Analysis of Research Strategies to Determine Individual Colour Preference: N-Alternative Forced Choice, Rank-Order, Rating and Paired Comparison

Luwen Yu,[1] Guobin Xia,*[1, 2] Stephen Westland,[1] Zhenhong Li,[3] Vien Cheung[1]

[1]School of Design, University of Leeds, Leeds, UK

[2]School of Textile and Design, Herriot-Watt University, UK

[3]School of Electronic and Electrical Engineering, University of Leeds, Leeds, UK

Abstract:
Exploring an efficient research method to understanding colour preference is important to researchers and designers. This work compares four experimental methods for individual colour preference research (N-alternative forced choice, rank-order, rating and paired comparison). Three psychophysical experiments have been carried out with 338 participants. Participants were presented with six colour patches (red, orange, yellow, green, blue and purple) arranged in a random order. This work suggested orange is the strongest preferred colour and green is the weakest preferred in three individual colour preference experimental methods with six hues. The Monte Carlo Analysis method further compares the result performance for four methods, which suggests the rating, paired comparison and rank-order methods are more stable than the N-alternative forced choice method when only small number participants take part in the experiment, such as for studies involving small numbers of participants (even less than 6), the rating, rank-order and pair comparison methods should be preferred.

Keywords: Colour Preferences, Experimental Method, Monte Carlo Analysis, Research Strategy.

INTRODUCTION

Colour preference *per se* has been studied by many researchers[1]. It is concerned with which colours individual prefer among alternatives[2]. Individual colour preferences vary from person to person[3]. Thus, finding an efficient way to understand individual colour preference is important to both researchers[4] and designers[5,6]. The previous studies have used standardised colours, multiple research strategies, and sophisticated statistical methods[7]. From the previous studies, a colour preference research strategy includes experimental material, experimental method and experimental environment[7]. The experimental methods could be classified as N-alternative forced choice (N-AFC, N≥2), rank-order, subjective rating, affective judgments, description method, physical, behavioural measurements etc.[7] In N-alternative forced-choice method, participants indicate which colour they 'prefer aesthetically' with all colours simultaneously presented on a visual display[8]. Additionally, the 2-AFC is also called 'Paired-Comparison'[9], which has the advantage of requiring a simple response to indicate the choice from only two samples. However, the paired comparison required becomes prohibitive with a large number work when the complete matrix of comparisons is carried out (for example, N(N-1)/2) of paired comparisons for N stimuli)[10]. Thus, researchers sometimes conduct incomplete-paired-comparison experiments when it has high number of stimuli[11]. The rank-order method requires participants to provide an order to all colours from the most to the least preferred for all colours simultaneously displayed[12]. The subjective rating[13] and affective judgment[14] methods are made by response scales for how much they prefer each single colour, such as N-point Likert scale or a line-mark rating.

This research concerned with finding an efficient research method to test individual colour preference. Four experimental methods have been chosen, n-alternative forced-choice method, rank-order, rating and paired comparison experimental methods. The research aim is: 1) to test the agreements of individual colour preference results between four colour preference research strategies: N-alternative forced-choice method, rank-order, rating and pair comparison; 2) to determine a lower limited number of participants with consistent and reliable results for four strategies.

EXPERIMENTAL

It has been suggested that there was no statistical significance between online and laboratory environments for individual colour preference test for *experimental environments*[7]. Although, the online environment is less-controlled than the laboratory environment in display technology or viewing conditions etc. However, the advantage is relatively easy to recruit a large number of participants, and robust estimates of individual colour preference for groups of participants (such as nationality or even socio-economic status). Thus, online questionnaire with non-controlled environment has been chosen in this research. Four experimental methods have been used, N-alternative forced-choice method (AFC), rank-order, rating and pair comparison.

Six colours (red, orange, yellow, green, blue and purple) were selected from Adobe HSB colour system [12, 15, 16] to determine participants' individual colour preference. This study aims to evaluate the colour preference research method. Since the hues exhibit their maximum chroma values at different lightness values; keeping lightness constant would limit the colours to different chroma values. Consequently, in this study there was some variation in lightness and chroma. Thus, hue is the only colour variable considered although in principle the work could be extended to include lightness and chroma in this study. Colour patched were defined by sRGB values and the actual colours that were displayed on the monitor (HP DreamColor LP2480zx Professional Display, the Y value of the display white was 218 cd/m^2) were measured using a Minolta CS100A colorimeter. Colour information is reported in TABLE 1 by using sRGB and CIE L*a*b* values.

TABLE 1 The sRGB and CIE L*a*b* colour coordinates of the six colour patches.

Colour coordinates of the six basic colour patches						
Coloured Squares	Red	Orange	Yellow	Green	Blue	Purple
sRGB	255, 0, 0	255, 127, 0	255, 255, 0	0, 255, 0	0, 0, 255	127, 0, 255
CIE L*a*b*	40.01, 55.08, 58.27	49.84, 28.88, 64.20	72.35, -17.18, 80.75	63.85, -68.14, 69.37	17.65, 75.93, -89.76	29.46, 61.16, -63.54

Participants were presented with six colour patches at a same time in a random location on a random display, such as a monitor, a smart phone etc. In the AFC test (Study I), participants were asked to indicate *one colour* which they prefer most; for the rank-order test (Study II), participants were asked to give the *sequence for six colours* of their individual colour preference from the most to least[17]; for the rating test (Study III), participants were asked to *scales their individual colour preference for each colour* (six scales) by line-mark ratings (from 0%-100%)[16]. A total 338 participants were recruited worldwide to participate, comprising of 192 participants for AFC test (including 72 males and 120 females), 85 participants for rank-order test (including 42 males and 43 females) and 41 participants for rating test (including 20 males and 21 females). Experimental details please see FIGURE 1.

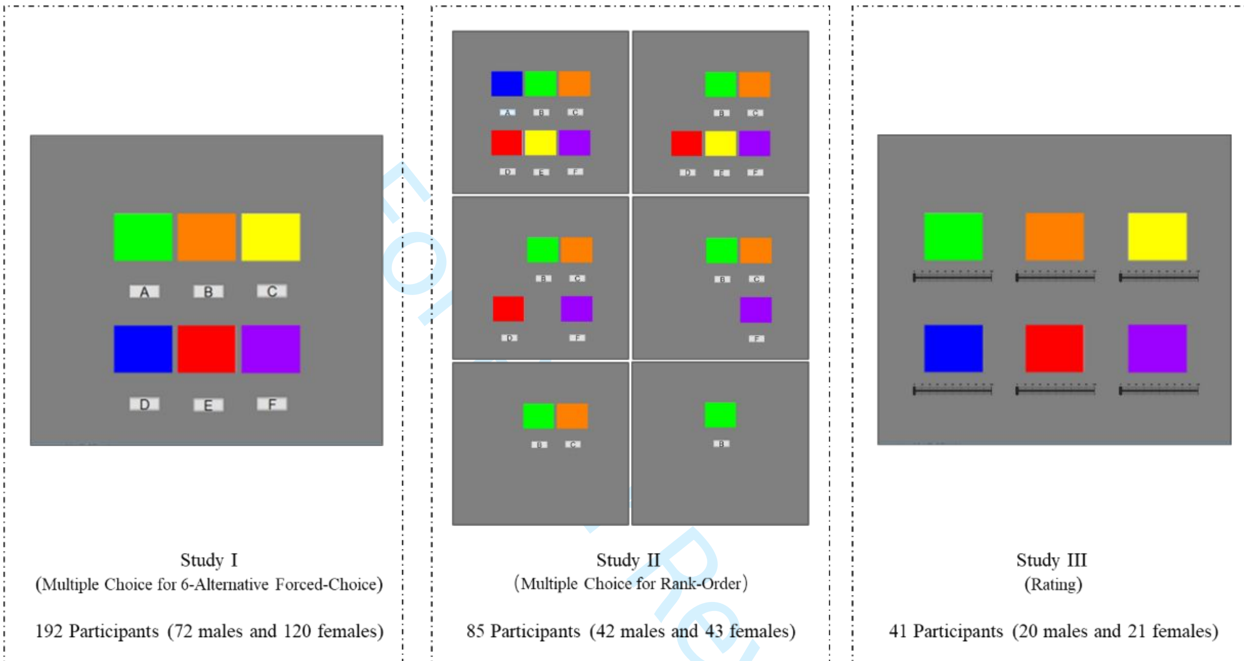

FIGURE 1. Graphical user interface for Study I, Study II and Study III in which the coloured patches were presented randomly for each participant.

RESULTS AND DISCUSSION

The Individual colour preference has been analysed by AFC, Rank-Order, Rating and Paired-Comparison methods. First of all, 338 responses have been collected as AFC method (where the colour was placed in the first choice was considered in the rank-order method, and the colour was obtained the highest rating was counted in the rating method); 126 responses have been collected as rank-order method (where the rating preference sequence was considered as a rank-order); 41 responses have been collected in rating method, and 41 responses have been collected in paired comparison method (where the rating score results comparison was consider as a paired comparison). More details please see TABLE 2.

TABLE 2 Data details for analysis.

Data type	Data from	Total number of data
AFC	Study I, Study II, and Study III	338
Rank-Order	Study II and Study III	126
Rating	Study III	41
Paired-Comparison	Study III	41

In order to compare four methods, the interval scale values *z* have also been used on AFC, rank-order, rating and paired comparation methods. The following four steps are:

- For AFC method, the individual colour preference percentage for each colour is the number of times that each colour has been selected as the most preferred (percentage from 0% to 100%)[18];
- For rank-order method, the ordinal rank and comparative data are converted to interval-data z scores[7]. The rank-order data from each participant were combined to mean rank data and subsequently proportion values (between 0% to 100%)[19];
- For rating method, each colour preference percentage has been averaged by participants' rating from 0% to 100% as non-prefer to most prefer[4];
- For the paired-comparation method, each individual rating data set made by each respondent for all possible pairs, the average of weighted score (between 0% to 100%) for each colour have been produced[14];
- The proportion values (four methods) were converted to interval scale values z using the inverse of the cumulative standardized normal distribution according to case V of Thurstones Law of Comparative Judgement[20]; additionally, the proportion of 0 and 1 are replaced with 1/999 and 999/1000 respectively when the proportions are exactly 0 or 1[7, 16].

The colour preference sequence for *z* scores for AFC method is blue > orange > red > yellow > purple > green; the preference sequence for rank-order method is orange > yellow > blue > red > green > purple; for rating is orange > yellow > blue > red > purple > green; which represents a similar sequence (please see FIGURE 2.1). However, it might be augured that the data were collected from different populations. In order to reduce the differentiation, the results from rating group have been re-analysed. The rank-order, rating and paired comparison methods have a same colour preference *z* scores sequence as, orange > yellow > blue > red > purple > green. Also, AFC method also obtained a similar sequence, orange > blue > yellow > purple > red > green (Please see FIGURE 2.2).

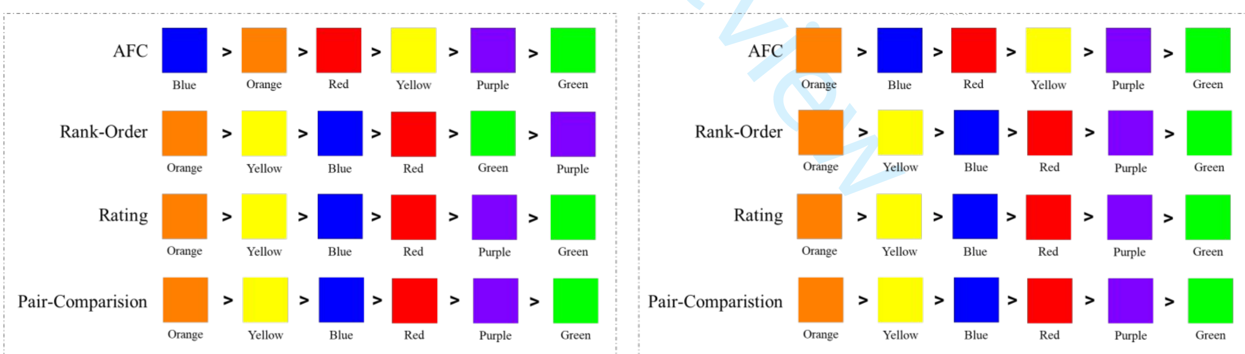

Figure 1.1 Individual colour preference colour preference rank between AFC, Rank-Order, Rating and Paired Comparison research methods for full data.

Figure 1.2 Individual colour preference colour preference rank between AFC, Rank-Order, Rating and Paired Comparison research methods for same data.

FIGURE 2. The individual colour preference rank sequences from N-alternative forced-choice, Rank-Order, Rating and Paired Comparison experimental methods.

Moreover, this study aims to explore the efficiency way on scaling individual colour preference, between experimental methods and participant numbers. The Monte Carlo simulation analysis was used to explore the level of agreement of these four methods[7, 16]. The data from rating were used as four methods. 41 responses were sub-sampled, and sub-sample the full population repeatedly taking a different sub-sample each time. The correlation coefficient (r-squared value which equals the square of the Pearson correlation coefficient) has been used to indicated the proportion of the variation in the dependent variable (where an

r^2 of 1 indicates that the regression predictions perfectly fit the original full data). The following four steps are:

- Subsample *n* (eight interactions) of the data from *the 41* participants *randomly* where *n* is [*36, 31, 26, 21, 16, 11, 6 and 1*] (Step 1);
- Calculate the individual colour preference *z* score for each *n* scale participants (n = 36, 31, 26, 21, 16, 11, 6 and 1) to obtain the three distributions of individual colour preference for each *n* scales from AFC, rank-order, rating and paired-comparation methods (Step 2);
- Construct the correlation coefficient (r^2) calculated between the individual colour preference from all the data (*n* = 41) and from the sub-sampled data [*n = 36, 31, 26, 21, 16, 11, 6 and 1,*] (Step 3, FIGURE 3);
- Compare the r^2 distributions between the four methods (Step 4, FIGURE 4).

FIGURE 3. The correlation for the individual colour preference between full samples (n=41) and each the sub-sampled samples (n=36, 31, 26, 21, 16, 11, 6 and 1) from N-alternative forced-choice, Rank-Order, Rating and Paired Comparison experimental methods (an example illustration for Step 3).

Figure 2 and Figure 3 illustrates the Step 3 and Step 4, presenting the correlation coefficient values with *n* scale samples (n=36, 31, 26, 21, 16, 11, 6 and 1). In Figure 3, the red, grey and blue lines show the r^2 distributions from AFC, rank-order and rating methods; each point represents the r^2 for each *n* scale samples. Comparing the small-scale samples tests (*n*=11, *n*=6 and *n*=1), the correlation coefficients hold on the high agreements) for both rating (0.84, 0.79 and 0.63), rank-order (0.82, 0.75 and 0.60) and paired comparison (0.88, 0.67 and 0.77) methods. However, the correlation coefficients decrease sharply to the low agreement positions for AFC method (0.43, 0.36 and 0.1). It suggests that the rating and rank-order methods are more stable than AFC in small number of responses.

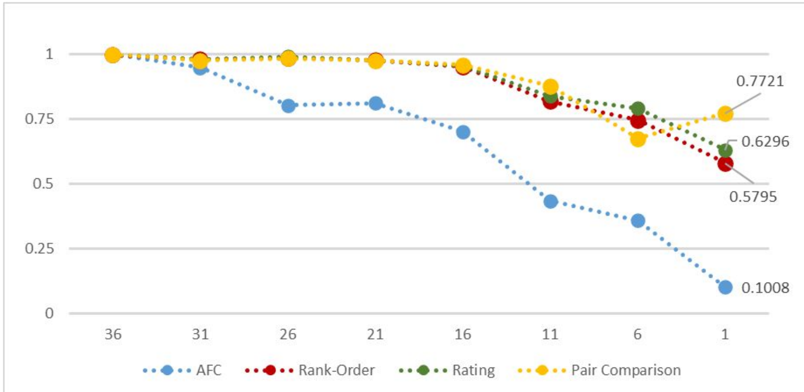

FIGURE 4. An example illustration for the correlation coefficient distributions from four experimental methods. The vertical and horizontal axes represent r^2 values and n scale samples (Step 4).

The Monte Carlo comparisons evaluate the stability of the four methods, the simulation was repeated 100 times. Each simulation starts with a different *random set of scale samples*, and the mean correlation coefficient was used as a measure of performance. FIGURE 5 displays the box plots for the r^2 distributions from N-alternative forced-choice, Rank-Order, Rating and Paired Comparison experimental method. The vertical axes and horizontal axes represent the correlation coefficient values and *n* scale samples. Notice that the median for rating, paired comparison and rank-order methods are higher than AFC method when samples less than 11 for example, especially, for the sample size as 11, 6 and 1, shows by the red horizontal line (for the size 1, $r^2 \sim 0.2$ in AFC method, $r^2 \sim 0.59$ in Rank-Order method, $r^2 \sim 0.67$ in rating method, and $r^2 \sim 0.65$ in paired comparison method). In other words, the results from Monte Carlo simulations suggests that the rating and rank-order methods are more relay and stable than the AFC method for individual colour preference when n is small. This suggests that, for studies involving small numbers of participants, the rank-order rating and rank-order should be considered.

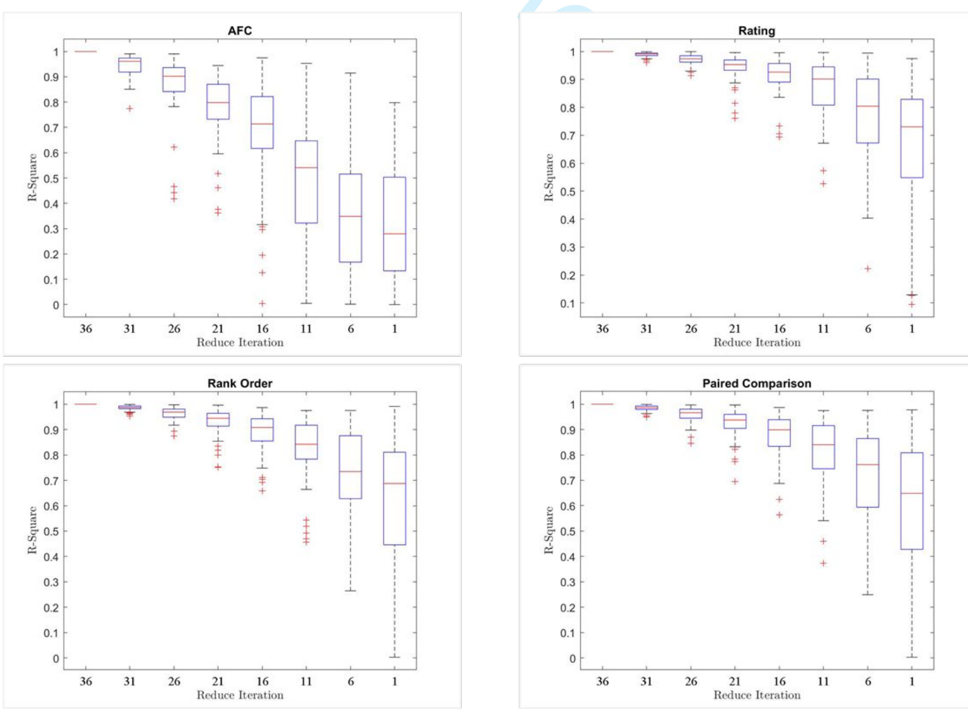

FIGURE 5.The box plots for the r^2 distributions from N-alternative forced-choice (AFC), Rank-Order, Rating and Paired Comparison experimental methods experimental methods.

DISSECTION

Colour is an important tool for products design[21], environment design[22], information design[23,24] etc. However, finding an efficient method to investigate individual colour preference is important for designer and researcher. The individual colour preference results from this work suggested orange is the strongest preferred colour in these six hues and green is the weakest preferred in four individual colour preference experimental methods (N-alternative forced-choice, Rank-Order, Rating and Paired Comparison). The Monte Carlo Analysis method also be employed to compare the result performance of the methods for individual colour preference by repeating 100 trials. The average correlation coefficients variation range for AFC method is wiled than both rating and rank-order methods in the small number of participants. That suggests the rating, paired comparison and rank-order method are more stable than the AFC method when only small number participants take part in the experiment, such as for studies involving small numbers of participants (less than 11 participants), the rating and rank-order method should be preferred. Even the exciting work suggests that, the uncontrolled experimental environment with a large number of participants have been encourage for individual colour preference[7,16]. However, the influence for controlled and uncontrolled experimental environments with different individual colour preference research methods are still unknow[25]. This study is noted that this work was concerned with that aspect of color known as hue. However, it is acknowledged that lightness and chroma may also be important for individual research. It is clear that much more work is required in this field.

REFERENCES

[1] Camgöz, N., Yener, C., & Güvenç, D. (2002). Effects of hue, saturation, and brightness on preference. *Color Research & Application, 27*(3), 199-207.

[2] Mikellides, B. (2012). Colour psychology: The emotional effects of colour perception. In *Colour Design* (pp. 105-128). Woodhead Publishing.

[3] Grieve, K. W. (1991). Traditional beliefs and colour perception. *Perceptual and motor skills*, *72*(3_suppl), 1319-1323.

[4] Jiang, L., Cheung, V., Westland, S., Rhodes, P. A., Shen, L., & Xu, L. (2020). The impact of color preference on adolescent children's choice of furniture. *Color Research & Application*, *45*(4), 754-767.

[5] Lee, G., Westland, S., & Cheung, V. (2019). Colour communication challenges: Exploring disciplinary divides. *Journal of the International Colour Association*, *23*, 25-35.

[6] Swasty, W., Mustikawan, A., & Naufalina, F. E. (2020). Visual Perception of Primary Display Panel of Coffee Packaging. *Jurnal Manajemen dan Kewirausahaan*, *22*(1), 73-79.

[7] Yu, L., Westland, S., & Li, Z. (2021). Analysis of experiments to determine individual colour preference. *Color Research & Application*, *46*(1), 155-167.

[8] Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual review of psychology*, *64*, 77-107.

[9] Ling, Y., Hurlbert, A., & Robinson, L. (2006). Sex differences in colour preference. *Progress in colour studies*, *2*, 173-188.

[10] David, H.A., 1987. Ranking from unbalanced paired-comparison data. *Biometrika, 74(2), pp.432-436.*

[11] Dittrich, R., Francis, B., Hatzinger, R. and Katzenbeisser, W., 2012. Missing observations in paired comparison data. *Statistical Modelling, 12(2), pp.117-143.*

[12] Holmes, C.B., Edward Fouty, H., Wurtz, P.J. and Burdick, B.M., 1985. The relationship between color preference and psychiatric disorders. *Journal of Clinical psychology*, *41(6), pp.746-749.*

[13] Adams, R.J., 1987. An evaluation of color preference in early infancy. *Infant Behavior and Development*, *10*(2), pp.143-150.

[14] Ou, Li-Chen, M. Ronnier Luo, Andrée Woodcock, and Angela Wright. A study of colour emotion and colour preference. Part I: Colour emotions for single colours. *Color Research & Application* 29, no. 3 (2004): 232-240.

[15] Yu, L., Westland, S., Li, Z., & Xia, G. (2021). The effect of decision time-length condition on consumer product-colour purchase decision. *Color Research & Application*.

[16] Yu, L., Westland, S., Cheung, V., & Xia, G. (2021). Analysis of Research Strategies to Determine Colour Preference II: AFC, Rank-Order and Rating. In *14th Congress of the International Color Association* (pp. 885-889). International Colour Association.

[17] Yu, L., Westland, S., Li, Z., Pan, Q., Shin, M. J., & Won, S. (2018). The role of individual colour preferences in consumer purchase decisions. *Color Research & Application*, *43*(2), 258-267.

[18] Yu, L., Westland, S., Li, Z., Pan, Q., Shin, M. J., & Won, S. (2018). The role of individual colour preferences in consumer purchase decisions. *Color Research & Application*, *43*(2), 258-267.

[19] Westland, S., Li, Y., & Cheung, V. (2014). Monte Carlo analysis of incomplete paired-comparison experiments. *Journal of Imaging Science and Technology*, *58*(5), 50506-1.

[20] Hohle, Raymond H. An empirical evaluation and comparison of two models for discriminability scales. *Journal of Mathematical Psychology* 3, no. 1 (1966): 174-183.

[21] Gong, H., Yu, L., & Westland, S. (2020, November). Simple primary colour editing for consumer product images. In *Color and Imaging Conference* (Vol. 2020, No. 28, pp. 270-276). Society for Imaging Science and Technology.

[22] Xia, G., Henry, P., Queiroz, F., & Westland, S. (2021). Effects of coloured lighting in the real world environment and virtual reality. *Journal of the International Colour Association*, *27*, 9-25.

[23] Xia, G., Henry, P., Queiroz, F., & Westland, S. (2019, November). Effects of object colour stimuli on human brain activities and subjective feelings in physical environment and virtual reality. In *Proceedings of the International Colour Association (AIC) Conference 2019*.

[24] Chen, Y., Yu, L., Westland, S., & Cheung, V. (2021). Investigation of designers' colour selection process. Color Research & Application, 46(3), 557-565.

[25] Zuffi, S., Brambilla, C., Eschbach, R., & Rizzi, A. (2009, January). A study on the equivalence of controlled and uncontrolled visual experiments. In *Color Imaging XIV: Displaying, Processing, Hardcopy, and Applications* (Vol. 7241, p. 724102).

Authors

Luwen Yu received her BA and MA degrees in Industrial Design from Huazhong University of Science and Technology, and received PhD degrees in Design from University of Leeds. She is now a Research Fellow in School of Design at the University of Leeds. Her research interests include industrial design, colour design, visual communication, immersive technologies, and design thinking and innovation.

Guobin Xia is the Assistant Professor of Digital Design and Innovation in Heriot-Watt University. His areas of interest are Colour design, Immersive and transformative experience design, and Data-driven design. Meanwhile, he is an invited reviewer for many prestigious publishers, including Sage, Routledge (Taylor & Francis) and Springer.

Stephen Westland received BS and Ph.D degrees from the University of Leeds. He is Professor of Colour Science and Technology at the University of Leeds where he leads the Experience Design Research Group. His research interests include color management, colou imaging, colou design, and data-driven design.

Zhenhong Li received Ph.D. degree in control engineering from the University of Manchester. He is now a research fellow in the School of Electronic and Electrical Engineering at the University of Leeds. His research interests include rehabilitation robot, human-robot interaction, distributed optimization and data-driven control.

Vien Cheung is an Associate Professor in Colour and Imaging Science in the School of Design, University of Leeds. Her research interests are colour vision, spectral imaging, colour reproduction and colour, all as applied to the art and design disciplines.