Towards Customized Emotional Design: An Explorative Study of User Personality and User Interface Skin Preferences

Willem-Paul Brinkman and Nick Fine

Brunel University

Uxbridge, Middlesex
UB8 3PH
United Kingdom
{willem.brinkman, nick.fine}@brunel.ac.uk

ABSTRACT

Two key elements in the area of cognitive ergonomics are user-system performance and the user similarity. However, with the introduction of skinnable user interfaces, a technology that gives the user interface a chameleon-like ability, elements such as aesthetic, fun, and especially user individuality and identity become more important. This paper presents two explorative studies on user personality in relation to skin preferences. In the studies participants were asked to rate their preference of a set of Windows Media Player skins and to complete the BIS/BAS and the IPIP-NEO personality inventories. The results of the first study suggest colour and similarity-attraction as two possible underlying factors for the correlations found between personality traits and skin preferences. The results of the second study partly confirm these findings, however not for similar personality traits and skin types correlations.

Keywords

skin, personality, customisation, user interface.

INTRODUCTION

Recently a number of people (e.g. McCarthy and Wright, 2004; Norman, 2004) have argued that the traditional focus on performance in product design might be too narrow, and things such as aesthetics, fun and the user self-image should be considered. These are all highly personal factors, with a large variation between users. Therefore, the traditional approach of designing for a single average user seems less effective here. Instead, user interfaces should be more fitted to the needs and desires of the individual user. However, most research on customisation focuses on function customisation, and less on form customisation. For example, personalised search engines and customised online purchase offers are both function customisations emphasising the importance of performance and usefulness of a product. When it comes to form customisation, which stresses aesthetic and the emotional need of the user, designers seem mostly guided by what is technically possible, and not always by what is desired by the users. Therefore, a better understanding of form customisation seems welcome.

Skinning

The introduction of skinning technology allows developers to change the look and feel of a system at runtime. Even after the system is taken into operation, developers can still redesign the appearance without reprogramming the system. Skinning can be very basic, such as changing images, or the font type, but also more extensive, such as changing the interaction style. A number of commercial applications already apply this technology, e.g. Windows Media Player, ICQPlus, and Winamp. The benefit for users is that they can change the entire interface in one instance, reducing the complexity of setting each system property individually. For the developers it means that they can customize their system to target specific user groups.

User personality

A research area that traditionally focuses on individual differences is Personality Psychology. The psychometric tests developed in this area could help designers to segment their user population. This would be useful since some research already suggests a link between personality and skin preference. For example, Nass, Moon, Fogg, Reeves and Dryer (1995) studied the similarity-attraction hypothesis and showed that a computer with a dominant interaction style was more preferred by dominant users than by submissive users and vice versa. Furthermore, Karsvall (2002) suggested a relationship between extraversion, and the colour and shape of an interface. Still, understanding in this area is limited and fragmented. Therefore, we conducted two explorative studies to develop hypotheses about the relation between several personality traits and the users' skin preference.

FIRST STUDY

The first study that we conducted aimed to develop hypotheses about the relation between several personality traits and the users' skin preference. Although this relation might evolve throughout the various usage phases: selection, initial use, and long term use of an application, this study only looked at the first phase —the selection. In this phase users often scroll through a list of skins from which they can select a skin to apply to an application. These lists often

Table 1: Skins used in the first study with their average rating.

No Title	Mean	No Title	Mean	No Title	Mean
0 Melvin	2.46				
1 Catwoman	3.34	21 Predator	4.36	41 Television	2.45
2 Tiny Player	3.60	22 Revert Superior	4.38	42 Koolskin for tutorial	3.71
3 DigitalAudioPlayer	2.90	23 Ducky	3.24	43 Divx	4.34
4 Longhorn Prototype for WMP	3.92	24 Pharaoh	3.18	44 Claw	2.30
5 Millennium	2.70	25 Pioneer AVIC-N1	4.43	45 Chrome Dome	3.62
6 Robotic Armature Device	3.79	26 Tube Frame	2.81	46 Back to the Future Trilogy	3.40
7 Eye of Africa	2.49	27 Ultranation	2.47	47 MERbar	3.72
8 Stealth	2.64	28 Raptor	3.78	48 Foo Fighters – One by One	2.38
9 US Army	3.89	29 Aoe (Age of Empires)	2.34	49 Gadget	2.85
10 Le Crapouillot et ses Geckos	2.30	30 Sportcartoon	2.30	50 PADD	2.69
11 Tokyo Matrix	2.49	31 Winp Media Player	3.86	51 To End All Wars	2.12
12 DMv2.0	4.36	32 Bluegrid	3.79	52 Military	2.17
13 Portals	2.73	33 Mtv411 Designer Jukebox	3.45	53 Dreamcatcher	2.55
14 The Last Samurai	3.09	34 Xebish1-0	3.17	54 Mtv411 Platinum Player	3.15
15 The Stick	3.44	35 Ocean	3.12	55 Combat Flight Simulator 3	2.48
16 Mac OS Remix	3.18	36 Sankofa Bird	2.42	56 Scooby-Doo 2	2.72
17 Official DVD Profiles Skin	4.28	37 Israeli	2.67	57 wcnwms	2.50
18 Blue Martian	2.86	38 Protozoa	4.04	58 M.D.U 38	2.95
19 Futurama Season 4	3.04	39 Modern blue	3.18	59 T3: Skynet	3.81
20 Age of Mythology for MP7	2.73	40 Shandy	3.61	60 Anatman	2.76

The slides can be downloaded from http://people.brunel.ac.uk/~csstwpb/skintest1.pdf

consist of a small image of the skin, its name, and sometimes a description. When users select these skins they often have not interacted with them. Therefore, this phase seems mainly driven by things such as their prior knowledge, their taste and expectation, and less by their actual experience with a particular skin.

METHOD

In the first study a total of 99 people, including both students and staff of Brunel University, rated 61 Windows Media Player skins (Table 1). The skins were randomly selected out of a set of 178 skins that were downloaded from the Internet. In the test each skin was presented on a single slide with screen shots of the player in four situations: (1) the idle situation, (2) the radio situation, (3) the video situation, and (4) the CD/MP3 situation. The participants saw the slides in a classroom setting, with a projector projecting the slides on a large screen. First the participants saw the complete set of skins. This would allow them to establish a frame of reference for their preference rating, and would minimize possible learning effects. After this the skins were again presented in sequences of 10 skins, with skin 0 as a trial session. The participants first saw, for two seconds, a green slide with the skin number. Next they saw for five seconds the slide with the four screen shots of the skin, and finally they saw a blue slide for four seconds with the statement and a seven points Likert scale ranging from 1 (unlikely) to 7 (likely). The participants had to rate the following statement: "I would try this skin on my media player". The participants were instructed to assume that they would have a PC on which the Windows Media Player was installed, and they were also reminded that once a skin was selected in this application it was always possible to change it again. After participants rated the 61 skins, they were asked to complete a 24 questions test to rate them on the BIS/BAS scales (Carver and White, 1994) assessing their motivation towards something desirable or undesirable. Finally the participants were asked to take the short version of the online IPIP-NEO test (Buchanan, Johnson, and Goldberg, 2005), and to hand in the results of this test at a later time. This inventory consists of 120 questions, and measures differences between individuals on five dimensions, the so-called five-factor model (McCrae and Costa, 1987), which are: extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience.

RESULTS

Table 1 shows the average rating for each of the skins. The rating ranges from 2.12 for the To End All Wars skin (51) to 4.43 for the Pioneer AVIC-N1 skin (25), with a standard deviation ranging from 1.52 to 2.14. The first analysis explored a possible underlying structure on which participants rated the skins. The skin preference data from each individual participant was standardized using a z-score transformation to reduce the individual variation in responding to a Likert scale, e.g. using the entire range of the scale or only a sub range, or individual tendency to give high ratings or only low ratings. The analysis, conducted with principle

component as factor extraction method, suggested a solution with 19 components each with an eigenvalue above 1. These 19 components explain 0.74 of the overall variance. Exploring the varimax rotated component matrix, did not reveal an easy interpretable component structure. Therefore this analysis does not suggest that the participants applied a small number of underlying factors to rate these skins.

Gender and Age

The next step was calculating and interpreting the correlations between preference rating and the individual characteristics, such as gender, age, and personality.

Table 2: Correlations between skin preference rating and the factors age and gender (female=0, male=1).

Factor	Skin no (correlation)
Age	18(.31**) 49(.23*) 58(.22*)
Gender	0(-.25*) 8(-.23*) 11(-.23*) 18(-.23*) 21(-.21*) 23(-.34**) 25(.37**) 32(-.34**) 40(.39**) 43(.21*) 48(-.41**) 52(.25*) 53(.30**) 59(.43**)

* sign. at a 0.05 level, ** sign. at a 0.01 level.

Of the 99 participants in the first study, 58 were male, 35 were female and 6 did not specify their gender. The age of participants was between 18 and 52 years with a 20 years median. Table 2 shows that the rating of three skins resulted in a significant Spearman correlation with age. Examining the three skins, however, did not reveal a common factor or theme in the skins. Table 2 also shows the 14 significant Spearman correlations between gender and skin preference. Examining the skins, it seems that female participants were more in favour than the male participants of skins with an image of a 'cute' penguin (skin no 0), or a 'cute' duck (skin no 23). Colour might be another common factor, as dominantly blue (skin no 18 and 32), purple (skin no 48), or white (skin no 11, and 21) coloured skins were more preferred by female participants than by male participants. On the other hand, the male participants were more in favour than the female participants of skins with themes such as the "Military" (no 52), "Dreamcatcher" (no 53), and "Terminator 3: Skynet" (no 59). Another interpretation of the results was the higher preference by male participants than by female participants for skins that might be described as large or complex by displaying a large number of buttons (no 25, 40, 43).

BIS/BAS

The significant Spearman correlation of the BIS/BAS scales and the skin preference rating are presented in Table 3. The scores on these scales have been reversed. Therefore a high score is seen as more sensitivity for the two general motivation systems. For the drive subscale, associated with persistent pursuit of desired goals, the table reveals five significant correlations with skin preferences. Examining these skins it seems that skins with negative correlations are more over-powering, or darker (no 9, 28 and 34), whereas skins with a positive correlation are lighter, or brighter (no 11 and 51). The fun seeking subscale, associated with a desire for new rewards and a willingness to persuade a potentially rewarding event impulsively, had only one significant correlation, a negative correlated with the preference for the penguin skin (no 0). The last subscale, reward responsiveness associated with the positive reaction towards rewards or anticipation of rewards, had a significant correlation with the preference of four skins. Examining these four skins, however, does not reveal a common factor. The interpretation of the results of the BIS scale, associated with motivation of avoiding undesirable situations, seems easier. All three skins (no 25, 32, 37) are mainly blue. The only puzzling factor is that these are both positive as negative correlations.

Table 3: Correlations between personality and skin preference.

(Sub)Scale	Skin no (correlation)
BAS	
-Drive	9(-.34**) 11(.27**) 28(-.26*) 34(-.28**) 51(.27**)
-Fun Seeking	0(-.26*)
-Reward Responsiveness	1(-.30**) 9(-.27**) 25(-.25*) 49(.29**)
BIS	25(-.23*) 32(.22*) 37(.23*)
IPIP-NEO	
Extraversion	10(.52*) 51(.55*)
-Friendliness	6(.67**) 10(.56*)
-Gregariousness	7(.55*) 44(.53*)
-Assertiveness	21(.58*) 48(-.68**)
-Activity level	35(.61*) 43(-.57*)
-Excitement-Seeking	19(.54*) 26(-.77**) 32(-.72**)
-Cheerfulness	28(-.53*)
Neuroticism	5(-.52*) 37(.68**)
-Anxiety	7(-.78**) 37(.62*)
-Anger	7(-.70**) 57(-.53*)
-Depression	56(.75**)
-Self-Consciousness	0(.54*) 13(-.72**) 49(-.58*) 54(.64**) 56(.54*)
-Immoderation	19(.55*)
-Vulnerability	5(-.63*)

* sign. at a 0.05 level, ** sign. at a 0.01 level.

Five-Factor Personality Inventory

Out of 99 participants 17 returned their results of the IPIP-NEO personality test. Table 3 shows the results of the significant partial Pearson correlation, controlled for age and gender, between the IPIP-NEO scales and the skin preference. To limit the bias that possible outliers can have in a small data set, the table only shows the results if the scale/skin combination also had a significant Spearman correlation. The analysis showed that the extraversion, agreeableness and neuroticism domains make up the main part of the correlations found. Interpreting these correlations for the agreeableness domain, however, was less clear than for the extraversion and neuroticism domain, and the

analyses, therefore, only focused on the last two domains here.

Similarity-Attraction

Nass' et al. (1995) hypothesis of similarity-attraction seems to be confirmed by findings such as a correlation between friendliness and a robotic armature device skin (no 6), but also with the gecko skin (no 10); both can be seen as harmless, or even friendly creatures. Other support is the negative correlation between cheerfulness and the "Raptor" skin (no 28), which is a large dark skin in shape of a reptile predator, but also the negative correlation between neuroticism, vulnerability, and the scary looking "Millennium" skin (no 5), or the positive correlation between the two sub-domain of neuroticism, depression and self-consciousness, and the skin rating of the spooky-mystery solvers "Scooby-Doo" (no 56). Finally this hypothesis is also supported by correlations between immoderation, excitement-seeking and the rating of the "Futurama" skin (no 19), a humoristic cartoon series.

Colour

A second observation is that of colour as an underlying factor explaining the correlations between personality domains such as extraversion and the sub-domain assertiveness, and skins such as the bright gecko skin (no 10), the mainly white coloured "To End All Wars" skin (no 51), and also the very shiny white and blue "Predator" (no 21) skin. More support can also be seen in the positive correlation between gregariousness and the yellow, green, red coloured "Eye of Africa" skin (no 7), and the negative correlation between this skin and anxiety, and anger. Colour as a factor can also be interpreted as the significant element in the relationship between neuroticism, anxiety and the mainly blue coloured "Israeli" skin (no 37), but also for the negative correlation between assertiveness and the mainly purple "Foo Fighter –One by One" skin (no 48). Finally, colour might also be the underlying factor for the negative correlation between activity level and grey coloured "Divx" skin (no 43).

SECOND STUDY

The results of the first study supported the idea that skin preferences have a relationship with individual characteristics such as gender, age and personality. However, as this was an initial explorative study, the interpretations of these results are very subjective, especially the explanative element such as skin colour or interpretation of the skin symbolism. The study does not provide reliable answers whether these are indeed the actual underlying factors that determine the link with individual characteristics. Next, the large number of correlations that were examined creates a considerable opportunity for capitalization on chance. The relative small sample size of IPIP- NEO personality data also makes the results related to personality potentially unstable, in that adding a number of new samples to the analysis might alter the findings considerably. Therefore, a second study was set-up to test the ideas of the first study. Instead of randomly selecting skins from Internet, as in the first study, skins were selected based on 17 categories of skin properties, such as their colour, their shape, their theme etc. These categories were identified from the findings of the first study as potential relevant skin properties that correlated with age, gender, or the users' personality. Table 4 shows the categories and the skins in these categories. Again skins were downloaded from the Internet; however, this time some skins were also specially designed for a category, as no suitable candidate could be found online. An asterisk indicates these skins in Table 4.

METHOD

The procedure of the second study was similar in set-up to the first study. However this time the skins were presented in sequences of 17 skins. To limit possible order, learning or fatigue effects, the skins of each category were distributed equally over the beginning, middle or end of the sequences. After rating the skins, the participants were asked to answer a reduced version of the BIS/BAS questionnaire. The questionnaire did not include the nine questions related to Fun Seeking and Reward Responsiveness sub-scale. To increase the sample size of the personality data, participants in the classroom were also asked to rate the 48 statements taken from the IPIP-NEO inventory that are related to extraversion and neuroticism scale and their sub-scale. A total of 130 undergraduate students of Brunel University participated in the second study.

RESULTS

Table 4 shows the average rating for the individual skins. The rating ranges from 1.99, for Pda skin (22), to 4.85, for the 10series skin (9), and the standard deviation ranged from 1.33 to 2.21. Examining the table shows also a relatively high overall rating for skins in category Q —Standard skins. It seems that the participants on average had a high preference for the standard Windows skins available for the Media Player. However, the table also shows the number of participants that gave a skin their highest rating. The standard skins 9 and 18 received the highest rating from 25% of the participants, whereas The Simpsons skin (37) received the highest rating from 36% of the participants. Assuming that the standard skins were designed to maximize average preference, this result shows the limitation of designing for similarity. The standard skins were not the most preferred skins for the largest number of participants. In other words, although standard skins might be on average the best design compromise, they were not as often as the Simpsons skin the first choice of participants.

Analysing the z-scores of the preference data it was also possible to study how many of the responding participants (n) on average agreed with each other. For each skin the numbers of z-scores above zero (a) and below zero (b) were counted. Next, for each skin the percentage of agreement was calculated as follows: $agreement = \frac{\max(a,b) - \frac{1}{2}n}{\frac{1}{2}n}$.

Table 4: Skins used in the second study with the average rating, and the frequency of the highest rating.

Category	no	Title	Rating: Mean	High	Category	no	Title	Rating: Mean	High
A: Blue skins	12	Colorchooser (blue)	2.67	2	J: Small and friendly creature	13	Eros (+)	2.71	8
	32	Color + OS Blue (+)	3.24	8		34	Anemone (+)	2.78	8
	40	Royale-X's player (+)	3.29	3		39	Springflower (+)	2.68	11
	54	StellarHub (New)	3.78	17		52	UglyPLayerKappa2-0-0 (+)	2.21	4
B: Purple/ Violet skins	14	Imusica	3.78	9	K: Friendly, charming 'baby' character	1	Bubbles (+)	2.94	9
	29	Colorchooser (Purple)	2.44	3		24	iRover	2.18	7
	38	Color + OS Violet	3.21	6		51	Toothy (+)	2.89	12
	57	XBOX_Musci_Mixer	3.90	15		62	Thomas (+)	2.91	8
C: Grey skins	6	Greyscale (+)	2.44	7	L: Humoristic skins	15	Spongebob squarepants* (+)	3.37	18
	21	Color + OS Grey*	3.24	5		27	Little Britain* (+)	2.13	3
	46	DTPM (+)	2.90	3		37	The Simpsons* (+)	4.84	47
	65	Colorchooser (Grey) (+)	2.08	1		59	Monty Python* (+)	2.55	6
D: Bright skins	4	Color + OS Yellow	2.88	4	M: Action theme skins	2	Crimson_skies	2.73	3
	23	mPod	4.17	23		26	Xbox Live Skin (+)	4.25	28
	48	Colorchooser (white)	2.58	1		50	NeedforSpeed…(+)	3.62	15
	63	Fathom WMP	2.42	5		60	BlueCruhs_MP7 (+)	4.15	21
E: Colourful skins	5	Aquaplayer	2.74	4	N: Predator skins	16	Snake* (+)	3.06	14
	19	Blinx (+)	3.95	18		28	T-Rex* (+)	2.77	7
	47	ColorTek1-1 (+)	2.39	3		36	Aliens* (+)	2.96	9
	67	DigitalDJ	3.18	7		58	Jaws (+)	2.71	7
F: Large, dark skins	3	Godey Classic	2.90	5	O: Scary skins	8	Halfskull (+)	2.85	9
	25	BP.net v1.0	2.89	9		20	KoRnIssues1-35	2.63	4
	49	SplinterCell	3.64	15		44	Blair witch project*	2.13	5
	61	Rook	2.18	3		66	EvilHeadspace1-0 (+)	2.91	13
G: Small skins	11	Basic_4 (BASIC.4) (+)	2.23	3	P: Mildly scary skins	17	Tore-upxpwmp	2.95	8
	31	Jskin (+)	2.17	2		30	Halloween (+)	3.15	8
	41	Luster (+)	2.69	3		35	MyHalloweenwmp (+)	2.69	7
	55	SleekMPv3 (+)	3.03	8		56	AlienwareTeleport	4.23	23
H: Medium size, squared shape skins	7	Anatta (+)	2.69	6	Q: Standard skins	9	10series (+)	4.85	32
	22	Pda (+)	1.99	2		18	9Seriesdefault (+)	4.50	32
	45	Slim (+)	2.83	1		43	Classic (+)	3.40	16
	64	FranzFerdinand	2.02	2		68	Corporate (+)	4.38	30
I: Complex interaction, with a lot of functionality	10	Cablemusic Player (+)	4.23	18					
	33	Erektorset	3.21	11					
	42	Kenwood KDC-X959 (+)	4.16	19					
	53	Main Street (+)	3.16	4					

* Specially created for the study. (+) Included in the aggregated category measure

The slides can be downloaded from http://people.brunel.ac.uk/~csstwpb/skintest2.pdf

This formula returns a zero if an equal number of participants rated a skin above and below their individual average (zero), and a one if all participants rated the skin either above or below their individual average (zero). The average agreement of all skins was 0.34 which is significantly ($t(67)=-7.30$, $p. <0.001$) below the 0.5 threshold to progress from a minority to a majority in agreement. Therefore, on average the participants were more in disagreement than in agreement about the rating of skins. This supports the idea of designing for diversity instead of similarity when developing multiple skins.

Instead of analysing correlations between the preference of individual skins and user characteristic, the analyses this time looked at the preference of skin categories. This approach increases the validity of the interpretation of the underlying factor as it studies the factor across a series of skins and thereby reducing the impact of other confounding factors that might be important for the rating of a particular individual skin. The first step therefore was to analyse the users' consistency in their rating of skins in a category by calculating the correlation between the ratings of skins within each category. The skins in a category that had a significant

positive correlation with other skins in that category are indicated by "+" in Table 4. The ratings of these skins were used in aggregated category measures by taking the average value of their z-scores. Categories B (purple/violet), D (bright), and F (large, dark), did not have positive correlating skin ratings, and were therefore excluded from the other analyses.

Gender and Age

Of the 130 participants in the second study, 83 were male, 44 were female and 3 did not specify their gender or age. The age of participants was between 18 and 49 years with a 21 years median. Table 5 shows a significant Pearson correlation between age and C-category, suggesting that older participants were more in favour of the grey skins than the younger participants. The other significant correlation was a negative correlation between age and the humoristic skins (L-category), which might suggest that ratings of older participants were less humour driven or that the humoristic reference presented in the skins was more appealing for younger participants.

Table 5 also shows significant Pearson correlations between gender and skin categories. The significant correlation between gender and the category K (friendly, charming 'baby' character) confirms the finding of the first study. Females rated these skins higher than the male participants. The females were also more in favour, compared to the males, of skins with small and friendly creatures (category J). Male participants on the other hand were more in favour of: small skins; medium size, squared shape skins; predator skins; and the scary skins (categories G, H, N, and O).

Table 5: Correlations between skin preference rating and the factors age and gender (female=0, male=1).

Factor	Category (correlation)
Age	C(.27**) L(-.29**)
Gender	G(.20*) H(.21*) J(-.31**)
	K(-.41)** N(.29**) O(.20*)

* sign. at a 0.05 level, ** sign. at a 0.01 level.

BIS/BAS

Table 6 shows the significant correlation between aggregated category rating and the personality scales. The table shows both the Pearson correlation and the partial correlation controlled for age and gender. No significant correlations were found this time for the BAS–Drive scale, associated with pursuit of desired goals. However, the BIS scale, associated with avoidance motivations, correlated with the colourful skins category (E) and with the friendly charming 'baby' character skins category (K).

Extraversion and Neuroticism

Reducing the IPIP-NEO personality inventory, by only focussing on the extraversion and neuroticism scale, meant that these questions could be included in the data gathering activity in the classroom setting, which resulted in 125 useful responses. The increased sample size addresses the problem mentioned earlier about the potential instability of the small data set of the first study. The significant correlation between colourful skins (E) and gregariousness is the only one to relate to the findings of the first study. However, this time the correlation was reversed, in that the more gregarious participants were less likely to select a colourful skin, than the less gregarious participants. They were also less likely to select a skin with a small and friendly creature (J). Furthermore, these gregarious participants and those scoring high on the assertiveness scale seem also more in favour of small skins (G), than participants that scored low on these scales.

Table 6: Correlations (*r*) and partial correlation (*pr*) (controlled for age and gender) between personality and categories of skin preference.

	Category (correlation)	
(Sub)Scale	**(*r*)**	**(*pr*)**
BAS (Drive)		
BIS	E(.17*) K(.28**)	K(.19*)
Extraversion		
-Gregariousness	**E(-.18*)** G(.21*) J(-.26**)	J(-.20*)
-Assertiveness	G(.22*)	
-Cheerfulness	M(.24**)	L(.20*) M(.22*)
Neuroticism		
-Depression	L(-.21*)	L(-.21*)

* sign. at a 0.05 level, ** sign. at a 0.01 level.

The similarity-attraction hypothesis seems to be confirmed by the findings of a positive correlation between action theme skins (M) and cheerfulness, but especially by the positive correlation between cheerfulness and humoristic theme skins (L), and the negative correlation between this category and depression. Whereas cheerfulness scales measures positive mood and feelings, the depression scale measures the tendency to feel sad and discouraged. The correlations with preference rating of humoristic skins, such as Monty Python and The Simpsons, seem to mirror these traits.

Factor analysis

As the categories were not designed to be mutually exclusive, an analysis of their internal relationship and the relation with the individual characteristics might help to understand the underlying factors of the skins ratings. Therefore, an explorative factor analysis was carried out on the aggregated category measures by means of the principle component extraction method, and resulting components were correlated with age, gender and personality.

The components extracted in the factor analysis were set to exceed an eigenvalue of 1. The results of the analyses reveal a solution with a Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy, an indicator of factorability, with a value of 0.59. This seems acceptable as it approaches the recommended 0.6 value

(Tabachnick and Fidell 2001). Next, the Bartlett's sphericity test was significant (p. < 0.001) indicating correlating categories. The results of the analysis revealed a solution with 4 components. Table 7 shows the factor loadings after varimax rotation of the categories on these 4 components. Instead of a fixed critical value, such as 0.4, Steven's (1996) recommendation was followed to relate it to the sample size by taking the critical value set for an ordinary Pearson correlation and doubling this.

Table 7: Rotated component loading matrix.

	Components			
Category	**1**	**2**	**3**	**4**
G: Small	-.78*	-.04	-.04	-.11
H: Medium size	-.66*	-.38	-.21	-.03
M: Action	.51*	-.09	.36	-.44
P: Mildly scary	.51*	.19	.23	.36
L: Humour	.05	.84*	.17	.18
K: Friendly character	.18	.67*	-.32	.33
C: Grey	-.39	-.54*	-.33	-.04
E: Colourful	.48*	.53*	-.10	.19
N: Predator	.19	.06	.79*	-.01
O: Scary	.26	-.15	.72*	.13
A: Blue	.14	-.37	-.57*	.02
J: Small and friendly creature	.28	.07	-.10	.77*
I: Complex functionality	.14	-.30	-.11	-.71*
Q: Standard	-.40	-.12	-.38	-.57*

* sign. at a 0.01 level.

Examining the loading matrix it seems that the first component relates to a shape and colour dimension of the skins. Categories G and H, related with small and medium size skins, but also with squared shaped skins, loaded negatively on this component, whereas categories M, P, and E loaded positive on the component. Although category M represents an action theme, and category P a mildly scary theme, examining the skins that make up these aggregated category measures, shows that none of these skins have a square shape, instead they have more smooth shapes. Next, they all use a variation of colours. Therefore, it seems that shape and colour is the common factor that links these categories. Table 8 shows that the male participants, more than the female participants, were in favour of colourful smoothly shaped skins. Next Table 9 shows also that this component also has a negative correlation with gregariousness, which agrees with the previous finding about the correlations between gregariousness and colourful (E) and small size skins (G) categories.

The second component seems to represent a skin dimension that could be described as soberness versus playfulness. Categories L (humoristic theme), K (friendly charming 'baby' character), and E (colourfulness) correlated positively with this component, whereas category C (grey skins) load negatively on this component. This second component correlated negatively with age, and the playful side was also more preferred than the sober side by the female than by the male participants. Next, regarding participants' personality, the component correlated positively with extraversion and its subscales gregariousness and cheerfulness, but negatively with depression (Table 9).

Table 8: Pearson correlations between components of skin preference categories and age (years), and gender (female=0, male=1).

Factor	**Category (correlation)**
Age	2(-.31**)
Gender	1(-.20*) 2(-.18*) 3(.35**) 4(-.19*)

* sign. at a 0.05 level, ** sign. at a 0.01 level.

The third component of the factor analysis can be interpreted as calm versus scary. Table 7 shows that category A (blue skins) loaded negatively on this component, while categories N (predator) and O (scary skins) loaded positively on the component. The positive side of the component was more appreciated by the male participants, while the female participants were more drawn toward the negative (calm) side of the component. The component also has a negative correlation with the BIS scale, which is related with the aversive motivational system that tries to inhibit behaviour that may lead to negative or painful outcomes (Carver and White, 1994).

Table 9: Correlations (*r*) and partial correlation (*pr*) (controlled for age and gender) between personality and components of skin preference categories.

	Component(correlation)	
(Sub)Scale	**(*r*)**	**(*pr*)**
BIS	3(-.21*)	
Extraversion	4(-.21*)	2(.20*) 4(-.19)
-Gregariousness	1(-.18*) 4(-.27**)	2(.20*) 4(-.23**)
-Cheerfulness	2(.21*) 4(-.24**)	2(.27**) 4(-.19*)
Neuroticism		
-Depression	2(-.18*)	2(-.20*)

* sign. at a 0.05 level, ** sign. at a 0.01 level.

The last component correlated positively with category J (small and friendly creature) and negatively with categories I (complex interaction with a lot of functionality) and Q (standard skins). Although the interpretation is less clear, it seems to represent polished versus sloppy, since the skins in category Q and the skins included in the aggregated measure of category I, seems all very sophisticated and stylish, while skins in category J are more cosy and loose. Table 8 shows that the female participants where more drawn to the negative side of the component, while the male participants more towards the positive side. The component also correlated with personality traits, such as negatively with extraversion and its subscale gregariousness and cheerfulness.

Colour

The skin set used in the second study also had two sub sets of skins that were identical except for their colour. These were the Colorchooser skins (12, 29, 65, 48) and Color + OS skins (4, 21, 32, 38). An analysis was conducted on these skins, to see whether the correlation between the preference rating and individual characteristics changed as a function of colour. Two MANCOVA's with repeated measures were conducted which took: as dependent variable the raw preference rating of either the Colorchooser or Color+OS skins; as covariates age, and the personality scales; and as independent variables gender. The results reveal a significant two-way interaction effect ($F(3,94) = 3.14$, $p. = .029$) for colour and BAS-Drive scale on the skin rating of the Colorchooser.

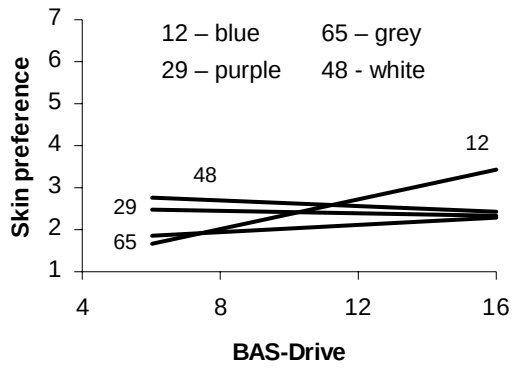

Figure 1: Regression lines predicting skin preference of the Colorchooser skin set from BAS-Drive scale.

Examining the individual correlations between the z-scores of the skin rating and BAS-Drive scale reveals a positive correlation ($r = .20$, $p. = .028$) between the blue version of the Colorchooser and BAS-Drive scale. As Figure 1 of the fitted linear regression lines shows, an increase of the BAS-Drive scale coincides with an increase in the rating of the blue skin, while the rating for the other skins remain more or less equal. However, the results of other MANCOVA on the rating of the Color+OS skin set, revealed no significant two-way effect ($F(3,89) = 2.3$, $p. = .083$) for colour and BAS-Drive. Examining the individual correlation with the z-scores of these skins reveals a relatively small, but again not significant correlation with the blue skin version of Color+OS ($r = -.16$, $p. = .064$). Note however that this time the correlation is in the opposite direction.

CONCLUDING REMARKS

In conclusion, the data collected in these two studies indicate that designing skins for specific user groups might be possible and more effective: possible because the results suggest several relationships between skin preference and user characteristics, such as gender and personality; and effective because so much diversity was found between the users preference of skins. Although the factor analysis on the data of the first study suggests that users' preference for skins could not be explained by a limited number of factors, the factor analysis on the data of the second study provide some ideas for underlying preference dimensions for some skins. The correlations between these dimensions and personality traits also support the similarity-attraction hypothesis, such as the correlation between BIS and a 'calm versus scary' dimension and extraversion and a 'soberness versus playfulness' dimension.

Still, the correlations between personality traits and skin preferences in the second study were all relatively small, which however is not uncommon when predicting behaviour directly from personality traits. Factors such as social pressures and perceived control of the technology might have been factors interfering with a direct relation. The Theory of Reasoned Action (Ajzen and Fishbein, 1980) and its successor, Theory of Planned Behaviour, seems to be interesting candidates for further research in this context, as they try to predict behaviour from a person's intention by looking at his/her attitude and the perceived social norm. These again are influenced by personality traits according to these theories.

REFERENCES

Ajzen, I. & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs : Prentice-Hall.

Buchanan, T., Johnson, J. A., & Goldberg, L. R. (2005). Implementing a five-factor personality inventory for use on the internet. *European Journal of Psychological Assessment.* 21, 115-127

Carver, C.S., & White, T.L. (1994). Behavioral inhibition, behavioural activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319-333.

Karsvall, A. (2002). Personality preferences in graphical interface design. *Proceedings NordiCHI*, 19-23.

McCarthy, J. & Wright, P.C. (2004). *Technology as experience*. Cambridge, MA: MIT press.

McCrae, R.R. & Costa, P.T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90.

Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, D. C. (1995). Can computer personalities be human personalities? *International Journal of Human-Computer Studies*, 43, 223-239.

Norman, A. D. (2004). *Emotional design: why we love (or hate) everyday things.* New York: Basic Books.

Steven, J., 1996, *Applied multivariate statistics for the social sciences* (Mahwah, NJ: Lawrence Erlbaum).

Tabachnick, B.G., & Fidell, L.S. (2001), *Using multivariate statistics* (Boston, MA: Allyn & Bacon).