Paul van Schaik[1]

Matthew Unwin[1]

David Eves[2]

Melissa Lara[3]

Alfred Mensah[1]

[1]Teesside University, [2]amBX UK Ltd, [3]McLaren Technology Group

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

To appear in *Quality and User Experience*

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

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 (up to 6): user-experience evaluation; magnitude-based inference; programmable surround lighting; media content; data analysis; comparative testing

Table of contents

1 Introduction

Over the past 20 years, user-experience (UX) has become an influential concept (Hassenzahl & Tractinsky, 2006). UX and its enhancement have been studied predominantly by targeting the design of products (Hekkert & Schifferstein, 2008). 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.

1.1 Affect and UX

Emotion, and more generally affect, plays a central role in UX. For example, in Thüring and Mahlke's (2007) 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 (Russell, 1980, 2003). 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 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 been 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 (Watson et al., 1988; MacKinnon et al., 1999). 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.

1.2 Framework for UX design and evaluation

Kurosu (2017) 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”, Kurosu, 2017, p. 32] and functionality [“the ability of an artifact to support the user functionally to achieve the goal”, Kurosu, 2017, p. 32]). We can also design subjective quality (attractiveness: appeal for needs [e.g., stimulation: the capability to meet the human need for stimulation; Sheldon, Elliot, Kim & Kasser, 2001] 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; Kurosu, 2017]). We can then evaluate objective quality in use (e.g., achieved usability as effectiveness and efficiency [International Standard Organization, 2011]). 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).

1.3 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 (Kurosu, 2017) 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 (amBX, 2010-2016) or in augmenting digital media (Cole & Eves, 2003). 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 (amBX, 2014). The research presented here has used this innovative solution, documented by various patents (Eves & Cole [2000], covering physical mark-up language, and Eves & Cole [2001], 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 (Eves & Cole, 2013). 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 (Neumeister, Goessler, Lucht, Kapitany, Bamas & Kasper, 1996). 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 (2000) 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 adrenalin. 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.

1.4 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 and 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., Carvalho, Leite, Galdo-Álvarez & Gonçalves, 2012). 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 unwieldly to measure affect, with several different ratings per video clip in one test session (Bradley & Lang, 1994). 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 (Watson et al., 1988; MacKinnon et al., 1999).

Inference in UX evaluation. An important aspect of UX evaluation is comparative testing of designs (Sauro & Lewis, 2012). 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 (Cumming, 2012; Murphy & Myors, 1999). Magnitude-based inference offers an attractive alternative (Buchheit, 2016; Hopkins & Batterham, 2016) and its use in UX evaluation, and user research more generally, has also been advocated and demonstrated (van Schaik & Weston, 2016). In the current study, we chose to use magnitude-based inference in UX evaluation, as it has several advantages (van Schaik & Weston, 2016). First, the approach requires the researcher to define a smallest important effect, instead of testing the 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 Wagenmakers et al., 2011). 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.

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 …

Note. Interpretation according to Batterham and Hopkins (2006).

1.5 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 (see Section 2 below) and (2) test the benefits of surround lighting (see Section 3 below). 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?

2 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 in Section 1.3).

2.1 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).

2.2 Participants

Participants were recruited by e-mail and word of mouth, and received 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]

2.3 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 (see Section 2.4) 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

---

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

that represented the displeasure dimension of negative affective quality evoking the most negative affect. From 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 (practice or main) series, 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 within 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).

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 *yellow* and 6 with white lighting, presented in quasi-random order	12 clips, positive affective quality, 6 with *green* and 6 with white lighting, presented in quasi-random order
M2	12 clips, positive affective quality, 6 with *yellow* and 6 with white lighting, presented in quasi-random order	12 clips, positive affective quality, 6 with *green* and 6 with white lighting, presented in quasi-random order
M3	12 clips, positive affective quality, 6 with *green* and 6 with white lighting, presented in quasi-random order	12 clips, positive affective quality, 6 with *yellow* and 6 with white lighting, presented in quasi-random order
M4	12 clips, positive affective quality, 6 with *green* and 6 with white lighting, presented in quasi-random order	12 clips, positive affective quality, 6 with *yellow* and 6 with white lighting, presented in quasi-random order
M5	12 clips, negative affective quality, 6 with *red* and 6 with white lighting, presented in quasi-random order	12 clips, negative affective quality, 6 with *purple* and 6 with white lighting, presented in quasi-random order
M6	12 clips, negative affective quality, 6 with *red* and 6 with white lighting, presented in quasi-random order	12 clips, negative affective quality, 6 with *purple* and 6 with white lighting, presented in quasi-random order
M7	12 clips, negative affective quality, 6 with *purple* and 6 with white lighting, presented in quasi-random order	12 clips, negative affective quality, 6 with *red* and 6 with white lighting, presented in quasi-random order
M8	12 clips, negative affective quality, 6 with *purple* and 6 with white lighting, presented in quasi-random order	12 clips, negative affective quality, 6 with *red* and 6 with white lighting, presented in quasi-random order

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

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 (153cm x 122cm). 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 in Section 1.3. 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 (Bradley & Lang, 1994; Watson et al., 1988) to enhance the sensitivity of measurement.

2.4 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 seconds. Then they were given 15 seconds 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.

3 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 means[2]) was used.[3] Inferential statistics were conducted through related *t* tests with SPSS, followed by magnitude-based inference[4] (see Section 1.3) (using the SPSS output as input) with dedicated spreadsheets (Hopkins, 2007). This was followed by generalised linear mixed-model analysis with SPSS to verify the consistency of results.

3.1 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’; Monk, 2004). 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.

3.1.1 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.

---

[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 (Sauro & Lewis, 2012).

[4] In magnitude-based inference, results are here presented for a small effect (according to Cohen’s [1988] 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	

Note . (Y): white within alternating series of yellow and white surround lighting. (G): white within alternating series of green and white surround lighting.

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 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, means remained unchanged, but 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 (2004) 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) (Heck et al., 2010). 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 in Section 3.2). 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

a. Data averaged over video clips

	Comparison	Mean	*SD*	90%-confidence interval of mean diff.	*d*	Chances			Odds ratio (B/H)	Inference
						Beneficial/+ive	Negligible/trivial	Harmful/-ive		
1	Yellow-	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
	white	6.72	1.05			very unlikely	very likely	most unlikely		P: very likely trivial; don't use
2	Green-	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
	white	6.32	1.77			most unlikely	very likely	very unlikely		P: very likely trivial; don't use
3	Yellow-	6.77	1.24	0.52, 0.084 to 0.95	0.31	76.02 %	23.87 %	0.11 %	2875.35	M: likely +ive
	green	6.25	2.02			likely	unlikely	most unlikely		P: likely beneficial; use
4	White(Y)-	6.72	1.05	0.4, 0.04 to 0.76	0.27	69.29 %	30.57 %	0.13 %	1680.99	M: possibly +ive
	white(G)	6.32	1.77			possibly	possibly	most unlikely		P: possibly beneficial; use
5	Y/W(Y)-	6.75	1.12	0.46, 0.07 to 0.85	0.30	74.29 %	25.61 %	0.11 %	2703.18	M: possibly +ive
	G/W(G)	6.29	1.87			possibly	possibly	most unlikely		P: possibly beneficial; use

b. Data averaged over participants

	Comparison	Mean	*SD*	90%-confidence interval of mean diff.	*d*	Chances			Odds ratio (B/H)	Inference
						Beneficial/+ive	Negligible/trivial	Harmful/-ive		
1	Yellow-	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
	white	6.72	0.70			possibly	possibly	unlikely		P: unclear; don't use; get more data
2	Green-	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
	white	6.32	0.72			very unlikely	very likely	very unlikely		P: very likely trivial; don't use
3	Yellow-	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
	green	6.25	0.60			very likely	very unlikely	most unlikely		P: very likely beneficial; use
4	White(Y)-	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
	white(G)	6.32	0.72			very likely	very unlikely	most unlikely		P: very likely beneficial; use
5	Y/W(Y)-	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
	G/W(G)	6.29	0.62			most likely	most unlikely	most unlikely		P: most likely beneficial; use

Note . Thresholds for important effect size *d* = ±0.2. *d* : observed effect size. B: benefit. H: harm. M: mechanistic inference. P: practical inference.

3.1.2 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, means remained unchanged, but 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 and for the alternation of red (*d* = 0.69) 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.

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	

Note . (R): white within alternating series of yellow and white surround lighting. (P): white within alternating series of green and white surround lighting.

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.

Table 6

Magnitude-based inference for negative affect

a. Data averaged over video clips

Comparison		Mean	*SD*	90%-confidence interval of mean diff.	*d*	Chances Beneficial/+ive	Chances Negligible/trivial	Chances Harmful/-ive	Odds ratio (B/H)	Inference
1	Red-	6.12	1.64	0.35, 0.12 to 0.57	0.25	69.90 %	30.10 %	0.00 %	97560.37	M: possibly +ive
	white	5.77	1.39			possibly	possibly	most unlikely		P: possibly beneficial; use
2	Purple-	5.74	1.69	0.15, -0.088 to 0.4	0.10	13.26 %	86.61 %	0.12 %	125.83	M: likely trivial
	white	5.58	1.58			unlikely	likely	most unlikely		P: likely trivial; don't use
3	Red-	6.12	1.64	0.38, 0.13 to 0.63	0.23	62.74 %	37.25 %	0.00 %	132791.09	M: possibly +ive
	purple	5.74	1.69			possibly	possibly	most unlikely		P: possibly beneficial; use
4	White(R)-	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
	white(P)	5.58	1.58			unlikely	likely	most unlikely		P: likely trivial; don't use
5	R/W(R)-	5.94	1.47	0.28, 0.074 to 0.49	0.19	43.05 %	56.95 %	0.00 %	38333.92	M: possibly +ive
	P/W(P)	5.66	1.59			possibly	possibly	most unlikely		P: possibly beneficial; use

b. Data averaged over participants

Comparison		Mean	*SD*	90%-confidence interval of mean diff.	*d*	Chances Beneficial/+ive	Chances Negligible/trivial	Chances Harmful/-ive	Odds ratio (B/H)	Inference
1	Red-	6.12	0.47	0.35, 0.067 to 0.63	0.69	92.84 %	6.39 %	0.77 %	1671.22	M: likely +ive
	white	5.77	0.53			likely	unlikely	very unlikely		P: unclear; don't use; get more data
2	Purple-	5.74	0.46	0.15, -0.069 to 0.38	0.30	65.32 %	31.58 %	3.10 %	58.97	M: possibly +ive
	white	5.58	0.57			possibly	possibly	very unlikely		P: unclear; don't use; get more data
3	Red-	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
	purple	5.74	0.46			very likely	very unlikely	very unlikely		P: unclear; don't use; get more data
4	White(R)-	5.77	0.53	0.19, 0.00084 to 0.37	0.34	76.19 %	23.05 %	0.76 %	416.93	M: likely +ive
	white(P)	5.58	0.57			likely	unlikely	very unlikely		P: unclear; don't use; get more data
5	R/W(R)-	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
	P/W(P)	5.66	0.47			very likely	very unlikely	most unlikely		P: very likely beneficial; use

Note . Thresholds for important effect size *d* = ±0.2. *d* : observed effect size. B: benefit. H: harm. M: mechanistic inference. P: practical inference.

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) (Heck et al., 2010). 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 in Section 3.2). In sum, the pattern of the results of mixed-modelling matches that of magnitude-based inference (presented in the previous subsection).

3.2 Comparisons by clip

In the main series, there were 12 clips with positive affective quality and 12 clips with negative affective 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 time, 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 (*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, y*ellow versus green*. In the comparison of yellow and green (Table 7; Figure 1), there were *likely positive/beneficial* small to moderate effects of

---

[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.

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 an 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; Figure 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

---

[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.

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.

4 Discussion

4.1 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.

---

[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.

Table 7

Descriptives for positive affect per positive clip

	Yellow		Green	
Clip	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

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

Table 8

Descriptives for negative affect per negative clip

	Red		Purple	
Clip	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

Clip	Yellow	Green	*d*
1	6.76	6.36	*0.17*
2	6.64		
3	6.94	6.27	*0.14*
4	5.39	4.85	*0.19*
5	7.06	6.09	*0.35*
6	6.82	6.03	*0.32*
7	7.52	7.33	*0.10*
8	7.39	6.33	*0.40*
9	5.82	6.03	*-0.08*
10	6.79	6.70	*0.03*
11	7.27	6.18	*0.39*
12	6.70	6.58	*0.05*

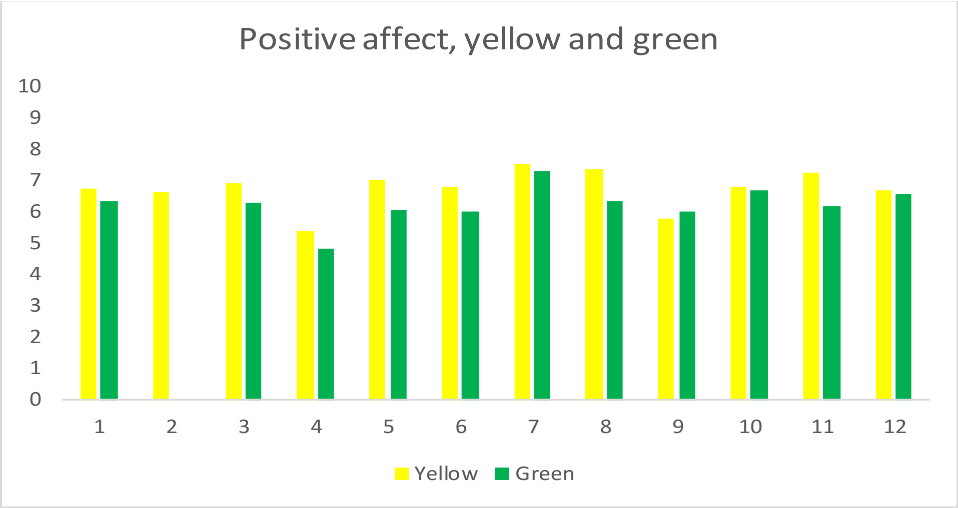


Figure 1 . Mean ratings and effect size of positive affect per positive clip.

Clip	Red	Purple	*d*
1	6.00	5.15	*0.29*
2	5.88	5.60	*0.12*
3	5.88	6.09	*-0.09*
4	6.18	5.48	*0.26*
5	5.36	6.15	*-0.31*
6	5.42	4.81	*0.21*
7	6.45	6.06	*0.19*
8	7.00	5.76	*0.47*
9	5.91	5.85	*0.02*
10	6.33	5.39	*0.42*
11	6.39	6.29	*0.04*
12	6.61	6.21	*0.14*

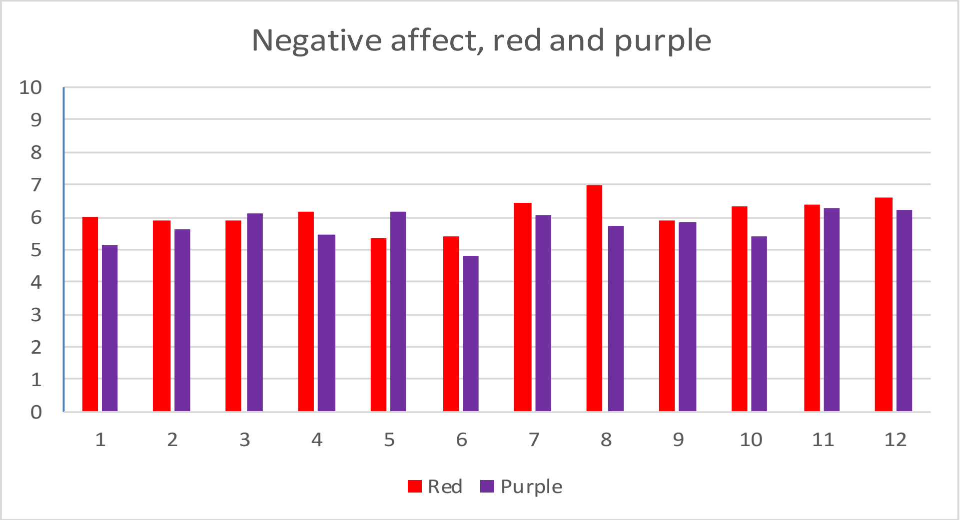


Figure 2 . Mean ratings and effect size of negative affect per negative clip.

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 (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 (see Section 3.2), we derived 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.

4.2 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

---

[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.

content (and media content more generally), assimilation effects are presumably inevitable.[10] Further evidence comes from various research studies published in the psychology literature. For example, in a review of perception research, Lockhead (2004) 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 (2013) 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 (2011) 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., Brown, Venkatesh & Goyal, 2014). 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 (Cypryanska,

---

[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.

Bedynska & Golec De Zavala, 2012), categorization and exemplar production (Zotov, Jones & Mewhort, 2012), and the visual perception of natural scenes (Howe & Purves, 2004).

4.3 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 (Gross & Levenson, 1995). 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.

4.4 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 (2017) 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 (Golden et al., 2005). 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 (Figueiro et al., 2014) and reduces cognitive decline (Van Hoof et al., 2009). Furthermore, pain is strongly linked to depression, and either condition can be cause or effect of the other (Gerrits et al., 2012). 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 (Canazei et al., 2014), as an alternative or adjunct to pharmacological (Baldwin et al., 2005) and psychological interventions (Smits et al., 2008).

5 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.

6 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 (2017) 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

amBX. (2014). The amBX light-scene engine for commercial lighting. Retrieved from *http://www.ambx.com/ambx-light-scene-engine-commercial-lighting*

amBX. (2016). Case studies (2010-2016). Retrieved from *http://www.ambx.com/case-studies*

Baldwin, D. S., Anderson, I. M., Nutt, D. J., Bandelow, B., Bond, A., Davidson, J. R., ... & Wittchen, H. U. (2005). Evidence-based guidelines for the pharmacological treatment of anxiety disorders: recommendations from the British Association for Psychopharmacology. *Journal of Psychopharmacology, 19*(6), 567-596.

Batterham, A. M., & Hopkins, W. G. (2006). Making meaningful inferences about magnitudes. *International Journal of Sports Physiology and Performance, 1*(1), 50-57.

Bradley and Lang (1994) – Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59.

Brown, S. A., Venkatesh, V., & Goyal, S. (2014). Expectation confirmation in information systems research: A test of six competing models. *MIS Quarterly, 38*(3), 729-A9.

Buchheit, M. (2016). The numbers will love you back in return-I promise. *International Journal of Sports Physiology and Performance, 11*(4), 551-554. doi:10.1123/IJSPP.2016-0214

Canazei, M., Dehoff, P., Staggl, S., & Pohl, W. (2014). Effects of dynamic ambient lighting on female permanent morning shift workers. *Lighting Research & Technology, 46*(2), 140-156.

Carvalho, S., Leite, J., Galdo-Álvarez, S., & Gonçalves, O. F. (2012). The emotional movie database (EMDB): A self-report and psychophysiological study. *Applied Psychophysiology Biofeedback, 37*(4), 279-294. doi:10.1007/s10484-012-9201-6

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cole, R., & Eves, D. (2003). PML: When your room becomes your browser. In E. Aarts, & S. Marzano (Eds.), *The new everyday: Views on ambient intelligence* (pp. 312-313). Rotterdam: 010 Publishers.

Cypryanska, M., Bedynska, S., & Golec De Zavala, A. (2012). The role of similarity and dissimilarity focus in sequential judgments of physical attractiveness. *Journal of Social Psychology, 152*(3), 266-269. doi:10.1080/00224545.2011.609846

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*(1), 7-29. doi:10.1177/0956797613504966

Eves, D., & Cole, R. (2000). *Physical mark-up language*. Patent No. WO 02/092183. Switzerland: World Intellectual Property Organisation.

Eves, D., & Cole, R. (2001). *Dynamic mark-up language*. Patent No. WO 03/100548.Switzerland: World Intellectual Property Organisation.

Eves, D., & Cole, R. (2013). *Unique intelligent software for digital lighting control* (white paper). Redhill, United Kingdom: amBX.

Figueiro, M. G., Hunter, C. M., Higgins, P. A., Hornick, T. R., Jones, G. E., Plitnick, B., ... & Rea, M. S. (2015). Tailored lighting intervention for persons with dementia and caregivers living at home. *Sleep Health, 1*(4), 322-330.

Garnett, O. (2000). *Colour: a social history*. Manvers, Rotherham (UK): National Trust.

Gerrits, M. M., Vogelzangs, N., Van Oppen, P., Van Marwijk, H. W., van der Horst, H., & Penninx, B. W. (2012). Impact of pain on the course of depressive and anxiety disorders. *Pain, 153*(2), 429-436

Golden, R. N., Gaynes, B. N., Ekstrom, R. D., Hamer, R. M., Jacobsen, F. M., Suppes, T., ... & Nemeroff, C. B. (2005). The efficacy of light therapy in the treatment of mood disorders: a review and meta-analysis of the evidence. *American Journal of Psychiatry, 162*(4), 656-662.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & Emotion*, *9*(1), 87-108.

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour and Information Technology, 25*(2), 91-97. doi:10.1080/01449290500330331

Heck, R. H., Thomas, S. L., & Tabata, L. N. (2010). *Multilevel and longitudinal modeling with IBM SPSS*. London: Routledge Academic.

Hekkert, P., & Schifferstein, H. (2008). *Product experience*. Amsterdam: Elsevier.

Hoof, J. van, Schoutens, A. M. C., & Aarts, M. P. J. (2009). High colour temperature lighting for institutionalised older people with dementia. *Building and Environment, 44*(9), 1959-1969.

Hopkins, W. G. (2007). A spreadsheet for deriving a confidence interval, mechanistic inference and clinical inference from a *p*-value. *Sport Science, 11*, 16-20.

Hopkins, W. G., & Batterham, A. M. (2016). Error rates, decisive outcomes and publication bias with several inferential methods. *Sports Medicine, 46*(10), 1563-1573. doi:10.1007/s40279-016-0517-x

Howe, C. Q., & Purves, D. (2004). Size contrast and assimilation explained by the statistics of natural scene geometry. *Journal of Cognitive Neuroscience, 16*(1), 90-102. doi:10.1162/089892904322755584

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*. Switzerland: Authors.

Kondo, A., Takahashi, K., & Watanabe, K. (2013). Influence of gender membership on sequential decisions of face attractiveness. *Attention, Perception, and Psychophysics, 75*(7), 1347-1352. doi:10.3758/s13414-013-0533-y

Kurosu, M. (2017). *Theory of user engineering*. CRC Press.

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. *Journal of Experimental Psychology: Human Perception and Performance, 37*(6), 1874-1886. doi:10.1037/a0025589

Lockhead, G. R. (2004). Absolute judgments are relative: A reinterpretation of some psychophysical ideas. *Review of General Psychology, 8*(4), 265-272. doi:10.1037/1089-2680.8.4.265

Monk, A. (2004). The product as a fixed-effect fallacy. *Human-Computer Interaction, 19*(4), 371-375.

Mackinnon, A., Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P. A., & 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. *Personality and Individual Differences, 27*(3), 405-416. doi:10.1016/S0191-8869(98)00251-7

Murphy, K. R., & Myors, B. (1999). Testing the hypothesis that treatments have negligible effects: Minimum-effect tests in the general linear model. *Journal of Applied Psychology, 84*(2), 234-248. doi:10.1037/0021-9010.84.2.234

Neumeister, A, Goessler, R., Lucht, M., Kapitany, T., Bamas, C. & Kasper, S. (1996). Bright light therapy stabilizes the antidepressant effect of partial sleep deprivation. *Biological Psychiatry, 39*, 16–21.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*, 145-172.

Sauro, J., & Lewis, J. R. (2012). *Quantifying the user experience*. Amsterdam: Morgan Kaufmann.

Schaik, P., van & Weston, M. (2016). Magnitude-based inference and its application in user research. *International Journal of Human Computer Studies, 88*, 38-50. doi:10.1016/j.ijhcs.2016.01.002

Sheldon, K. M., Elliot, A. J., Kim, Y., & Kasser, T. (2001). What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of Personality and Social Psychology, 80*(2), 325-339. doi:10.1037/0022-3514.80.2.325

Smits, J. A., Berry, A. C., Tart, C. D., & Powers, M. B. (2008). The efficacy of cognitive-behavioral interventions for reducing anxiety sensitivity: A meta-analytic review. *Behaviour Research and Therapy, 46*(9), 1047-1054.

Thüring, M., & Mahlke, S. (2007). Usability, aesthetics and emotions in human-technology interaction. *International Journal of Psychology, 42*(4), 253-264. doi:10.1080/00207590701396674

Wagenmakers, E., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology, 100*(3), 426-432. doi:10.1037/a0022790

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, *54*(6), 1063.

Zotov, V., Jones, M. N., & Mewhort, D. J. K. (2011). Contrast and assimilation in categorization and exemplar production. *Attention, Perception, and Psychophysics, 73*(2), 621-639. doi:10.3758/s13414-010-0036-z

Online Appendix A1 – video clips used in the study

Table A1.1

Details of video clips

Emotion	P/M	Code	Video content name	Clip name	Video quality
Negative					
	Practice Trial				
		1/A	Anonymous X - Marks the Spot in Melbourne 2014-HD	kids in dark	good
		2/B	Hammerhead Shark Attacks Kayakers!!!	shark 1	good
		3/C	Crescent Foods Halal Chicken at Processing Plant	chicken factory	good
		4/D	Lamma Oil Spill 2013	oil spill 3	good
		5/E	Resident Aliens Episode IV- Agents of Decay, Part 2 (Flies, Maggots, Weevils, Wasps)	ants on fish	good
		6/F	James's Dirty Dishes- A Documentary	dirty dishes	good

(Negative) Main Trial

1/A	Dung Beetles - Underground army, enriching soils (7min)	dung beetles	good
2/B	American Ubiquity (2012)-SD	plastic bag lq	good
3/C	Anonymous x 2014 Homeless Trailer -HD.mp4	clothes sorting	good
4/D	Mandaluyong Animal Shelter dogs. Please adopt. Don't buy.	dogs cages negative	medium
5/E	Graphic HD Video- Marines in combat firefight against enemy in Afghanistan (1)	marine gun 1	good
6/F	Resident Aliens Episode IV- Agents of Decay, Part 2 (Flies, Maggots, Weevils, Wasps)	flies	good
7/G	Four Corners Video-HD	plastic bag 2	good
8/H	The 3 Causes of Traffic Jams, A Rant	traffic pic	good
9/I	motorbike crash	Z trim motorcycle neg	medium
10/J	Power - Electricity Generation through Bio Gas (Cow dung) Plant - Digester at PAU	man in poo	good
11/K	Homeless in Hollywood-HD	homeless	good
12/L	gorakhpur express-HD	train crash	poor

Positive

Practice Trial

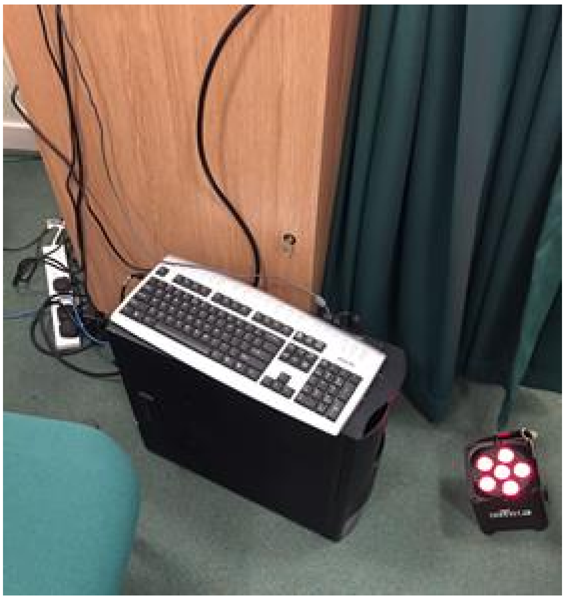
1/A	Sicily 2014-HD.mp4	water party	good
2/B	BEAN TIME LAPSE	bean time lapse	good
3/C	SAN LUCA BRANCA (Potenza) - PIROTECNICA MODERNA di PADOVANO Giovanni (Piromusicale - 2014)-HD	fireworks 2	good
4/D	WHEN IN WOMAD-HD	music festival	good
5/E	Occupy Austin trip to Occupy Congress J17	friends	good
6/F	Dune%20Jumping-HD	dune jump	good

(Positive) Main Trial

1/A	cheerleaders - preview	cheerleader trick	good
2/B	Mylne%20Bolt%2018%20-%20Electric%20Performance-HD	speed boat	good
3/C	Head Shoulders Knees and toes and kissing hugging baby in mirror	baby laughing in mirror	good
4/D	we love funny faces	funny faces	good
5/E	'Puppies on Slides Compilation - PART 2' – CFS	dog on slide	good
6/F	Mango Season Records Debut Island Soul Jazz CD!-HD	love	good
7/G	Time Lapse Sunrise (HD 720p)	sunrise	good
8/H	Baby and his puppies	Baby kisses dog	good
9/I	Tropical Breeze Hybrid Bowling Ball Reaction Video Review	bowling	good
10/J	Funny Dogs Playing Sports Compilation 2014 funny dog videos	dog plays drums	good
11/K	Laughing baby playing with dog and water	baby plays with dog	good
12/L	Around the Alps-HD	night sky	good

Note. Source: Creative Commons (*http://search.creativecommons.org* and *http://vimeo.com/creativecommons*).

Online Appendix A2 – experiment set-up



Online Appendix A3 – generalised linear mixed-model analysis for positive or negative affect

Table A3.1

Generalised linear mixed-model analysis, positive affect

Comparison			Source	Coefficient	*SE*	*t* (788)	*p*	*OR*	*CI* (*OR* , 95%)		Cumulative *OR* across response scale
									Lower limit	Upper Limit	
Yellow	-	White									
			Intercept	-0.289	0.443	-0.653	0.514	0.749	0.314	1.786	
			Positive affect, Level 1 (per clip)	0.000	0.033	0.009	0.993	1.000	0.937	1.068	1.000
			Positive affect, Level 2 (over clips)	0.048	0.071	0.682	0.495	1.050	0.913	1.206	1.629
Green	-	White									
			Intercept	0.279	0.284	0.985	0.325	1.322	0.758	2.308	
			Positive affect, Level 1 (per clip)	-0.001	0.037	-0.021	0.983	0.999	0.928	1.075	0.990
			Positive affect, Level 2 (over clips)	-0.017	0.054	-0.310	0.757	0.983	0.884	1.094	0.842
Yellow	-	Green									
			Intercept	-1.185	0.294	-4.034	0.000	0.306	0.172	0.544	
			Positive affect, Level 1 (per clip)	0.000	0.036	0.002	0.998	1.000	0.932	1.074	1.000
			Positive affect, Level 2 (over clips)	0.192	0.055	3.484	0.001	1.211	1.087	1.350	6.783
White (yellow)	-	White (green)									
			Intercept	-1.009	0.332	-3.305	0.002	0.365	0.190	0.700	
			Positive affect, Level 1 (per clip)	0.001	0.034	0.025	0.980	1.001	0.936	1.070	1.010
			Positive affect, Level 2 (over clips)	0.183	0.058	3.139	0.002	1.201	1.071	1.347	6.244
Yellow and white	-	Green and white									
			Intercept	-1.243	0.327	-3.801	0.000	0.288	0.152	0.548	
			Positive affect, Level 1 (per clip)	0.000	0.041	-0.008	0.994	1.000	0.923	1.083	1.000
			Positive affect, Level 2 (over clips)	1.960	0.060	3.261	0.001	1.217	1.081	1.370	7.127

Note. Model: comparison = intercept + uniform distribution + positive affect (Level 1) + positive affect (Level 2)
The target distribution and relationship (link function) with the model was binary logistic regression.
The uniform distribution was added to the model, as without this addition the results for some of the comparisons were incomplete (either the Level-1 predictor or Level-2 predictor was not included in the results). Note that this was not due to collinearity, as the Level-1 predictor and the Level-2 predictor were not highly correlated for any of the comparisons.
The coefficient for the uniform distribution as a predictor was always 0.000 and this result is therefore not presented in the table.
The analysis did not include subject (participant) as a random effect. This is because the finding of a significant random effect of subject is expected and was not of interest.
For the tests of the option of robust (rather than model-based) covariances was used to handle violations of model assumptions.
The residual method was used for degrees of freedom (the Satterthwaite correction produced identical test results).

Table A3.2

Generalised linear mixed-model analysis, negative affect

Comparison			Source	Coefficient	*SE*	*t* (788)	*p*	*OR*	C I(*OR* , 95%)		Cumulative *OR* across response scale
									Lower limit	Upper Limit	
Red	-	White									
			Intercept	-0.890	0.321	2.775	0.006	0.411	0.219	0.771	
			Negative affect, Level 1 (per clip)	0.000	0.035	0.010	0.992	1.000	0.933	1.071	1.000
			Negative affect, Level 2 (over clips)	0.156	0.059	2.627	0.009	1.168	1.040	1.312	4.725
Purple	-	White									
			Intercept	-0.206	0.291	-0.706	0.480	0.814	0.459	1.442	
			Negative affect, Level 1 (per clip)	0.002	0.036	0.047	0.962	1.002	0.933	1.076	1.020
			Negative affect, Level 2 (over clips)	0.056	0.057	0.985	0.325	1.058	0.946	1.183	1.757
Red	-	Purple									
			Intercept	-0.858	0.297	-2.884	0.004	0.424	0.237	0.760	
			Negative affect, Level 1 (per clip)	0.000	0.036	-0.003	0.998	1.000	0.932	1.073	1.000
			Negative affect, Level 2 (over clips)	0.142	0.057	2.504	0.012	1.152	1.031	1.287	4.116
White (red)	-	White (purple)									
			Intercept	-0.416	0.310	-1.341	0.180	0.660	0.359	1.213	
			Negative affect, Level 1 (per clip)	0.000	0.035	-0.007	0.995	1.000	0.933	1.071	1.000
			Negative affect, Level 2 (over clips)	0.087	0.060	1.451	0.147	1.091	0.970	1.227	2.389
Red and white	-	Purple and white									
			Intercept	-0.783	0.322	-2.435	0.015	0.457	0.243	0.859	
			Negative affect, Level 1 (per clip)	0.000	0.042	0.002	0.998	1.000	0.920	1.087	1.000
			Negative affect, Level 2 (over clips)	0.126	0.064	1.981	0.048	1.135	1.001	1.286	3.548

Note. Model: comparison = intercept + uniform distribution + negative affect (Level 1) + negative affect (Level 2)
The target distribution and relationship (link function) with the model was binary logistic regression.
The uniform distribution was added to the model, as without this addition the results for some of the comparisons were incomplete (either the Level-1 predictor or Level-2 predictor was not included in the results). Note that this was not due to collinearity, as the Level-1 predictor and the Level-2 predictor were not highly correlated for any of the comparisons.
The coefficient for the uniform distribution as a predictor was always 0.000 and this result is therefore not presented in the table.
The analysis did not include subject (participant) as a random effect. This is because the finding of a significant random effect of subject is expected and was not of interest.
For the tests of the option of robust (rather than model-based) covariances was used to handle violations of model assumptions.
The residual method was used for degrees of freedom (the Satterthwaite correction produced identical test results).

Online Appendix A4 – magnitude-based inference for positive or negative affect per video clip

Table A4.1

Magnitude-based inference for positive affect (individual clips): yellow versus green surround lighting

Clip	Comparison	Mean	*SD*	90%-confidence interval of mean diff.	*d*	Chances that the true value of the effect statistic is …			Odds ratio (benefit/harm)
						beneficial or substantially +ive	negligible or trivial	harmful or substantially -ive	
1	Yellow-	6.76	2.17	0.39, -0.089 to 0.88	0.17	40.26 %	59.48 %	0.25 %	266.47707
	green	6.36	2.47			possibly	possibly	most unlikely	
3	Yellow-	6.64	2.28	0.36, -0.61 to 1.3	0.14	39.96 %	53.13 %	6.91 %	8.96467
	green	6.27	2.81			possibly	possibly	unlikely	
4	Yellow-	5.39	2.65	0.55, -0.12 to 1.2	0.19	48.30 %	51.28 %	0.42 %	221.79719
	green	4.85	2.97			possibly	possibly	most unlikely	
5	Yellow-	7.06	2.46	0.97, 0.18 to 1.8	0.35	80.99 %	18.89 %	0.12 %	3514.78685
	green	6.09	3.08			likely	unlikely	most unlikely	
6	Yellow-	6.82	2.35	0.79, 0.21 to 1.4	0.32	80.10 %	19.86 %	0.04 %	11472.05610
	green	6.03	2.59			likely	unlikely	most unlikely	
7	Yellow-	7.52	1.70	0.18, -0.38 to 0.75	0.10	30.69 %	63.37 %	5.94 %	7.00762
	green	7.33	1.81			possibly	possibly	unlikely	
8	Yellow-	7.39	2.40	1.1, 0.31 to 1.8	0.40	88.02 %	11.92 %	0.06 %	12398.84560
	green	6.33	2.85			likely	unlikely	most unlikely	
9	Yellow-	5.82	2.64	-0.21, -0.98 to 0.55	-0.08	5.70 %	69.41 %	24.89 %	0.18241
	green	6.03	2.58			unlikely	possibly	unlikely	
10	Yellow-	6.79	2.53	0.091, -0.58 to 0.76	0.03	12.16 %	82.39 %	5.45 %	2.40463
	green	6.70	3.05			unlikely	likely	unlikely	
11	Yellow-	7.27	2.39	1.1, 0.31 to 1.9	0.39	88.17 %	11.77 %	0.06 %	12072.52388
	green	6.18	2.95			likely	unlikely	most unlikely	
12	Yellow-	6.70	2.52	0.12, -0.87 to 1.1	0.05	24.01 %	62.64 %	13.35 %	2.05119
	green	6.58	2.85			unlikely	possibly	unlikely	

Note. Thresholds for important effect size *d* = ±0.2. *d* : observed effect size.

Table A4.2

Magnitude-based inference for negative affect (individual clips): red versus purple surround lighting

Clip	Comparison	Mean	*SD*	90%-confidence interval of mean diff.	*d*	Chances that the true value of the effect statistic is …			Odds ratio (benefit/harm)
						beneficial or substantially +ive	negligible or trivial	harmful or substantially -ive	
1	Red-	6.00	2.99	0.85, 0.14 to 1.5	0.29	73.61 %	26.31 %	0.08 %	3387.63567
	purple	5.15	2.81			possibly	possibly	most unlikely	
2	Red-	5.88	2.29	0.27, -0.51 to 1.1	0.12	34.69 %	59.21 %	6.10 %	8.17997
	purple	5.60	2.29			possibly	possibly	unlikely	
3	Red-	5.88	2.03	-0.21, -0.88 to 0.46	-0.09	5.00 %	68.09 %	26.91 %	0.14298
	purple	6.09	2.53			unlikely	possibly	possibly	
4	Red-	6.18	2.83	0.7, -0.046 to 1.4	0.26	63.35 %	36.26 %	0.39 %	437.57866
	purple	5.48	2.62			possibly	possibly	most unlikely	
5	Red-	5.36	2.61	-0.79, -1.7 to 0.087	-0.31	0.85 %	29.01 %	70.14 %	0.00365
	purple	6.15	2.51			very unlikely	possibly	possibly	
6	Red-	5.42	3.03	0.61, -0.037 to 1.3	0.21	53.28 %	46.53 %	0.19 %	586.35423
	purple	4.81	2.79			possibly	possibly	most unlikely	
7	Red-	6.45	1.86	0.39, -0.0026 to 0.79	0.19	47.92 %	51.99 %	0.09 %	1058.34365
	purple	6.06	2.19			possibly	possibly	most unlikely	
8	Red-	7.00	2.49	1.2, 0.45 to 2	0.47	93.25 %	6.72 %	0.03 %	43926.38003
	purple	5.76	2.77			likely	unlikely	most unlikely	
9	Red-	5.91	3.20	0.061, -0.43 to 0.55	0.02	3.30 %	95.35 %	1.35 %	2.49960
	purple	5.85	2.96			very unlikely	very likely	very unlikely	
10	Red-	6.33	2.25	0.95, 0.38 to 1.5	0.42	92.68 %	7.31 %	0.01 %	108390.09470
	purple	5.39	2.22			likely	unlikely	most unlikely	
11	Red-	6.39	2.25	0.1, -0.77 to 0.98	0.04	24.13 %	62.23 %	13.64 %	2.01442
	purple	6.29	2.45			unlikely	possibly	unlikely	
12	Red-	6.61	2.97	0.39, -0.42 to 1.2	0.14	36.91 %	60.29 %	2.80 %	20.32443
	purple	6.21	2.57			possibly	possibly	very unlikely	

Note. Thresholds for important effect size *d* = ±0.2. *d*: observed effect size.