical axis. The number of participants in each column is shown at the bottom.

content only a few times a year, and there is little difference between age-groups. There are significant differences between TV groups ($p = 0.01$, Kruskal–Wallis). The median response was "a few times a year" for the A group, but "less than once a year" for the B and C groups. However, the B group does contain one family where two members report viewing S3D weekly or more (household code H235B). Five households report that at least two members view S3D TV at least once a month. Two of these are in the A group (H132A, H246A), one in the B group (the previously mentioned H235B) and two in the C group (H104C, H134C). Thus our 5 "keen S3D" households are distributed as evenly as possible amongst the 3 groups. Although the inter-group differences should be borne in mind when interpreting the results, our 3 groups are at least comparable in their previous exposure to S3D, with most people experiencing minimal S3D at present. There were no gender differences in the frequency of S3D viewing ($p = 0.49$, Kruskal–Wallis).

In summary, then, our study population is aged 4–67 years, concentrated on teens and middle-aged, and with equal numbers

of males and females. Participants in the three different TV groups were similar, though at study onset, C-group participants watched slightly more TV and A-group participants watched slightly more S3D. The three groups had similar levels of experience with S3D displays, with most people viewing S3D content only a few times in a year.

3.2. Reliability checks

How reliable are these self-reports? Previous work has suggested that subjective reports of viewer experience are reliable [12], and that data collected online is very similar to that collected in the laboratory [13,14]. However, we have already seen reason to suspect some errors in the recruitment questionnaires. There are some internal checks we can do to examine this.

Fig. 5 shows that TV viewing time reported in the daily questionnaires correlated well with the initial reports on the recruitment questionnaire (Spearman correlation $\rho = 0.58$, $p < 10^{-10}$), giving us some confidence in these self-reported measures. One participant (H224B001, labelled in Fig. 5) did report viewing < 60 min TV initially, but averaged 3–5 h on the daily questionnaires. Of course we do not know why this might be, but we note that this participant was an adult and was the person who filled in his household's recruitment form. Thus he reported his own TV viewing, reducing the scope for mistakes. He was in a "B" group household which received a S3D TV, but S3D was not the reason for his increased TV viewing, as he reported viewing S3D on only two out of the 64 days on which he filled out questionnaires, both within four days of receiving the TV.

The daily questionnaire first asked participants "How long did you spend watching TV?" and then "How long did you spend watching S3D TV?" Clearly the answer to the latter should not exceed the former. Reassuringly, this occurred only 4 times in the 6501 responses collected.

One household, H83C, did often report spending more time on S3D cinema or gaming than on all cinema or gaming, suggesting either an error in reporting or a misunderstanding of what was being asked. This household was unusual in other ways. First, it was a household of four adults, a woman in her mid-40 s, a woman aged 19 and two men in their 20 s. None reported any qualifications above GCSE and none were in full-time work (the men were unemployed, the younger woman was a student and the older worked part-time). All reported very large amounts of all sorts of screen time, on average 10 h a day, including exceptionally large amounts of S3D screen time. For example, the woman in her 40 s
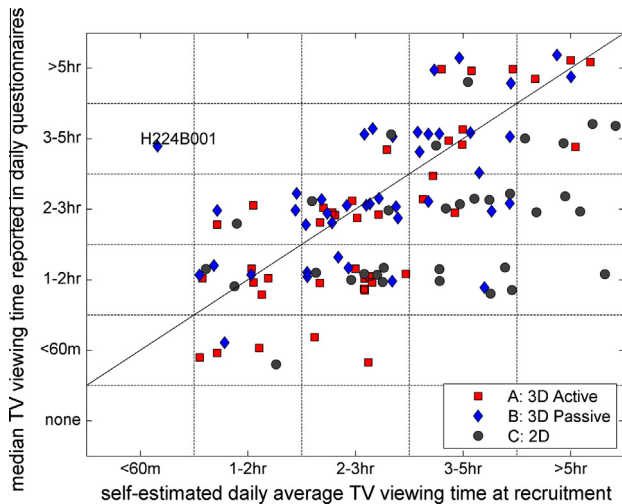
Fig. 5. Scatterplot relating each participant's reported average daily TV viewing at recruitment, as reported in the initial recruitment questionnaire, to the median TV viewing time from their (ideally) daily reports throughout the study. The diagonal line marks the identity. So that points do not overlap, data points are given a random jitter before plotting; this is not large enough to move them out of their grid square. The correlation was of course calculated on the raw data without jitter. Participant H224B001, who is discussed in the text, is marked.
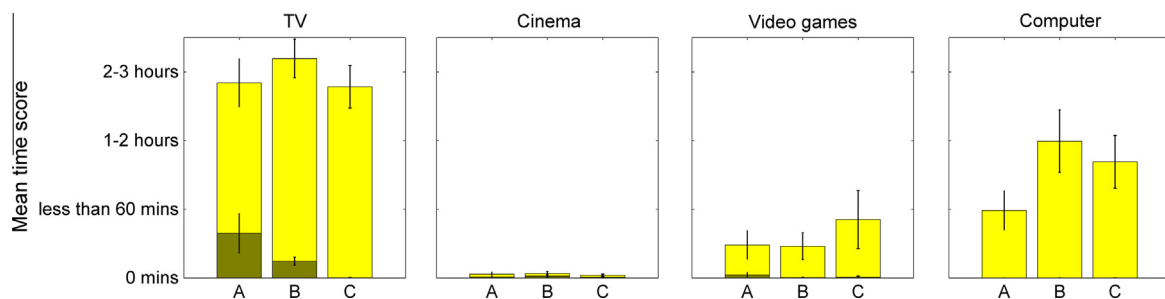
3.3. Reported screen time

Fig. 6 shows the average daily screen time reported by participants in the three different groups, for four different types of screen use. TV viewing (Fig. 6A) is similar across the 3 groups. The C group reported slightly more viewing time at recruitment (still true after the removal of household H83C), but this is not reflected in their answers to the daily questionnaires.

The dark bars in Fig. 6 show the time for S3D. Since the S3D enthusiasts in household H83C have now been removed from the analysis, only 4 out of the remaining 33 C-group participants reported watching any S3D TV during the study, and only on one occasion each. In contrast, 39 out of the 41 A-group participants watched at least some S3D TV, and 41 out of the 42 B-group participants. So, when people are given the opportunity to watch S3D TV, almost all of them choose to do so.

Fig. 7 shows how the probability of watching S3D TV decayed over the course of the study. When filing the first few reports, shortly after having received their new S3D TV, just under half of participants in the S3D groups were viewing S3D TV each day. After 3 weeks or so, as the novelty presumably wore off, this had decreased to around 25% in the A group and 10% in the B group. However, in both cases it then remained relatively steady at this level until the end of the study.

often reported having spent 1–2 h that day playing S3D video games, often as well as spending 1–2 h the same day watching S3D movies in the cinema. One of the men regularly reported having spent 3–5 h that day watching S3D movies in the cinema. For all for members of this household, the median time reported watching S3D movies in the cinema each day was at least 1–2 h. Given the relatively small number of S3D movies released, this implies that they must have gone to see the same film multiple times. This household was in the 2D control group, meaning that they were given a 2D TV for the study. At recruitment, they reported having 4 HD TVs in their home and no S3D TV, but during the study, they reported spending large amounts of time watching S3D TV. The median S3D TV was "1–2 h" for the older woman and at least "2–3 h" for the others. Of course, this household may have bought a S3D TV independently during the course of the study, or they may have watched S3D TV outside the home e.g. in sports bars. However, given the unusual viewing patterns reported, we felt that these reports might simply be unreliable. For this reason, household H83C is excluded from the subsequent analysis.

3.4. Reported enjoyment of 2D and S3D

As well as reporting how much time they spent in various activities, participants also reported how much they enjoyed it, on a seven-point Likert scale. Fig. 8 shows the distribution of these enjoyment reports for all screen time (top row), and for S3D screen time (bottom row). Recall that participants are not initially asked about 2D TV and S3D TV. Rather, they are asked about "TV", and if they report watching some TV, they are then asked how much was S3D. Thus, if they have watched both 2D and S3D TV, the figures in the top row should relate to both types. The legends indicate the number of responses (R) these distributions are calculated from, and the number of participants (P) making those responses. For example, only 24 participants ever reported viewing S3D cinema, on a total of 38 occasions, so the distribution in Fig. 8e is calculated from just 38 responses collected from 24 individuals. In contrast, the distribution in Fig. 8e is calculated from 5856 responses collected from 116 individuals, making it a much more reliable estimate of TV enjoyment. The vertical lines in Fig. 8 mark the median (solid) and quartiles (dotted).

Fig. 6. Screen time reported on the daily questionnaires. Participants were asked to choose one of 6 options to describe the amount of time spent on different types of screen time. To obtain the data in this figure, each option is assigned a number from 0 to 5 and these numbers are then averaged across the full set of responses completed for each individual participant, resulting in an "average time score" for each participant. This is of course not literally the number of hours, firstly because a score 5 = " > 5h" does not necessarily represent five times the viewing time of 1 = " < 60 min", and secondly because no account was taken of the time since the last response, even though participants were asked to report for the entire time-period since their last questionnaire. After having obtained this average time score for each participant, these were then further averaged across all participants in the 3 different TV groups, A B and C, to obtain the bars shown here. Error-bars are standard errors of the mean across participants. The portion of the bars shaded dark show mean time reported watching S3D content. For "Computer", where the question was "How long did you spend using a computer, but not for games?", no S3D option was available. We chose to report mean time as described above rather than median, because median time was zero for cinema (all participants) and for S3D TV (all B group participants).
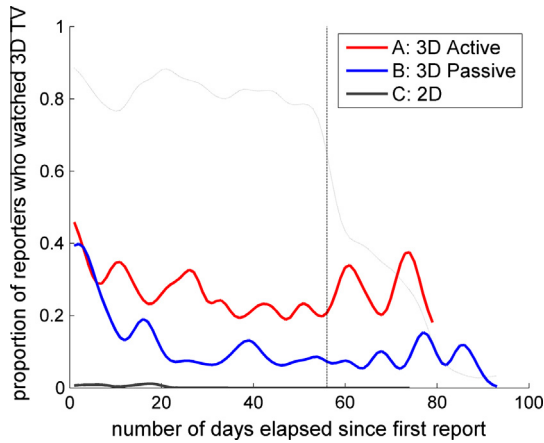
Fig. 7. Frequency of S3D TV viewing during the study. The thick colored lines show, out of all participants in each group who submitted a report on that day, the proportion who reported watching S3D TV, as a function of the number of days since each participant's first report. For comparison, the dashed line shows the proportion of all participants who submitted a report on each day. Data were smoothed with a moving Gaussian window with a standard deviation of 4 days. The vertical dashed line marks the 8 weeks for which participants were asked to submit daily reports. Around 80% complied; a minority continued submitting reports beyond this.

Our data suggest that people get particular enjoyment from watching S3D TV. Comparing Fig. 8a vs. d, and Fig. 8b vs. e, people are more likely to give the strongest enjoyment ratings to S3D content rather than 2D. Bootstrap resampling indicated that this difference was significant for TV ($p = 0.01$), though not cinema ($p = 0.1$). To calculate this for TV, we take Fig. 8a as our estimate of the distribution of enjoyment for TV, under the null hypothesis that 2D and S3D TV is equally enjoyable. We generated resampled data-sets by picking 990 values, randomly with replacement, from this set of 6145 ratings. We calculated the median of this resampled data-set, and repeated this 10,000 times. The resampled

median equalled the median for S3D viewing only 1% of the time, and never exceeded it.

Fig. 8d also shows that users are more likely to give ratings of "poor" and "not very good" to S3D content. This raises the possibility that as well as producing higher enjoyment on average, S3D also evokes a wider range of responses both positive and negative. However, our bootstrap resampling did not indicate a significant difference in the interquartile range for S3D vs. unspecified content, thus not supporting this possibility.

Interestingly, S3D video games were experienced as less enjoyable than video games in general, and this difference in medians was also judged significant under bootstrap resampling ($p = 0.02$).
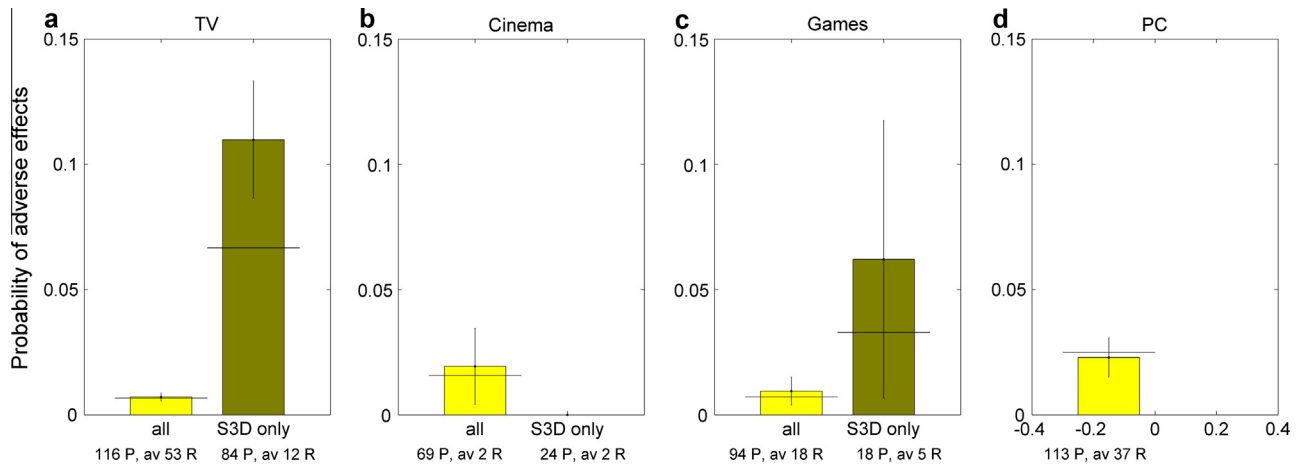
3.5. Adverse effects

Participants were also asked to report whether they experienced any adverse effects after the screen time they reported. For each participant, we calculated the proportion of times on which they reported adverse effects after each type of screen use. Fig. 9 shows this mean probability averaged across participants. Bright bars show the rate of adverse effects reported after unspecified "TV", "cinema" etc; dark bars show it specifically after "S3D TV", "S3D cinema".

The bars in Fig. 9 show the mean probability of reported adverse effects for each participant. An alternative approach would be to calculate the rate of adverse responses after pooling responses over participants, and this is shown in Fig. 9 by horizontal lines. To illustrate how these two approaches could potentially give very different results, suppose that we only had 2 participants. Suppose one watched S3D TV only once, and reported experiencing an adverse effect; their "rate of adverse effects" would therefore be estimated, with low confidence, at 100%. Suppose the other watched S3D TV on 40 occasions, and reported no adverse effects; their "rate of adverse effects" would therefore be estimated, with rather more confidence, as 0%. The mean rate of adverse effects would be 50%. This figure would be shown by the height of the bars in Fig. 9. The disadvantage of this calculation is that it gives equal
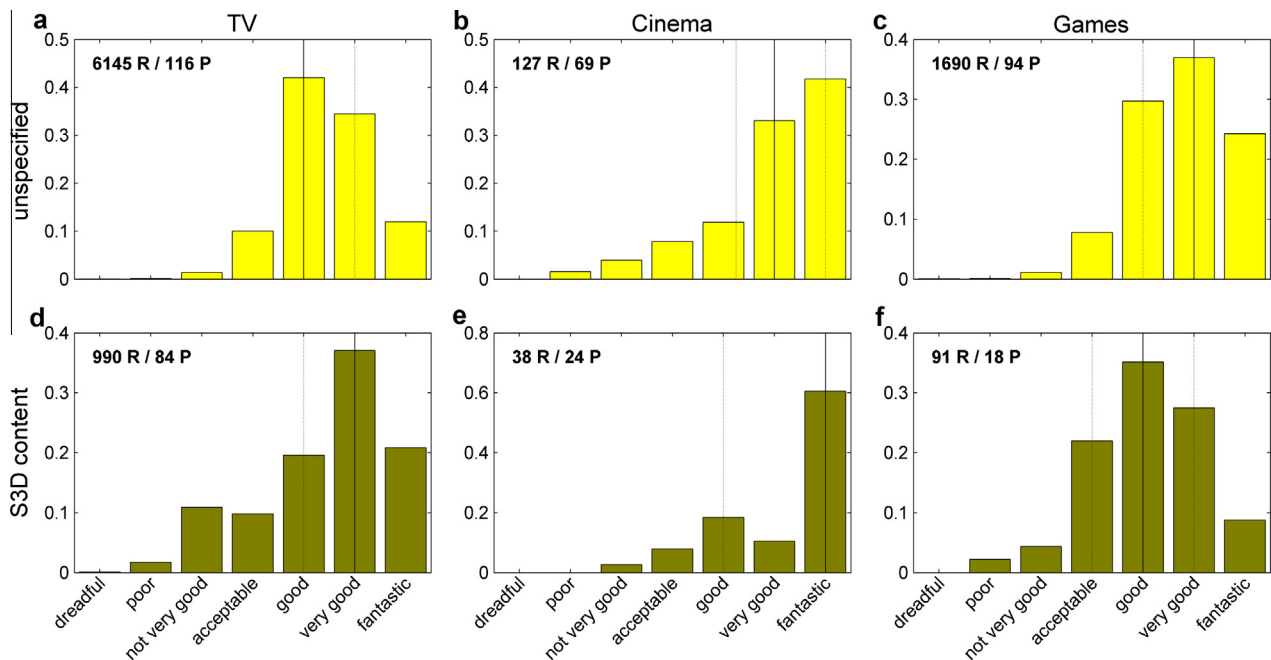
Fig. 8. Enjoyment of screen time. We examined enjoyment ratings reported from participants for each activity, and calculated the distribution of enjoyment ratings, i.e. how often each activity was given which enjoyment rating. The numbers "6145 R/116 P" etc mean that the specified distribution was calculated from "6145 responses collected from 116 participants" over the course of the study. Vertical solid lines mark the median; dotted lines show the interquartile range.

Fig. 9. Probability that adverse effects are reported after viewing, for each participant. To obtain this figure, for each participant, we take all the responses where they reported a particular type of screen use, and calculate their individual probability of reporting adverse effects after that screen type. We then average this over all participants who reported that kind of screen use. The numbers below each bar show the number of participants who went into the calculation, and the mean number of responses per participant. For example, only 24 participants reported viewing S3D cinema, on a total of 38 occasions, and none of them ever reported experiencing adverse effects. Their mean rate of adverse effects is therefore 0, calculated from an average of under 2 responses per participant. The error bar shows the standard error on the mean rate of adverse effects. "PC" indicates "screen use at a computer, not for games". We did not ask about S3D for this case.

weight to the two estimates, even though our confidence is much higher for the second. Alternatively, we could say that we have 41 reports of S3D TV viewing, and only 1 report of an adverse effect, and calculate the rate of adverse effects as 1/41 = 2.4%. This would be shown by horizontal lines in Fig. 9. The disadvantage of this calculation is that it gives most weight to participants who made the most responses. If, as seems likely, participants who are least likely to experience adverse effects are most likely to choose to view S3D, this could underestimate the incidence of adverse effects in the general population. In fact, as Fig. 9 shows, the two calculations give rather similar results in our data-set, making the choice less critical. Unsurprisingly, the rate is slightly lower after pooling (horizontal lines are generally lower than bars). This is what we would expect if different people are more or less susceptible to adverse effects after screen use, and people who are less susceptible are more likely to engage in that kind of screen use.

The highest rate of adverse effects was associated with S3D TV, where it was around 10%. This is much lower than the rates reported in the lab, where we have recently reported rates in excess of 20% [11]. We estimated there that around 14% of the population may experience adverse effects which are really due to S3D, rather than to other factors such as the glasses or negative expectations. Here, the rate is lower still, at around 10% over all participants. There are a number of possible reasons for this: the artificial environment of lab studies may make participants more likely to report adverse effects, or people may be more likely to report adverse effects when they are asked about them immediately after viewing than when they fill in a report some time later. It may also indicate that people "acclimatise" to S3D. Interestingly, adverse effects were never reported after watching S3D cinema. Adverse effects after S3D TV were also not reported by the 4 participants in the 2D group who reported viewing S3D TV, presumably after going out of their way to watch it. This is potentially an important finding, given that most scientific work on adverse effects has been carried out in the lab, often in much less comfortable conditions than in [11] (see Fig. 10).

Our results agree with previous work suggesting that the most common adverse effects reported after viewing S3D content are
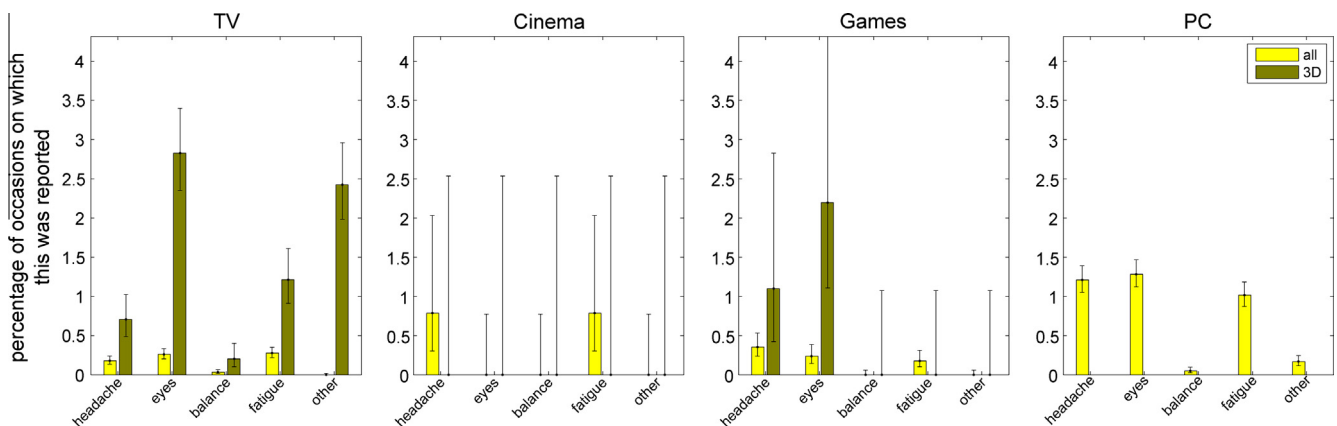
Fig. 10. For this figure, adverse responses were grouped into 5 broad classes, as described in the Methods. For each panel, we collected all responses on which the participant reported spending time on the type of screen use under consideration. The bars show on what percentage of responses the participant reported one or more of the adverse effects within each group. Bright bars are where 2D/S3D was unspecified; dark bars are for S3D. We did not ask about S3D for PC ("computer, non-game") screen use. The overall rate of adverse effects in S3D TV was 6.7% (990 questionnaires filed on which people reported viewing some S3D TV; adverse effects reported on 66 of these). The bars in panel (a) sum to 7.4%, slightly more than this, because of a small number of questionnaires in which participants reported adverse effects from two different groups. Error-bars show 68% confidence intervals assuming simple binomial statistics. The confidence intervals on TV are particularly small because participants viewed TV so often; conversely confidence intervals on cinema are particularly high.

some form of eyestrain. Problems with the eyes were reported 2–3% of the time after viewing S3D TV or games. Eye problems were also reported 1% of the time after using a computer for non-game, presumably 2D, purposes.

In agreement with our previous paper [11], we found that females were more likely to report adverse effects after viewing S3D than males. On average across days where they reported viewing S3D TV, males reported adverse effects 7% of the time. For females, this rose to 15%.

4. Discussion

Previous studies of S3D viewer experience have either been in the laboratory, where results are highly reliable but may have little ecological validity [1,2,5–7,11,15–17], or survey-based studies asking people to recall their experience over long periods of time, where the results may be inaccurate [9,18,19]. This study is unusual in that it examines viewer experience in participants' own homes, but does so by asking participants to report on their screen use and experience on a daily basis, while the experience is still fresh in their minds. Thus, it strikes a balance between reliability and relevance to normal viewing conditions. Our methodology is closest to that of [8], but examines experience in the home with TV, video games and computers as well as in cinemas.

The main problem with the type of data obtained in this study is that it depends on self-report by untrained participants. We collected large amounts of data by asking participants to complete on-line questionnaires: a detailed questionnaire at recruitment, including demographic information, and short daily reports on screen use. We have no way of knowing how accurate these reports are; as we discuss below, internal evidence suggests that some participants misinterpreted some questions. Not all participants did report screen use daily, as requested. Finally, this paper examines only participants' subjective experience, rather than quantitative measures such as pupillary changes [5] or vergence movements [20]. Thus, this study complements rather than replaces laboratory studies.

The trade-off is that our data have high ecological validity. Our results are more likely than those of lab studies to apply to people's actual experience of S3D displays in normal life. Critically, in our study participants watched S3D content of their own choosing, at a time of their choosing. They viewed it in real-life conditions, in their own homes or in other locations such as pubs or cinemas. The long timescale of the study allowed participants to have become used to viewing S3D displays, and gave time for any initial
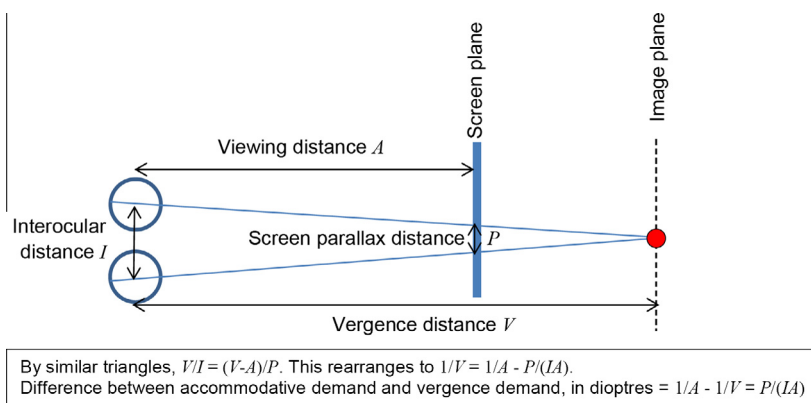
novelty to have worn off. Previous studies have also tried to assess viewers' experience of S3D outside the lab, but the design of the study makes our data of higher quality than studies which, for example, asked participants to report remembered experiences of S3D viewed many months ago [9]. In our study, participants reported on content viewed that day. Previous non-lab studies have asked participants about aspects of their experience such as visual comfort or sense of immersion, but only for S3D content [9,18,19], meaning that it is impossible to draw conclusions about how S3D differs from 2D. Crucially, we asked participants about both general and S3D screen use. Together, these aspects of study design enable us to draw robust conclusions about how S3D affects viewers' experience in naturalistic settings.

Even when they own a S3D TV, people spent a relatively small proportion of time viewing S3D content, typically less than an hour a day out of 2–3 h total TV viewing. On any given day, around 1 in every 5 participants in the two S3D groups would report viewing any S3D content. Due to the nature of the study, we do not have any insight into what content participants chose to view.

Our study reveals that different S3D displays produce different user experience. Our participants reported finding S3D TV and cinema more enjoyable than TV and cinema in general, whereas S3D games were less enjoyable than games in general. S3D TV and games also produced significantly higher rates of adverse effects than TV and games in general, whereas none of our participants reported any adverse effects with S3D cinema (admittedly from a small sample).

One possible reason for this difference could be that these three forms of S3D produce different amounts of conflict between vergence and accommodation. Suppose an object is presented at screen parallax P (in metres) on a screen viewed at a distance A (Fig. 11). The accommodation required to focus on the screen is $1/A$ dioptres. However, if viewers converge on the object's simulated distance V, then the natural accommodation for the vergence distance is $1/V$. The difference between the natural and required accommodation is $(1/V-1/A)$, and several studies have shown that the larger this difference, the more uncomfortable and fatiguing the viewing [1,2,4,6,21]. Fig. 11 shows that the difference $(1/V-1/A)$ equates to $P/(IA)$, where I is the interocular distance. Thus, the amount of discomfort is predicted to scale with the ratio between the screen parallax, in metres, and the viewing distance.

Of course, screen parallax also varies with viewing distance. If the same S3D content were presented on different-sized displays, simply rescaled to fit the display, then screen parallax P would be proportional to the screen height H. If viewers then chose to sit at a constant multiple of screen height, as is often assumed as

By similar triangles, $V/I = (V-A)/P$. This rearranges to $1/V = 1/A - P/(IA)$.
Difference between accommodative demand and vergence demand, in dioptres $= 1/A - 1/V = P/(IA)$

Fig. 11. Quantifying the accommodation/vergence conflict. A virtual object (disk) is presented with left and right images a distance P apart on a screen viewed at a distance A. The viewer fixates the virtual object, at a vergence distance V. The vergence/accommodation conflict is then P/IA. The calculation is shown for positive screen parallax (object behind screen); the same result holds for negative screen parallax, with a change in sign since the accommodative demand is then less than the vergence.

a rule of thumb, then both P and A would be proportional to H. The vergence/accommodation conflict would then be constant, predicting that all the displays would be equally comfortable.

In fact, people watch different content on different displays, and stereographers take great care to avoid divergent images where parallax exceeds the interocular distance, $P > I$. If we assume that the parallax P presented on any display never exceeds the interocular distance I, then the maximum vergence/accommodation conflict is simply the reciprocal of viewing distance, $1/A$. Human depth of field is around 0.2 dioptres, so under these assumptions, the vergence/accommodation conflict ceases to be noticeable at viewing distances longer than about 5 m. This may explain why so few adverse effects were associated with S3D cinema, and why S3D games were experienced as less enjoyable and more uncomfortable.

Our study is particularly valuable when compared with our own results in the lab [11]. Comparing these two studies, people report fewer adverse effects when viewing S3D TV in the comfort of their own home at a time of their own choosing than they do when studied in a lab setting. Thus, lab studies by their nature may tend to over-estimate the adverse effects associated with S3D. A second key finding is that people report greater enjoyment of S3D TV than of general TV. The most obvious difference between this and previous work is that ours is an observational study, in which people chose whether and when to watch S3D content. The picture that is emerging from this and other studies is that some people do not like S3D and a minority even report adverse effects such as headache. However, most people do not experience adverse effects with S3D. Given the opportunity, many people choose to watch S3D content at least once a week, and they tend to report that it enhances their viewing experience.

Acknowledgments

References

[1] M. Hoffman, A.R. Girshick, K. Akeley, M.S. Banks, Vergence–accommodation conflicts hinder visual performance and cause visual fatigue, J. Vis. 8 (2008). 33 31–30.

[2] T. Shibata, J. Kim, D.M. Hoffman, M.S. Banks, The zone of comfort: predicting visual discomfort with stereo displays, J. Vis. 11 (2011).

[3] P.A. Howarth, Potential hazards of viewing 3-D stereoscopic television, cinema and computer games: a review, Ophthalmic Physiol. Opt. 31 (2011) 111–122.

[4] K.J. MacKenzie, D.M. Hoffman, S.J. Watt, Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control, J. Vis. 10 (2010) 22.

[5] H. Oyamada, A. Iijima, A. Tanaka, K. Ukai, H. Toda, N. Sugita, M. Yoshizawa, T. Bando, A pilot study on pupillary and cardiovascular changes induced by stereoscopic video movies, J Neuroeng. Rehabilit. 4 (2007) 37.

[6] S. Yano, S. Ide, T. Mitsuhashi, H. Thwaites, A study of visual fatigue and visual comfort for 3D HDTV/HDTV images, Displays 23 (2002) 191–201.

[7] S. Oliveira, J. Jorge, J.M. Gonzalez-Meijome, Dynamic accommodative response to different visual stimuli (2D vs. 3D) while watching television and while playing Nintendo 3DS Console, Ophthal. Physl. Opt. 32 (2012) 383–389.

[8] A.G. Solimini, Are there side effects to watching 3D movies? A prospective crossover observational study on visually induced motion sickness, Plos One 8 (2013) e56160.

[9] A.G. Solimini, A. Mannocci, D. Di Thiene, G. La Torre, A survey of visually induced symptoms and associated factors in spectators of three dimensional stereoscopic movies, BMC public health 12 (2012) 779.

[10] M. Prasad, M. Arora, I. Abu-Arafeh, G. Harding, 3D movies and risk of seizures in patients with photosensitive epilepsy, Seizure (2011).

[11] J.C. Read, I. Bohr, User experience while viewing stereoscopic 3D television, Ergonomics 57 (2012) 1140–1153.

[12] P.A. Howarth, H.O. Istance, The validity of subjective reports of visual discomfort, Human Factors 28 (1986) 347–351.

[13] J. Sprouse, A validation of Amazon mechanical turk for the collection of acceptability judgments in linguistic theory, Behav. Res. Meth. 43 (2011) 155–167.

[14] J.K. Goodman, C.E. Cryder, A. Cheema, Data collection in a flat world: the strengths and weaknesses of mechanical turk samples, J. Behav. Decis. Making 26 (2013) 213–224.

[15] Y. Nojiri, H. Yamanoue, A. Hanazato, M. Emoto, F. Okano, Visual comfort/discomfort and visual fatigue caused by stereoscopic HDTV viewing, Stereosc. Displays Virtual Real. Syst. Xi 5291 (2004) 303–313.

[16] S. Yano, M. Emoto, Two factors in visual fatigue caused by stereoscopic HDTV images, P. Soc. Photo-Opt. Ins. 4864 (2002) 157–166.

[17] S. Yano, M. Emoto, T. Mitsuhashi, Two factors in visual fatigue caused by stereoscopic HDTV images, Displays 25 (2004) 141–150.

[18] M. Polonen, M. Salmimaa, V. Aaltonen, J. Hakkinen, J. Takatalo, Subjective measures of presence and discomfort in viewers of color-separation-based stereoscopic cinema, J. Soc. Inf. Display 17 (2009) 459–466.

[19] M. Polonen, M. Salmimaa, J. Takatalo, J. Hakkinen, Subjective experiences of watching stereoscopic Avatar and U2 3D in a cinema, J. Electron. Imaging 21 (2012).

[20] A. Iijima, S. Komagata, T. Kiryu, T. Bando, I. Hasegawa, Vergence eye movements signifying 3D depth perception from 2D movies, Displays 33 (2012) 91–97.

[21] Y. Okada, K. Ukai, J.S. Wolffsohn, B. Gilmartin, A. Iijima, T. Bando, Target spatial frequency determines the response to conflicting defocus- and convergence-driven accommodative stimuli, Vision. Res. 46 (2006) 475–484.

Jenny Read works on all aspects of stereoscopic (S3D) vision, including computational models of disparity encoding, human psychophysics, clinical applications in ophthalmology, S3D displays, and S3D vision in an invertebrate system, the praying mantis. After a doctorate in theoretical astrophysics, Read moved into neuroscience career with a Welcome Trust Training Fellowship in mathematical biology. In 2001, she moved to the U.S. to work at the National Eye Institute at the National Institutes of Health in Maryland. She moved to Newcastle University in the northeast of England on a Royal Society University Research Fellowship, and is now a Reader in Vision Science at Newcastle's Institute of Neuroscience.