RESEARCH ARTICLE

User-experience in design and use: enhancing the experience of media content with programmable surround lighting

Paul van Schaik¹ · Matthew Unwin¹ · David Eves² · Melissa Lara³ · Alfred Mensah¹

Received: 15 December 2016 / Published online: 15 November 2017
© Springer International Publishing AG 2017

Abstract Programmable surround lighting has the potential to enhance user-experience of media content, but there is a lack of research demonstrating this. Building on existing work in user-experience and Kurosu's framework for user-experience design and evaluation, we developed a method for testing people's experience of video content with added programmable controlled surround lighting. We employed simple video content to evoke a response of positive or negative affect. Using a repeated-measures design ($N = 33$), we manipulated the colour of surround lighting to enhance the affect response (yellow and green for positive affect; red and purple for negative affect) and then tested the benefits of added surround lighting. Yellow surround colour enhanced positive affect in response to video content and red surround colour enhanced negative affect. There was evidence of assimilation effects as a result of alternating coloured (e.g., yellow) and white surround lighting on affect. This work has implications for the choice of surround lighting colour to enhance user-experience, research design and substantive future research.

Keywords User-experience evaluation · Magnitude-based inference · Programmable surround lighting · Media content · Data analysis · Comparative testing

Introduction

Over the past 20 years, user-experience (UX) has become an influential concept [22]. UX and its enhancement have been studied predominantly by targeting the design of products [24]. Here, we take a novel approach to enhancing UX by way of programmable surround lighting. In particular, we study how this lighting can enhance the affective response to media content. As a basis for our study, we first review work on affect and UX, a framework for UX design and evaluation, the enhancement of UX by way of programmable surround lighting and UX evaluation.

Affect and UX

Emotion, and more generally affect, plays a central role in UX. For example, in Thüring and Mahlke's [44] Components of User Experience (CUE) model emotional reaction is a major antecedent of a user's evaluation of product quality. Contemporary research examines emotion from a dimensional perspective [38, 39]. The idea is that at all times people have conscious access to affect as a simple feeling (core affect); this is a mix of two dimensions: valence (pleasure–displeasure) and arousal (activation–deactivation). Affective quality is the ability of things (objects) in the environment to cause a change in affect. An emotion can be seen as a particular combination of components. This can include a person's affect, the perception of affective quality that the person links to a particular thing in the environment and the person's evaluation of

Electronic supplementary material The online version of this article (doi:[10.1007/s41233-017-0013-6](https://doi.org/10.1007/s41233-017-0013-6)) contains supplementary material, which is available to authorized users.

✉ Paul van Schaik
p.van-schaik@tees.ac.uk

¹ Teesside University, Middlesbrough, UK

² amBX UK Ltd, Redhill, Surrey, UK

³ McLaren Technology Group, Woking, UK

what is felt. A distinction needs to be made between the typical form (prototype) of a particular emotion and the particular occurrence of the emotion (emotional episode). The latter resembles the typical form well enough to be experienced as an instance of the prototype. According to Russell, many or most instances of affect may not be emotions, even though people can still report their affect. Moreover, a person may not experience a potential emotion because the components of the experience that could give rise to an emotional episode do not resemble the person's prototype of a particular emotion well enough. It is therefore important to measure affect in UX research. Affect may be more sensitive to any manipulation than emotion is, as a person may not recognise a particular emotion in what they experience.

As an illustration Russell's conceptualisation of affect consider the following example. Mary's emotion of fear can be seen as a combination of Mary's affect, her perception of the affective quality of a fear-inducing spider and her evaluation of what she feels (fear). In contrast to Mary's prototype of fear, stands an emotional episode of Mary's fear in response to seeing a particular spider in her virtual food cupboard when playing a life simulation video game. Mary would not experience fear if the components of the experience that could give rise to the emotional episode do not resemble her prototype of fear well enough.

In our study, we wanted to examine both positive affect and negative affect rather than both valence and arousal. We considered positive affect and negative affect separately because these have been found to be independent factors in factor-analytic studies [35, 46]. We interpret positive affect in response to stimuli with positive affective/pleasure quality as valence (pleasure). We also interpret negative affect in response to stimuli with negative affective/displeasure as valence (displeasure). Therefore, although core affect is a possible framework for studying affect, we decided not to use this in full.

Framework for UX design and evaluation

Kurosu [31] presents a useful two-dimensional framework for the design and evaluation of UX. The two dimensions are quality (objective or subjective) and activity (design or use). According to this framework we can design objective quality (e.g., objective usability ("the capability of an artifact to be used with ease", [31]) and functionality ("the ability of an artifact to support the user functionally to achieve the goal", [31])). We can also design subjective quality (attractiveness: appeal for needs (e.g., stimulation: the capability to meet the human need for stimulation; [42]) and appeal for apperception (a user's projection of their mental state onto an external stimulus and their interpretation based on their thoughts and feelings; [31])).

We can then evaluate objective quality in use (e.g., achieved usability as effectiveness and efficiency [29]). We can also evaluate subjective quality in use, in other words UX (e.g., affect). In this study, as an example of subjective quality in design, we examine affective quality: appeal for an affect response to an artefact. As an example of subjective quality in use, we examine affect in response to an artefact. In terms of Kurosu's framework, the current study is directed at designing subjective quality (affective quality) to enhance subjective quality in use (UX of media content in terms of affect).

Programmable surround lighting to enhance UX

Traditionally, surround lighting has played an important role in media experiences, for example in the lighting of live performances of stage plays and music by enhancing the audience's affect response, and in other contexts such as work by making sure that surround lighting matches work requirements. With the advent of programmable surround lighting, the possibilities to enhance objective quality and subjective quality in use through appropriate objective quality and subjective quality in design [31] as well as the flexibility and efficiency of designing enhancements have increased substantially. These enhancements are made possible by developments in hardware and software.

Hardware and software

Light-emitting diode (LED) technology has been available since the early 1960s. LEDs are two-lead semiconductor light sources and have various advantages over incandescent light sources (e.g., light bulbs). These include reduced energy consumption, increased lifetime, greater physical robustness, reduced size and increased speed of switching. LEDs are not only used for lighting, but have also been used to build new types of display.

Computer-controlled LED lighting allows for the flexible control of many types of environment, such as multi-use hospitality and entertainment venues, for example Manchester Arena or 8 Northumberland Avenue [1, 2] or in augmenting digital media [11]. However, a challenge to achieving its potential of enhancing UX and other outcomes has been a lack of powerful and flexible software and control systems. The amBX Light-Scene Engine system now provides this capability [1]. The research presented here has used this innovative solution, documented by various patents [14], covering physical mark-up language, and Eves and Cole [15], covering dynamic mark-up language). In particular, the solution offers (among other capabilities) full ambient, dynamic and interactive lighting control, direct audio-and-video-to-light control effects and

interactive effects from a wide range of triggers, sensors and external systems [16]. These can facilitate the integration of media experiences in design and use, and make media experiences more interactive.

Enhancement of objective and subjective quality

Programmable surround-lighting technology offers opportunities for objective and subjective quality in design, and for quickly changing lighting depending on specific requirements. For instance, research has demonstrated the potential of light to influence people's affect. In particular, in depressive patients exposure to bright light prevented a relapse to depressive symptoms after one night of partial sleep deprivation [37]. By demonstrating that lighting can have a positive effect on an individual's affect, this existing work provides a basis for the current research into the enhancement of UX.

Programmable surround lighting may, for instance, be effective in adaptively maintaining appropriate lighting for the purpose of work in an internal space, in response to external lighting. In the context of the current study, Garnett [18] provides the following guidance on how hue (the wavelength of a colour, which is identified with names such as blue or green) may be manipulated to influence affect response. Yellow is associated with uplifting and illuminating emotions (due to it being the lightest hue of the spectrum), which offers positive feelings of hope, happiness, cheerfulness and fun. Green is associated with growth, the colour of spring, renewal and life, which suggests it is an emotionally positive colour. Therefore, both yellow and green may be effective in enhancing positive affective quality of media content. Red is an energising colour that motivates arousal and adrenaline. It can stimulate love and sex on the positive side, and revenge and anger on the negative. Red, in society, is also associated to identify danger through warning signs and signals, and this has an evolutionary basis associated with blood and the feeling of danger it evokes. Purple can be effective by combining the richness of red and coolness of blue; the resulting colour demands respect, suggesting it could be used to enhance negative emotions like fear and depression. Therefore, both red and purple may be effective in enhancing negative affective quality of media content.

UX evaluation

In the UX literature there has been a lack of research evaluating UX of media content in relation to the design of lighting enhancements to support people's experience of media content. However, such research is becoming increasingly important as new technology supports the design of programmable surround lighting enhancements.

Therefore, their effect on UX in response to media content needs to be tested. In response to this need, the current study presents a method for evaluating the effect of programmable surround lighting to enhance UX of media content.

UX evaluation of media experiences enhanced by surround lighting

Research measuring emotion in response to film/video content has been published, with more than 25 journal papers since 1980 (e.g., [9]). However, despite the potential of (programmable) surround lighting, this existing research has not examined the effect of surround lighting on people's experience of video content. Much of the research has used multi-item and multi-dimensional validated psychometric instruments. Their use can be justified when the affect response is evaluated to longer pieces of media content that have been designed to evoke a potentially multi-dimensional affect response. However, this may not be appropriate for shorter pieces (e.g., with a duration of 6 s, as in the current study) that do not target a multi-dimensional affect response. Instead, for such pieces measuring valence (positive or negative) in response to stimuli with positive affective/pleasure or negative affective/displeasure quality will be more appropriate. This is because completing such instruments would be unwieldy to measure affect, with several different ratings per video clip in one test session [5]. In particular, this is likely to easily take a factor 10 more time (or even more, depending on the length of the particular multi-item scale that is used) to complete than the duration of the content that is being evaluated and therefore lead to potential response fatigue when multiple pieces of content are presented and affect response is measured. Therefore, the current study evaluates the effect of simple enhancements of stimulus affective quality (by way of surround lighting) on affect (measured with single-item scales) in response to brief video extracts. Moreover, we measure positive and negative affect separately, as these have been found to be independent factors in factor-analytic studies [35, 46].

Inference in UX evaluation

An important aspect of UX evaluation is comparative testing of designs [40]. Traditionally, inferential testing of the merits of alternative designs has been through testing the null hypothesis of no effect. However, several shortcomings of this approach have been identified [13, 36]). Magnitude-based inference offers an attractive alternative [7, 27] and its use in UX evaluation, and user research more generally, has also been advocated and

demonstrated [41]. In the current study, we chose to use magnitude-based inference in UX evaluation, as it has several advantages [41]. First, the approach requires the researcher to define a smallest important effect, instead of testing the unrealistic null hypothesis of no effect. Second, the approach uses the smallest important effect size, together with the observed effect, as an integral part of inference. As a consequence, inferences are not an artefact of sample size. Third, the approach provides a rigorous and principled approach to infer practical significance, and provides a rigorous distinction between mechanistic significance (tests to what extent and effect is positive, negative or trivial) and practical significance (tests to what extent an effect is beneficial, harmful or negligible). Inference can thereby address benefit and harm, and by doing so facilitates decision-making on the practical relevance of an effect. Fourth, the approach provides a refined classification of inferences that can be made with descriptors of the probability of each of three outcome ranges (positivity/benefit, triviality/negligibility and negativity/harm) (Table 1). A similar type of refined classification rather than an accept/rejection decision strategy is also used in other statistical inference (Jeffreys 1961, cited in [45]). In sum, in magnitude-based inference, the outcome is unclear or clear with a qualification of 21 combinations of outcome range and probability, but in null hypothesis-testing the outcome is limited to rejection or retaining of the null hypothesis.

Rationale and research questions

Programmable surround lighting has the potential to enhance UX of media content, but there is a lack of research demonstrating this. Therefore, our research aims are to (1) develop a method for testing UX in the use of video content with added programmable (controlled) surround lighting and (2) test the benefits of surround lighting. Given the novelty of this type of research, we chose to employ a simple, but realistic type of content as a starting point for future research that may use more complex content. Specifically, we address these aims using short video clips as media content with positive affective (pleasure) or negative affective (displeasure) quality that were selected to evoke a response of positive or negative affect, respectively. Moreover, this choice of content facilitates the manipulation of surround colour to enhance this affect. Specifically, we address the following research questions.

- (1) Can surround lighting enhance positive affect response to short pieces of video content with positive affective/pleasure quality?

- (2) Can surround lighting enhance negative affect response to short pieces of video content with negative affective/displeasure quality?

Method

Short pieces of video content were studied with the aim of evoking a response of mild positive affect or mild negative affect. This response might be enhanced by way of suitable programmable surround lighting (as discussed above).

Design

The design of this study allows for an examination of particular colours that may enhance participants' positive and negative affect. An experimental, repeated-measures design was used, with two presentation orders for counterbalancing (Table 2). The independent variable was added (surround) lighting colour. The levels were neutral lighting (white) (for all clips), yellow, green (both for clips with positive affective [pleasure] quality), red and purple (both for clips with negative affective [displeasure] quality). The dependent variable was experienced affect (positive or negative, in response to video clip content with positive affective/pleasure quality and negative affective/displeasure quality, respectively).

Participants

Participants were recruited by e-mail and word of mouth, and received a reimbursement of £8. Of the 33 participants (25 male, 8 female), all but one were students. Mean age was 25 (SD = 6). All spoke English (17 as a first language).¹

Materials and equipment

Video clips having content with negative or positive affective quality were selected from <http://search.creativecommons.org> and <http://vimeo.com/creativecommons>. A listing of all clips, including description, is presented in Online Appendix A1. Before the main study that is reported here, a pilot test was carried out, using the same procedure and measurement scales as in the main study, and only white surround lighting. This was to identify 18 clips that represented the pleasure dimension of positive affective quality evoking the most positive affect and 18 clips that represented the displeasure dimension of negative affective quality evoking the most negative affect. From

¹ Participants were not screened for colour blindness. We acknowledge this as a limitation of the study.

Table 1 Interpreting probabilities

Probability	Chances	Odds	The effect ... beneficial/trivial/harmful
<0; 0.005	<0; 0.5%	<0; 1:99.5	is almost certainly not ...
<0.005; 0.05	<0.5%; 5%	<1:995; 1:19	is very unlikely to be ...
<0.05; 0.25	<5%; 25%	<1:19; 1:3	is unlikely to be ..., is probably not ...
<0.25; 0.75	<25%; 75%	<1:3; 3:1	is possibly (not) ..., may (not) be ...
<0.75; 0.95	<75%; 95%	<3:1; 19:1	is likely to be..., is probably ...
<0.95; 0.995	<95%; 99.5%	<19:1; 995:1	is very likely to be ...
<0.995; 1	<99.5; 100	<99.5:1; ∞ >	is almost certainly ...

Interpretation according to Batterham and Hopkins [4]

Table 2 Presentation orders

Series	Order 1	Order 2
P1	8 clips, white surround lighting, positive affective quality	Same
P2	8 clips, white surround lighting, negative affective quality	Same
M1	12 clips, positive affective quality, 6 with <i>yellow</i> and 6 with white lighting, presented in quasi-random order	12 clips, positive affective quality, 6 with <i>green</i> and 6 with white lighting, presented in quasi-random order
M2	12 clips, positive affective quality, 6 with <i>yellow</i> and 6 with white lighting, presented in quasi-random order	12 clips, positive affective quality, 6 with <i>green</i> and 6 with white lighting, presented in quasi-random order
M3	12 clips, positive affective quality, 6 with <i>green</i> and 6 with white lighting, presented in quasi-random order	12 clips, positive affective quality, 6 with <i>yellow</i> and 6 with white lighting, presented in quasi-random order
M4	12 clips, positive affective quality, 6 with <i>green</i> and 6 with white lighting, presented in quasi-random order	12 clips, positive affective quality, 6 with <i>yellow</i> and 6 with white lighting, presented in quasi-random order
M5	12 clips, negative affective quality, 6 with <i>red</i> and 6 with white lighting, presented in quasi-random order	12 clips, negative affective quality, 6 with <i>purple</i> and 6 with white lighting, presented in quasi-random order
M6	12 clips, negative affective quality, 6 with <i>red</i> and 6 with white lighting, presented in quasi-random order	12 clips, negative affective quality, 6 with <i>purple</i> and 6 with white lighting, presented in quasi-random order
M7	12 clips, negative affective quality, 6 with <i>purple</i> and 6 with white lighting, presented in quasi-random order	12 clips, negative affective quality, 6 with <i>red</i> and 6 with white lighting, presented in quasi-random order
M8	12 clips, negative affective quality, 6 with <i>purple</i> and 6 with white lighting, presented in quasi-random order	12 clips, negative affective quality, 6 with <i>red</i> and 6 with white lighting, presented in quasi-random order

P practice, *M* main. For details of clips see Table A1 (Online Appendix A1)

the selected 18 ‘positive’ clips, the 12 most positive were selected for the main trials in the experiment and the remainder for the practice trials in the main study. Each of these clips was presented four times (Table 2). Similarly, from the selected 18 ‘negative’ clips, the 12 most negative were selected for the main trials in the experiment and the remainder for the practice trials. Each of these clips was presented four times (Table 2). In each series (practice or main), clips with either positive or negative affective quality were presented.

In each series of clips, we alternated surround lighting between a specified colour to enhance viewers’ experience (yellow or green to enhance positive affect; red or purple to enhance negative affect) and white as neutral surround lighting. We chose alternation because without this (e.g., a series of only yellow) the affect response to all video clips would be identically enhanced by the coloured surround lighting and the affect response to the content of individual

clips might be ‘drowned’ out by the overriding effect of surround light colour in the series.

There were two main series for each colour: in the first series 6 out of 12 clips with a particular affective quality were presented with coloured surround light and the other 6 six with white surround lighting (e.g., yellow in Series 1 under Order 1 with positive affective quality); in the second series the remaining 6 out of 12 clips were presented with coloured surround light and the other 6 six with white surround light (e.g., yellow in Series 2 under Order 1 with positive affective quality).

A standard PC (Steatite NDURA-R-433OTE) was used, with a Phillips 200w 20-inch LCD monitor, Phillips wired multimedia keyboard and wired HP mouse. A computer front projector (Hitachi CPX 250 XGA LCD Data/Video Projector) in the ceiling projected the video image onto a Sahara projector screen (153 cm × 122 cm). On the computer, bespoke presentation software, the amBX light

scene engine, controlled the presentation of (a) video clips on a standard projection screen by way of a standard (front) projector and (b) surround lighting by way of wireless LED lights (Freedom PAR TRI-6, CHAUVET DJ). A hardware unit (ENTTEC DMX USB PRO no. 2057300) was used for controlling the lights through the software, which was sent to the wireless LEDs via an antenna (D-FI TX 2.4).

The selection process for positive and negative colours to provide surround lighting was influenced by the considerations presented above. In particular, we selected yellow and green as colours to enhance positive affect, but red and purple as colours to enhance negative affect.

The surround lighting used five varieties of colour (with RGB co-ordinates in brackets): yellow (R = 175, G = 183, B = 12), green (R = 73, G = 127, B = 0), red (R = 51, G = 0, B = 0), purple (R = 48, G = 0, B = 99) and white (R = 48, G = 48, B = 48). There was additional surround lighting during the presentation of each clip (yellow, green or white for each clip with positive affective quality; red, purple or white for each clips with negative affective quality). Photographs of the set-up are presented in Online Appendix A2.

Per video clip, positive affect was measured with a single-item scale, with stem “Did this video make you feel positive?” and end-points ‘Not at all’ (0) and ‘Very much’ (10), and midpoint ‘Neutral’ (5). For negative affect, in the wording of the stem ‘positive’ was replaced with ‘negative’. We used 11-point scales rather than 5-point scales [5, 46] to enhance the sensitivity of measurement.

Procedure

Participants took part individually, alone or in small groups (2–4) under the direction of an experimenter. They first read an information sheet. They then signed a consent form. Next, the room lights were dimmed. Two practice series followed, with six ‘positive’ video clips in the first and six ‘negative’ clips in the second. On each trial, participants were presented with a video clip for 6 s. Then they were given 15 s to rate their affect in response to the clip. After the practice trials, room light was fully restored. The experimenter answered any queries from the participants. The room lights were dimmed again and eight main series followed, with 12 ‘positive’ video clips in the first four and 12 ‘negative’ clips in the second four. The same procedure for presentation and response was followed as in the practice series. After completion of the main series, participants were debriefed.

Results

Data for clips with positive affective quality and negative affective quality were analysed separately. In the analysis of positive affect, the following specific comparisons were made between different surround colours: yellow surround lighting with white, green with white and yellow with green. Given potential assimilation effects, specific comparisons were also made between white within series of yellow and white within series of green, and between the alternation of yellow and white and the alternation of green and white. Similar comparisons were made in the analysis of negative affect, where surround lighting was red, purple and white. Data analysis involved descriptive statistics and effect sizes the effect size measure, d (the standardised difference between two means²) was used.³ Inferential statistics were conducted through related t tests with SPSS, followed by magnitude-based inference⁴ (see above) (using the SPSS output as input) with dedicated spreadsheets [26]. This was followed by generalised linear mixed-model analysis with SPSS to verify the consistency of results.

Aggregated comparisons

When data are analysed aggregated over products the results may or may not show the same pattern as when they are aggregated over participants (‘product as a fixed-effect fallacy’; [34]). Therefore, inferential analyses examined the effect of surround lighting with participants as cases and data aggregated over clips, and also with clips as cases and data aggregated over participants.

Positive affect

Descriptives. Over clips with positive affective quality, on a scale with range [0; 10], the overall mean rating was 6.5 and mean scores for the different lighting conditions varied from 6.3 to 6.8 (Table 3). Therefore, the clips evoked slightly positive affect, as intended.

The descriptives over clips (Table 3) show small effects, in terms of mean differences, for yellow versus green ($d = 0.31$), for white (within yellow) versus white (within

² For example, the mean difference between red and white divided by the standard deviation of white as a baseline in the comparison, or the mean difference between red and purple divided by their pooled standard deviation (in the absence of a baseline in this comparison).

³ d is commonly employed as an effect size measure in user research [40].

⁴ In magnitude-based inference, results are here presented for a small effect (according to Cohen’s [10] conventions for this effect size measure) as the threshold for the smallest important beneficial or positive effect ($d = 0.2$) and for the smallest harmful or negative effect ($d = -0.2$).

Table 3 Descriptives for positive affect

Lighting	Mean	SD
Yellow	6.77	1.24
Green	6.25	2.02
White (Y)	6.72	1.05
White (G)	6.32	1.77
Alternation of yellow and white (Y)	6.75	1.12
Alternation of green and white (G)	6.29	1.87
Average	6.52	

Y white within alternating series of yellow and white surround lighting, G white within alternating series of green and white surround lighting

green) ($d = 0.27$) and for the alternation of yellow and white versus the alternation of green and white ($d = 0.30$). They also show extremely small effects for yellow versus white ($d = 0.04$) and for green versus white ($d = -0.04$).

In the analysis over participants, the means remained unchanged, but the standard deviations were reduced. Therefore, the effect sizes from descriptives over participants were large for yellow versus green ($d = 0.83$) and for the alternation of yellow and white versus the alternation of green and white ($d = 0.76$) and moderate for white (within yellow) versus white (within green) ($d = 0.56$). As before, the effects for yellow versus white ($d = 0.07$) and for green versus white ($d = -0.11$) were extremely small.

In sum, the pattern of effect sizes of the analysis aggregated over clips matched that of the analysis aggregated over participants, with meaningful positive effects of yellow over green, the alternation of yellow and white over the alternation of green and white, and white (within yellow) versus white (within green).

Magnitude-based inference. According to mechanistic inference, the effect of yellow in comparison with green was likely positive (76.02%); given that yellow was likely beneficial (76.02%) and most unlikely (0.11%) harmful, this should be chosen over green. As further evidence for the advantage of yellow over green, the effect of the alternation of yellow and white in comparison with the alternation of green and white was possibly positive (74.29%); given that the alternation yellow and white was possibly beneficial (74.29%) and most unlikely (0.11%) harmful, this should be chosen over the alternation of green and white. The effect of white (within yellow) in comparison with white (within green) was possibly positive (69.29%). This suggests two assimilation effects: within a combined series of yellow and white, the ratings of video clips with white may increase due to the relatively high ratings with yellow within the same series, but within a combined series of green and white, the ratings with white may decrease due to the relatively low ratings with green

within the same series. Because the clear result was in favour of white within yellow, this provides also further evidence for the advantage of yellow over green.

The difference between yellow and white, and that between green and white were very likely trivial (95.18 and 98.48%, respectively); given that, according to practical inference, the effects were very likely negligible (95.18% and 98.48%, respectively), very/most unlikely harmful (4.36 and 0.08%, respectively) and most/very unlikely beneficial (0.46 and 1.46%, respectively), the results indicate that yellow or green should not be chosen over white. However, these results provide further evidence for the assimilation effects reported above and need to be considered with these effects in mind.

The results aggregated over participants (Table 4b) show the same pattern as those aggregated over video clips (Table 4a), providing further evidence of internal validity, but with larger effect sizes. In sum, the results of mechanistic inference consistently show clear effects in favour of yellow over green, both in the analysis aggregated over clips and in the analysis aggregated over participants.

Generalised linear mixed-model analysis. Our mixed-model analysis was motivated by Monk's [34] emphasis on including product as a factor. Therefore, in our analysis, we distinguished two levels: clip (at Level 1, 12 hazards existed) and subject (or participant; at Level 2, 33 participants existed); see Online Appendix A3, Table A3.1, for the complete model specification. The aim was to predict the binary comparison that was made [yellow versus white; green versus white; yellow versus green; white (within yellow) versus white (within green); alternation of yellow/white versus alternation green/white], so the outcome variable was the specific comparison that was examined. The predictors were positive affect per clip (at Level 1) and positive affect over clips (at Level 2). These allowed us to analyse the compositional effect: the extent to which the size of a relationship at a higher level (overall, across clips) adds to the effect at a lower level (for individual clips) [23]. The results of mixed-model analysis (see Online Appendix A3, Table A3.1) show the same pattern as those of magnitude-based inference: first, the positive effect of yellow over green was statistically significant, as were the positive effects of white (within yellow) over white (within green) and of the alternation of yellow and white over the alternation of green and white; second, the remaining effects were not significant. The significant results were found at Level 2 (overall, across clips) and with larger effect size, presumably due to variability among clips, thereby reducing the effect at Level 1 (see also further analysis below). In sum, the pattern of the results of mixed-modelling matches that of magnitude-based inference (presented in the previous subsection).

Table 4 Magnitude-based inference for positive affect

Comparison	Mean	SD	90%-confidence interval of mean diff.	d	Chances		Odds ratio (B/H)	Inference
					Beneficial/+ive	Negligible/trivial		
a. Data averaged over video clips								
1 Yellow-white	6.77	1.24	0.047, − 0.11 to 0.2	0.04	4.36%	95.18%	0.46%	9.90 M: very likely trivial
	6.72	1.05			Very unlikely	Very likely	Most unlikely	P: very likely trivial; don't use
2 Green-white	6.25	2.02	− 0.071, − 0.28 to 0.14	− 0.04	0.08%	98.48%	1.43%	0.06 M: very likely trivial
	6.32	1.77			Most unlikely	Very likely	Very unlikely	P: very likely trivial; don't use
3 Yellow-green	6.77	1.24	0.52, 0.084 to 0.95	0.31	76.02%	23.87%	0.11%	2875.35 M: likely +ive
	6.25	2.02			Likely	Unlikely	Most unlikely	P: likely beneficial; use
4 White(Y)-white(G)	6.72	1.05	0.4, 0.04 to 0.76	0.27	69.29%	30.57%	0.13%	1680.99 M: possibly +ive
	6.32	1.77			Possibly	Possibly	Most unlikely	P: possibly beneficial; use
5 Y/W(Y)-G/W(G)	6.75	1.12	0.46, 0.07 to 0.85	0.30	74.29%	25.61%	0.11%	2703.18 M: possibly +ive
	6.29	1.87			Possibly	Possibly	Most unlikely	P: possibly beneficial; use
b. Data averaged over participants								
1 Yellow-white	6.77	0.64	0.047, − 0.31 to 0.41	0.07	33.55%	47.19%	19.26%	2.12 M: unclear; get more data
	6.72	0.70			Possibly	possibly	unlikely	P: unclear; don't use; get more data
2 Green-white	6.25	0.60	− 0.071, − 0.32 to 0.18	− 0.11	0.62%	95.86%	3.52%	0.17 M: very likely trivial
	6.32	0.72			Very unlikely	very likely	very unlikely	P: very likely trivial; don't use
3 Yellow-green	6.77	0.64	0.52, 0.28 to 0.76	0.83	99.31%	0.66%	0.03%	429820.93 M: very likely +ive
	6.25	0.60			Very likely	very unlikely	most unlikely	P: very likely beneficial; use
4 White(Y)-white(G)	6.72	0.70	0.4, 0.18 to 0.61	0.56	97.26%	2.69%	0.05%	68217.28 M: very likely +ive
	6.32	0.72			very likely	very unlikely	most unlikely	P: very likely beneficial; use
5 Y/W(Y)-G/W(G)	6.75	0.58	0.46, 0.31 to 0.61	0.76	99.91%	0.09%	0.00%	71448655.63 M: most likely +ive
	6.29	0.62			Most likely	most unlikely	most unlikely	P: most likely beneficial; use

Thresholds for important effect size $d = \pm 0.2$. d observed effect size. B benefit, H harm, M mechanistic inference, P practical inference

Table 5 Descriptives for negative affect

Lighting	Mean	SD
Red	6.12	1.64
Purple	5.74	1.69
White (R)	5.77	1.39
White (P)	5.58	1.58
Red and white (R)	5.94	1.47
Purple and white (P)	5.66	1.59
Average	5.80	

R white within alternating series of yellow and white surround lighting, *P* white within alternating series of green and white surround lighting

Negative affect

Descriptives. Over clips with negative affective quality, on a scale with range [0; 10], the overall mean rating was 5.8 and mean scores for the different lighting conditions varied from 5.6 to 6.1 (Table 5). Therefore, the clips evoked slightly negative affect, as intended.

The descriptives over clips (Table 5) show small effects, in terms of mean difference, for red versus white ($d = 0.25$), for red versus purple ($d = 0.23$), and for the alternation of red and white versus the alternation of purple and white ($d = 0.19$). They also show very small effects for purple versus white ($d = 0.10$) and for white (within red) versus white (within purple) ($d = 0.13$).

In the analysis over participants, the means remained unchanged, but the standard deviations were reduced. Therefore, the effect sizes from descriptives over participants were large for red versus purple ($d = 0.81$), and moderate for red versus white ($d = 0.69$) and for the alternation of red and white versus the alternation of purple and white ($d = 0.63$). As before, the effects for purple versus white ($d = 0.30$) and for green versus white ($d = 0.34$) were substantially smaller.

In sum, the pattern of effect sizes of the analysis aggregated over clips matched that of the analysis aggregated over participants, with meaningful positive effects of red over white, red over purple, and the alternation of red and white over the alternation of purple and white.

Magnitude-based inference. According to mechanistic inference, the effect of red in comparison with white was possibly positive (69.90%); given that, according to practical inference, red was possibly beneficial (69.90%) and most unlikely (0.002%) harmful, this should be chosen over white. The effect of red in comparison with purple was possibly positive (62.74%), and given that red was possibly beneficial (62.74%) and most unlikely harmful (0.001%), this should be chosen over purple. The effect of the alternation of red and white in comparison with the

alternation of purple and white was possibly positive (43.05%); given that the effect was possibly beneficial (43.05%) and most unlikely (0.001%) harmful, the former should be chosen over the latter. Combined, these results show an advantage of red over purple and white.

The effect of purple in comparison with white was likely trivial (86.61%); given that purple was unlikely beneficial (13.26%) and most unlikely harmful (0.12%), this should not be chosen over white. The effect of white (within red) in comparison with white (within purple) was unlikely positive (20.83%) and likely trivial (79.12%), thus providing evidence against an assimilation effect.

The results aggregated over participants (Table 6b) show the same pattern as those aggregated over video clips (Table 6a), providing further evidence of internal validity, but with larger effect sizes. In sum, the results of mechanistic inference consistently show clear effects in favour of red over purple and white, both in the analysis aggregated over clips and in the analysis aggregated over participants.

Generalised linear mixed-model analysis. In our mixed-model analysis, two levels can be distinguished: clip (at Level 1, 12 hazards existed) and subject (or participant; at Level 2, 33 participants existed); see Online Appendix A3, Table A3.2, for the complete model specification. The aim was to predict the binary comparison that was made [red versus white; purple versus white; red versus purple; white (within red) versus white (within purple); alternation of red/white versus alternation purple/white], so the outcome variable was the specific comparison that was examined. The predictors were negative affect per clip (at Level 1) and negative affect over clips (at Level 2). These allowed us to analyse the compositional effect: the extent to which the size of a relationship at a higher level (overall, across clips) adds to the effect at a lower level (for individual clips) [23]. The results of mixed-model analysis (see Online Appendix A3, Table A3.2) show the same pattern as those of magnitude-based inference: first, the positive effect of red over white was statistically significant, as were the positive effects of red over purple and of the alternation of red and white over the alternation of purple and white; second, the remaining effects were not significant. The significant results were achieved at Level 2 (overall, across clips) and with larger effect size, presumably due to variability among clips, thereby reducing the effect at Level 1 (see also further analysis below). In sum, the pattern of the results of mixed-modelling matches that of magnitude-based inference (presented in the previous subsection).

Comparisons by clip

In the main series, there were 12 clips with positive affective quality and 12 clips with negative affective

Table 6 Magnitude-based inference for negative affect

Comparison	Mean	SD	90%-confidence interval of mean diff.	<i>d</i>	Chances			Odds ratio (B/H)	Inference	
					Beneficial/ + ive	Negligible/ trivial	Harmful/- ive			
a. Data averaged over video clips										
1 Red-white	6.12	1.64	0.35, 0.12 to 0.57	0.25	69.90%	30.10%	0.00%	97560.37	M:	possibly +ive
	5.77	1.39			Possibly	Possibly	Most unlikely		P:	possibly beneficial; use
2 Purple-white	5.74	1.69	0.15, -0.088 to 0.4	0.10	13.26%	86.61%	0.12%	125.83	M:	likely trivial
	5.58	1.58			Unlikely	Likely	Most unlikely		P:	likely trivial; don't use
3 Red-purple	6.12	1.64	0.38, 0.13 to 0.63	0.23	62.74%	37.25%	0.00%	132791.09	M:	possibly +ive
	5.74	1.69			Possibly	Possibly	Most unlikely		P:	possibly beneficial; use
4 White(R)-white(P)	5.77	1.39	0.19, -0.041 to 0.41	0.13	20.83%	79.12%	0.05%	504.04	M:	unlikely +ive; likely trivial
	5.58	1.58			Unlikely	Likely	Most unlikely		P:	likely trivial; don't use
5 R/W(R)-P/W(P)	5.94	1.47	0.28, 0.074 to 0.49	0.19	43.05%	56.95%	0.00%	38333.92	M:	possibly +ive
	5.66	1.59			Possibly	Possibly	Most unlikely		P:	possibly beneficial; use
b. Data averaged over participants										
1 Red-white	6.12	0.47	0.35, 0.067 to 0.63	0.69	92.84%	6.39%	0.77%	1671.22	M:	likely +ive
	5.77	0.53			likely	unlikely	very unlikely		P:	unclear; don't use; get more data
2 Purple-white	5.74	0.46	0.15, -0.069 to 0.38	0.30	65.32%	31.58%	3.10%	58.97	M:	possibly +ive
	5.58	0.57			possibly	possibly	very unlikely		P:	unclear; don't use; get more data
3 Red-purple	6.12	0.47	0.38, 0.095 to 0.67	0.81	95.09%	4.29 %	0.62%	3098.77	M:	very likely +ive
	5.74	0.46			very likely	very unlikely	very unlikely		P:	unclear; don't use; get more data
4 White(R)-white(P)	5.77	0.53	0.19, 0.00084 to 0.37	0.34	76.19%	23.05%	0.76%	416.93	M:	likely +ive
	5.58	0.57			likely	unlikely	very unlikely		P:	unclear; don't use; get more data
5 R/W(R)-P/W(P)	5.94	0.43	0.28, 0.11 to 0.46	0.63	96.52%	3.35%	0.13%	21393.01	M:	very likely +ive
	5.66	0.47			very likely	very unlikely	most unlikely		P:	very likely beneficial; use

Thresholds for important effect size $d = \pm 0.2$. *d* observed effect size. *B* benefit, *H* harm, *M* mechanistic inference, *P* practical inference

quality.⁵ Because individual clips were not selected based on a particular research question or hypothesis, the analysis of individual clips cannot provide answers; rather, it may provide questions for future research. Therefore, a detailed exploratory analysis of the results per clip is potentially useful to identify specific surround lighting design characteristics that can enhance video content. In this analysis, the aim is, first, to quantitatively identify individual clips that show a pronounced effect of surround lighting (e.g., red in comparison with white purple) and, second, to

qualitatively analyse these clips to identify particular film design characteristics that might make them more likely to be enhanced through surround lighting. From this exploratory analysis, hypotheses regarding film design characteristics in relation to the enhancing effect of surround lighting may be tested in future research. Because many clips are involved in the comparison of enhancements (e.g., 12 clips are involved in the comparison of red with white), the chance of incorrect inferences (e.g., the chance of a positive or negative effect) is inflated considerably (by more than 10 times, given the number of clips) and inference is not meaningful anymore. For this reason, in the following results by clip no inference is made. Instead, emphasis should be placed on the size of the effect

⁵ In Series 4, Clip 2 was, incorrectly, presented with white rather than green surround lighting. Therefore, the response to this clip was not analysed in comparisons in any of the Series 1–4.

Table 7 Descriptives for positive affect per positive clip

Clip	Yellow		Green	
	Mean	SD	Mean	SD
1	6.76	2.17	6.36	2.47
2	6.64	2.28		
3	6.94	2.21	6.27	2.81
4	5.39	2.65	4.85	2.97
5	7.06	2.46	6.09	3.08
6	6.82	2.35	6.03	2.59
7	7.52	1.70	7.33	1.81
8	7.39	2.40	6.33	2.85
9	5.82	2.64	6.03	2.58
10	6.79	2.53	6.70	3.05
11	7.27	2.39	6.18	2.95
12	6.70	2.52	6.58	2.85

Due to a design fault, Positive Clip 2 was not presented with green surround colour

(d), with probabilities of a beneficial/positive, negligible/trivial and harmful/negative effect as additional information only (see Online Appendix A4).

Positive affect by clip, yellow versus green. In the comparison of yellow and green (Table 7; Fig. 1), there were *likely positive/beneficial* small to moderate effects of yellow for Clips 5, 6, 8 and 11.⁶ The beneficial effect of yellow surround lighting for *Clip 5* (Dog on slide) may have been due to a reference by the yellow lighting to the sun in the outdoor scene; it could link to the playful nature and happiness of the clip. The beneficial effect for *Clip 6* (Love) may have been due to the yellow lighting matching the outside colours in the scene, and a reference to sun, warmth and love. Again, yellow could also evoke happiness in the audience. The beneficial effect for *Clip 8* (Baby kisses dog) may have been due again to the yellow lighting linking to the sun in the outdoor scene. The clip is of a playful happy nature where a child kisses puppies, linking to the happy lovingness of the child. Therefore, yellow could communicate these feelings to the audience. The beneficial effect for *Clip 11* (Baby plays with dog) may be explained as follows. This clip, again, is of a happy playful nature where a child plays with a dog. In the audience, yellow could trigger these happy feelings. Furthermore, the clip is taken from outside and the yellow colour references the sun and its warmth to help communicate these feelings to the audience.

Negative affect by clip, red versus purple. In the comparison of red and purple (Table 8; Fig. 2), small to moderate *likely positive/beneficial* effects of red surround

lighting were found in particular for Clips 8 and 10.⁷ The beneficial effect of red surround lighting for *Clip 8* (Traffic jam) may be explained as follows. The red surrounding lighting on this clip references red traffic light signals, stopping and the red cars within the clip. However, purple does not possess the colour references that red does in the traffic environment. Therefore, the matching of red surround colour with colours in the scene may be the cause for the beneficial effect of red over purple in this clip. The beneficial effect for *Clip 10* (Stepping into dirt) may be due to again the red surrounding light colour matching what is on screen, the red bucket, red clay and mud, whereas purple does not.

Discussion

The effect of surround lighting on UX

It is important to note that the results for positive affective quality presented in this section, occurred within the context of assimilation effects (see next section) that are specific to the research design that was used.⁸ In response to Research Question 1, the results over clips with positive affective quality indicate that yellow surround lighting can enhance people's positive affect more than and should be used in favour of green surround lighting. Moreover, alternating presentations of yellow and white surround lighting can enhance people's positive affect more than and should be used in favour of alternating presentations of green and white surround lighting. Because of assimilation effects, no advantage of yellow or green over white could be demonstrated.

In response to Research Question 2, red surround lighting can enhance people's negative affect more than and should be used in favour of purple surround lighting. Moreover, the results over clips with negative affective quality indicate that red surround lighting can enhance people's negative affect more than and should be used in favour of white surround lighting. Furthermore, alternating presentations of red and white surround lighting can enhance people's negative affect more than and should be used in favour of alternating presentations of purple and white surround lighting.

The results for positive and negative affect can be considered remarkable and encouraging, because affect response was enhanced by particular colours even though

⁶ There was one negative effect, with size very close to 0, for Clip 9. All other effects were positive.

⁷ There was a negative effect with size very close to 0 for Clip 3 and a negative small negative effect for Clip 5. All other effects were positive.

⁸ If a different research design would have been used, potential assimilation or contrast effects would still occur that would be specific to that design.

Fig. 1 Mean ratings and effect size of positive affect per positive clip

Table 8 Descriptives for negative affect per negative clip

Clip	Red		Purple	
	Mean	SD	Mean	SD
1	6.00	2.99	5.15	2.81
2	5.88	2.29	5.60	2.29
3	5.88	2.03	6.09	2.53
4	6.18	2.83	5.48	2.62
5	5.36	2.61	6.15	2.51
6	5.42	3.03	4.81	2.79
7	6.45	1.86	6.06	2.19
8	7.00	2.49	5.76	2.77
9	5.91	3.20	5.85	2.96
10	6.33	2.25	5.39	2.22
11	6.39	2.25	6.29	2.45
12	6.61	2.97	6.21	2.57

(a) the video clips had not been specifically designed to create a specific affect response, but were rather selected from existing material, (b) the enhancements were simple: constant light of a particular colour without change of colour, (c) the enhancements were not specifically designed to enhance the affective content of each individual selected clip and (d) the clips were only of (very) short duration (six seconds).

The results per clip highlight specific clips that were notable in terms of benefit from surround lighting to positive affect response or to negative affect response. From the analysis of these clips, we derive the following hypotheses regarding surround lighting design characteristics (subjective UX in design) that make video content more likely or less likely to be enhanced in terms of a particular affect response. These hypotheses provide a starting point, to be refined based on studying relevant existing literature, before they are tested in future work.

Hypothesis 1: surround lighting (e.g., yellow) that is associated with positive affective quality of the video content enhances positive affect.

Hypothesis 2: surround lighting (e.g., red) that is associated with negative affective quality of the video content enhances negative affect.

Hypothesis 3: surround lighting that references the (dominant) background colour or colour of objects used in the video content enhances positive or negative affect.

Assimilation effects

Any results regarding positive affect from this study must be interpreted in the light of assimilation effects that were found and that are specific to this study. In the analysis of both clips with positive affective quality and those with negative affective quality, evidence for an *assimilation effect* was found. In particular, the ratings of clips with positive affective quality presented with white surround lighting in a series alternating yellow and white were higher (possibly positive effect) than the ratings of the same clips presented with white surround lighting in a series alternating green and white.⁹ Presumably partly as a consequence of the possibly positive assimilation effect (of yellow as a context over green), a beneficial effect of yellow over white could not be demonstrated. However, a small to moderate likely beneficial effect of yellow over green could be demonstrated in the analysis aggregated over clips and a large very likely positive effect in the analysis aggregated over participants.

Although assimilation effects may be undesirable when the aim is to test the effectiveness of surround lighting in enhancing people's affect response to video content (and media content more generally), assimilation effects are

⁹ Moreover, the positive effect of the alternation of yellow and white over the alternation of green and white was equally large as the positive effect of yellow over green.

Fig. 2 Mean ratings and effect size of negative affect per negative clip

presumably inevitable.¹⁰ Further evidence comes from various research studies published in the psychology literature. For example, in a review of perception research, Lockhead [33] concludes that the judgement of the quality of the current stimulus is influenced by the quality of previously presented stimuli. Similarly, Kondo, Takahashi and Watanabe [30] found that people's judgements of the attractiveness of visual stimuli depend on the attractiveness of previously presented stimuli in the same series. Moreover, Kusev, Ayton, van Schaik, Tsaneva-Atanasova, Stewart and Chater [32] found consistently across sensory domains that people's frequency judgement of a series of stimuli in terms of a stimulus feature (e.g., colour) is influenced by the category (e.g., red) of the first repeated sequence of stimuli within the series (so the frequency of red as a stimulus colour over the series is overestimated if the first repeated sequence of stimuli has the colour red). In sum, assimilation effects cannot be eliminated in the design of studies; instead, they need to be carefully considered and managed in research design.

However, although previously presented stimuli influence the perception of the current stimulus, assimilation is not universal. Assimilation happens when, for example, two stimulus categories are sufficiently similar, but contrast happens when they are sufficiently different (e.g., [6]. A contrast effect occurs when the difference in perception between, for example, two stimulus categories is increased rather than decreased when they are presented closely in time or simultaneously. Both contrast effects and assimilation effects have been observed, for example, in judgements of physical attractiveness [12], categorization and exemplar production [47], and the visual perception of natural scenes [28].

¹⁰ As an advantage of assimilation effects, we demonstrated that even if affect-enhancing surround lighting is presented half of the time in a series (e.g., yellow in a series of alternating yellow and white surround lighting), it can still have a positive/beneficial effect.

Evaluation method

The measurement of affect in response to video content was limited in the sense that only one response from one type of affect (positive affect or negative affect) was recorded per participant per video clip. In response to clips selected for positive affective quality, positive affect was measured, and in response to clips selected for negative affective quality, negative affect was measured. However, according to previous research, the approach of measuring only one type of affect (e.g., positive) in response to a particular type of content affective quality (e.g., positive) can be misleading [21]. For example, a manipulation of surround lighting may enhance more than one type of affect. In other words, our results demonstrate sensitivity (e.g., positive affect was enhanced by yellow surround lighting compared to green), but not specificity (e.g., we do not know whether the effect of yellow versus green was specific to the response variable positive affect).

Similarly, the enhancement of affect type was confounded with colour of lighting. In particular, positive affect was tested with yellow, green and white light, but negative affect was tested with red, purple and white light. Therefore, it remains unknown to what extent red or purple might enhance positive affect or yellow or green might enhanced negative affect. Testing each colour on both positive and negative affect would provide further evidence of specificity.

It was our objective to study how surround lighting can enhance the valence dimension of human affect response rather than manipulate arousal (activation/deactivation) as subjective product quality and study the effect on arousal response. Therefore, we did not measure arousal. Furthermore, we did not select our clips to be high or low on arousal, so arousal was not controlled. However, pleasure as affective quality and displeasure as affective quality were controlled and our results of both the pilot study and the main study provide supportive evidence.

Future work

Complexity of UX. Given the novelty of this type of research, we chose to employ a simple, but realistic type of content as a starting point for future research that may use more complex content. Specifically, we examined how the affect response to content with positive or negative affective quality could be enhanced by coloured surround lighting. Future work may analyse how more complex content can be enhanced. For example, complex media content with a particular story line may aim to evoke different specific affect responses (or emotions) (UX in use) in sequence or simultaneously. This may require enhancement by various specific complex combinations or sequences of coloured surround lighting. It follows from Kurosu's [31] two-dimensional framework for the design and evaluation of UX that subjective UX in design and subjective UX in use need to be carefully co-ordinated to make sure that the desired effect (affect response or a specific emotional response) can be achieved and measured. In the current study, this was achieved by selecting simple video clip content with positive or negative affective quality and then measuring positive or negative affect, respectively. However, this will be more challenging with more complex content and more complex subjective UX in design.

Therapeutic application. The findings of the current research regarding the effect of programmable surround lighting on affect provide a starting point for research into its use in interventions to ease or prevent a range of affective disorders. For example, research has highlighted the therapeutic utility of light in treating depression [20]. Golden and colleagues specifically found that bright-light therapy was effective in the treatment of seasonal and non-seasonal depression with large and moderate effect sizes, respectively. Note that the existing literature focused on the therapeutic effects of bright lights. Programmable surround lighting not only has the capability to deliver this, but also offers flexibility, for example in terms of type of lighting, intensity and duration, and therefore facilitates research to improve existing therapies or develop new ones.

Other potential therapeutic applications of programmable surround lighting are in conditions such as Alzheimer's Disease and Related Dementia (ADRD), chronic and acute pain, and anxiety. This is because bright-light therapy in an ADRD population significantly improves night time sleep [17] and reduces cognitive decline [25]. Furthermore, pain is strongly linked to depression, and either condition can be cause or effect of the other [19]. Therefore, the therapeutic application of programmable surround lighting may improve mood in chronic-pain patients and thereby reduce pain. Moreover, because of its calming effect, programmable ambient

lighting could provide a therapy to ease anxiety disorders [8], as an alternative or adjunct to pharmacological [3] and psychological interventions [43].

Recommendations

Regarding our first aim (developing a method for testing UX), for studies with multiple short clips that are rated in terms of affect, research designs need to address both sensitivity and specificity in terms of the manipulation of surround lighting and the measurement of affect. As regards our second aim (testing the benefits of controlled surround lighting), in order to enhance positive affect in response to short clips of video with positive content affective quality, yellow surround lighting should be considered. In order to enhance negative affect in response to short clips of video with negative content affective quality, red surround lighting should be considered.

Conclusion

Our test results show how coloured surround lighting can enhance the affect response to media content and provide guidance for UX in design. In addition, we have shown how specific hypotheses can be derived from the results for individual clips, as a basis to further support this enhancement. Furthermore, our results have implications for the interpretation of test results and the design of evaluation studies. We look forward to future research building on our approach to designing for and evaluating UX enhanced by programmable surround lighting and applying Kurosu's [31] framework for the design and evaluation of UX.

Acknowledgements The authors are grateful to Innovate UK for funding this research and to Vasiliki Sergi for assistance with data collection.

References

1. amBX (2014) The amBX light-scene engine for commercial lighting. <http://www.ambx.com/ambx-light-scene-engine-commercial-lighting>. Accessed 24 Oct 2017
2. amBX (2016) Case studies (2010–2016). <http://www.ambx.com/case-studies>. Accessed 24 Oct 2017
3. Baldwin DS, Anderson IM, Nutt DJ, Bandelow B, Bond A, Davidson JR, Wittchen HU (2005) Evidence-based guidelines for the pharmacological treatment of anxiety disorders: recommendations from the British association for psychopharmacology. *J Psychopharmacol* 19(6):567–596
4. Batterham AM, Hopkins WG (2006) Making meaningful inferences about magnitudes. *Int J Sports Physiol Perform* 1(1):50–57

5. Bradley MM, Lang PJ (1994) Measuring emotion: the self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry* 25(1):49–59
6. Brown SA, Venkatesh V, Goyal S (2014) Expectation confirmation in information systems research: a test of six competing models. *MIS Q* 38(3):729
7. Buchheit M (2016) The numbers will love you back in return-I promise. *Int J Sports Physiol Perform* 11(4):551–554. doi:[10.1123/IJSP.2016-0214](https://doi.org/10.1123/IJSP.2016-0214)
8. Canazei M, Dehoff P, Staggl S, Pohl W (2014) Effects of dynamic ambient lighting on female permanent morning shift workers. *Lighting Res Technol* 46(2):140–156
9. Carvalho S, Leite J, Galdo-Álvarez S, Gonçalves OF (2012) The emotional movie database (EMDB): a self-report and psychophysiological study. *Appl Psychophysiol Biofeedback* 37(4):279–294. doi:[10.1007/s10484-012-9201-6](https://doi.org/10.1007/s10484-012-9201-6)
10. Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Erlbaum, Hillsdale
11. Cole R, Eves D (2003) PML: when your room becomes your browser. In: Aarts E, Marzano S (eds) *The new everyday: views on ambient intelligence*. 010 Publishers, Rotterdam, pp 312–313
12. Cyprianska M, Bedynska S, Golec De Zavala A (2012) The role of similarity and dissimilarity focus in sequential judgments of physical attractiveness. *J Soc Psychol* 152(3):266–269. doi:[10.1080/00224545.2011.609846](https://doi.org/10.1080/00224545.2011.609846)
13. Cumming G (2014) The new statistics: why and how. *Psychol Sci* 25(1):7–29. doi:[10.1177/0956797613504966](https://doi.org/10.1177/0956797613504966)
14. Eves D, Cole R (2000) Physical mark-up language. Patent No. WO 02/092183. Switzerland: World Intellectual Property Organisation
15. Eves D, Cole R (2001) Dynamic mark-up language. Patent No. WO 03/100548. Switzerland: World Intellectual Property Organisation
16. Eves D, Cole R (2013) Unique intelligent software for digital lighting control (white paper). Redhill, United Kingdom: amBX
17. Figueiro MG, Hunter CM, Higgins PA, Hornick TR, Jones GE, Plitnick B, Rea MS (2015) Tailored lighting intervention for persons with dementia and caregivers living at home. *Sleep Health* 1(4):322–330
18. Garnett O (2000) Colour: a social history. National Trust, Manvers
19. Gerrits MM, Vogelzangs N, van Oppen P, van Marwijk HW, van der Horst H, Penninx BW (2012) Impact of pain on the course of depressive and anxiety disorders. *Pain* 153(2):429–436
20. Golden RN, Gaynes BN, Ekstrom RD, Hamer RM, Jacobsen FM, Suppes T, Nemeroff CB (2005) The efficacy of light therapy in the treatment of mood disorders: a review and meta-analysis of the evidence. *Am J Psychiatry* 162(4):656–662
21. Gross JJ, Levenson RW (1995) Emotion elicitation using films. *Cognit Emot* 9(1):87–108
22. Hassenzahl M, Tractinsky N (2006) User experience - a research agenda. *Behav Inf Technol* 25(2):91–97. doi:[10.1080/01449290500330331](https://doi.org/10.1080/01449290500330331)
23. Heck RH, Thomas SL, Tabata LN (2010) Multilevel and longitudinal modeling with IBM SPSS. Routledge Academic, London
24. Hekkert P, Schifferstein H (2008) Product experience. Elsevier, Amsterdam
25. van Hoof J, Schoutens AMC, Aarts MPJ (2009) High colour temperature lighting for institutionalised older people with dementia. *Build Environ* 44(9):1959–1969
26. Hopkins WG (2007) A spreadsheet for deriving a confidence interval, mechanistic inference and clinical inference from a *p* value. *Sport Sci* 11:16–20
27. Hopkins WG, Batterham AM (2016) Error rates, decisive outcomes and publication bias with several inferential methods. *Sports Med* 46(10):1563–1573. doi:[10.1007/s40279-016-0517-x](https://doi.org/10.1007/s40279-016-0517-x)
28. Howe CQ, Purves D (2004) Size contrast and assimilation explained by the statistics of natural scene geometry. *J Cognit Neurosci* 16(1):90–102. doi:[10.1162/089892904322755584](https://doi.org/10.1162/089892904322755584)
29. International Organization for Standardization/International Electrotechnical Commission (2011) ISO/IEC 25010—Systems and software engineering—systems and software Quality Requirements and Evaluation (SQuaRE)—system and software quality models. Authors, Switzerland
30. Kondo A, Takahashi K, Watanabe K (2013) Influence of gender membership on sequential decisions of face attractiveness. *Atten Percept Psychophys* 75(7):1347–1352. doi:[10.3758/s13414-013-0533-y](https://doi.org/10.3758/s13414-013-0533-y)
31. Kurosu M (2017) Theory of user engineering. CRC Press, Boca Raton
32. Kusev P, Ayton P, van Schaik P, Tsaneva-Atanasova K, Stewart N, Chater N (2011) Judgments relative to patterns: how temporal sequence patterns affect judgments and memory. *J Exp Psychol Human Percept Perform* 37(6):1874–1886. doi:[10.1037/a0025589](https://doi.org/10.1037/a0025589)
33. Lockhead GR (2004) Absolute judgments are relative: a reinterpretation of some psychophysical ideas. *Re Gen Psychol* 8(4):265–272. doi:[10.1037/1089-2680.8.4.265](https://doi.org/10.1037/1089-2680.8.4.265)
34. Monk A (2004) The product as a fixed-effect fallacy. *Human-Comp Interact* 19(4):371–375
35. Mackinnon A, Jorm AF, Christensen H, Korten AE, Jacomb PA, Rodgers B (1999) A short form of the positive and negative affect schedule: evaluation of factorial validity and invariance across demographic variables in a community sample. *Personal Individ Differ* 27(3):405–416. doi:[10.1016/S0191-8869\(98\)00251-7](https://doi.org/10.1016/S0191-8869(98)00251-7)
36. Murphy KR, Myers B (1999) Testing the hypothesis that treatments have negligible effects: minimum-effect tests in the general linear model. *J Appl Psychol* 84(2):234–248. doi:[10.1037/0021-9010.84.2.234](https://doi.org/10.1037/0021-9010.84.2.234)
37. Neumeister A, Goessler R, Lucht M, Kapitany T, Bamas C, Kasper S (1996) Bright light therapy stabilizes the antidepressant effect of partial sleep deprivation. *Biol Psychiatry* 39:16–21
38. Russell JA (1980) A circumplex model of affect. *J Personal Soc Psychol* 39(6):1161
39. Russell JA (2003) Core affect and the psychological construction of emotion. *Psychol Rev* 110:145–172
40. Sauro J, Lewis JR (2012) Quantifying the user experience. Morgan Kaufmann, Amsterdam
41. van Schaik P, Weston M (2016) Magnitude-based inference and its application in user research. *Int J Human Comp Stud* 88:38–50. doi:[10.1016/j.ijhcs.2016.01.002](https://doi.org/10.1016/j.ijhcs.2016.01.002)
42. Sheldon KM, Elliot AJ, Kim Y, Kasser T (2001) What is satisfying about satisfying events? Testing 10 candidate psychological needs. *J Personal Soc Psychol* 80(2):325–339. doi:[10.1037/0022-3514.80.2.325](https://doi.org/10.1037/0022-3514.80.2.325)
43. Smits JA, Berry AC, Tart CD, Powers MB (2008) The efficacy of cognitive-behavioral interventions for reducing anxiety sensitivity: a meta-analytic review. *Behav Res Ther* 46(9):1047–1054
44. Thüring M, Mahlke S (2007) Usability, aesthetics and emotions in human-technology interaction. *Int J Psychol* 42(4):253–264. doi:[10.1080/00207590701396674](https://doi.org/10.1080/00207590701396674)
45. Wagenmakers E, Wetzels R, Borsboom D, van der Maas HLJ (2011) Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). *J Personal Soc Psychol* 100(3):426–432. doi:[10.1037/a0022790](https://doi.org/10.1037/a0022790)
46. Watson D, Clark LA, Tellegen A (1988) Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Personal Soc Psychol* 54(6):1063
47. Zotov V, Jones MN, Mewhort DJK (2011) Contrast and assimilation in categorization and exemplar production. *Atten Percept Psychophys* 73(2):621–639. doi:[10.3758/s13414-010-0036-z](https://doi.org/10.3758/s13414-010-0036-z)