Space, color, and perceived qualities of indoor environments

Gerald Franz
Max Planck Institute for Biological Cybernetics, Tübingen, Germany
gerald.franz@tuebingen.mpg.de

Abstract
When relating productivity measures to qualities of indoor environments, emotional concepts play an important role as intermediating variables (e.g., arousal in the classic Yerkes-Dodson law or workplace satisfaction). While several studies have analyzed influences of individual aspects of the physical environment, little is known about the interplay of multiple factors. Therefore the presented internet-based study exemplarily addressed both individual and interactive effects of room colors and primary spatial properties on affective responses to the presented environments. A comparison between affective appraisals and quantitative scene descriptions found strong correlations for all rating categories (overall explained variance $.43<R^2<.86$). In seven of eight rating categories a simple super-imposition of factors appeared to be the most plausible model when comparing the integrative experiment to corresponding single factor studies. Yet also the detected differences appear meaningful, suggesting that reliable real world predictions require a proper operationalization of baseline levels and context.

Keywords: Affective appraisals; color; space; integration; web experiment

Introduction

It is an ongoing debate to what degree indoor environmental quality related pathologic phenomena such as sick building syndrome can be mainly ascribed to the air quality or to the psychosocial work conditions (e.g., recently Marmot, Eley, Stafford, Stansfeld, Warwick, & Marmot, 2006; Mendell & Fisk, 2006). Despite this dispute concentrates on mainly two primary groups of factors, several recent studies have also highlighted the likely relevance of indirect psychological phenomena such as general workplace satisfaction (Zapf, 1991) and attitudes towards the workplace (Bischof, Bullinger-Naber, Kruppa, Müller, & Schwab, 2004). Also classic studies and theories on productivity such as the Yerkes-Dodson law (1908) rather stress the importance of emotional aspects, especially the general arousal level. It seems therefore well arguable that truly comprehensive measures of indoor environmental quality should equally take further physical factors affecting these emotional variables into account.

Indeed, numerous studies could demonstrate significant correlations between single perceptual properties and affective responses to environments (e.g., see the compilations of Mehrabian & Russell, 1974; Stamps, 2000). However, despite this substantial body of research, it is still infeasible to come up with reliable predictions on the likely contribution of visual and aesthetic properties on affective responses to work places. Besides the multitude of potentially relevant factors, a main reason for this is that, due to methodological reasons, most studies concentrated on selected individual factors and kept others fixed or counteracted potential interactions by the experimental design. However, in real world buildings many factors are parallelly present and lead to coherent and distinct experiences. Beyond a basic awareness of this, little is known about the interaction and integration of different factors that actually takes place when humans experience architectural environments. Hence, this question was exemplarily addressed in this study covering the basic design aspects room color and spatial dimensions, which both are known to be individually relevant for the experience of indoor spaces.

An empirical investigation of interactions of multiple factors requires an experimental design varying them parallelly over a set of stimuli, implicating a substantial increase of the necessary trials and scenes as compared to single factor studies. Therefore, the main experiment was designed as an internet-based study. In order to make sure that the observed main effects were not just an artifact of these rather loosely controlled experimental conditions, the study was accompanied by two well-controlled laboratory experiments that provided baseline conditions on the individual effects of the two design aspects.

Background

Emotion and affect. Emotion or synonymously affect is a diverse and complex psychic phenomenon comprising physiological responses by the vegetative nervous system, involuntary moods and feelings, as well as cognitions on these internal states and the related external context (Russell & Snodgrass, 1987). For practical empirical purposes, in particular the framework of Mehrabian and Russell (1974) has turned out to be suitable. It describes emotions along three basic underlying dimensions (pleasure/valence, arousal, and dominance) that widely imply related concepts such as preference or beauty. Their framework allows an effective quantification of affective responses using introspective verbal scaling techniques. These affective appraisals characterizing the mood altering capacity of a stimulus are seen as especially important for attitudes and decisions (Russell & Snodgrass, 1987), and likely provide at least an approximation to further e.g., physiological responses (Russell, 1988). Furthermore, several studies summarized by Stamps (2000, pp. 114-138) strongly suggest that, despite obviously existing individual differences, averaged appraisals indicate meaningful and stable main trends offering a basis for generalizable predications.

Emotional responses to color in architecture. Despite a large body of respective text books and general literature (e.g., most comprehensively Nemcsics, 1993), it is still difficult to subsume the current state of empirically backed knowledge on the influence of colors on emotions. The current state of knowledge mainly allows for predications concerning the affective dimensions arousal and dominance. Generally, warm and saturated colors are seen as more arousing, whereas the dominance dimension corresponds to differences in light-dark. These initially phenomenal observations are well backed by physiological data (Küller, 2001). As regards valence, mainly positive relations to brightness and saturation have been reported, whereas effects of color hues are secondary (Mehrabian & Russell, 1974, pp. 56-59) and probably depending on the color-bearing object category. Related to that, tendential preferences for more saturated and darker floor colors have been reported, whereas for ceilings lighter and less saturated colors are preferred (Frieling, 1974).

Emotional responses to architectural space. Although spatial form is traditionally seen as the primary dimension of architecture (Giedion, 1941), little is known about the influences of spatial properties on the affective experience of architecture beyond the mere fact, that they have an enormous emotional potential, as apparent in space-related phobias. Primary affective dimensions of spaces are probably their size (Joedicke, 1985) and rate of enclosure (Franz, von der Heyde, & Bülthoff, 2005a; Stamps, 2005). Humans seem to have preferential tendencies for spacious rooms that provide vistas into their surrounding, and offer at the same time some sort of near space protection (Appleton, 1988; Newman, 1996). However, all these findings and theories are still not sufficiently empirically backed and integrated, for example their context dependency, transferability, and generality still has to be investigated. Finally, phenotypically oppositional space related psychoses such as claustrophobia and agoraphobia suggest that individual predisposition and context may play dominant roles regarding reactions to particular spaces.

With respect to the emotional dimension of arousal, Franz, von der Heyde, & Bülthoff (2005b) have found that the absolute number and relative density of spatial differentiations are promising predictor variables. Their observations fit well into more general information theory related frameworks of environmental preferences (Berlyne, 1972) and findings on building facades (Stamps, 2000). Regarding the influences of spatial form on the third dimension dominance, common architectural practice suggests that an increased scale are an effective means to increase the dominance of a building. Furthermore, the recurrence of symmetries and particularly repetitions in totalitarian architecture likely indicates that also these properties contribute to the experienced dominance of a space.

Interactions between color and space perception. Since both spatial dimensions and colors are primarily visually conveyed, their perception and probably also experience is not completely independent. Indeed, normative architectural knowledge provides for some rules of thumbs concerning the influence of colors on perceived room dimensions. Generally, cool, desaturated, and light colors are considered to increase the experienced spaciousness, whereas dark, saturated, and warm colors tendentially have opposite effects. These characterizations fit well to the at least qualitatively well documented phenomenon of color perspective (e.g., Nemcsics, 1993; Bailey, Grimm, & Davoli, 2006). Analogously, in applied color design it is recommended to make extensive use of saturated dark only in large rooms (Frieling, 1974).

Integration of multiple factors into emotional responses. While there are several theories on the influence of single factors on affective responses, theoretic frameworks or empirical findings explicitly addressing the interaction of environmental factors influencing affective responses are rare. However, single factors studies provide at least a basis for specific hypotheses: Often significant linear correlations between rated affective qualities and multiple variables became apparent (e.g., Nasar, 1988; Stamps, 2005). In other words, these statistical observations suggest the existence of at least partially independent linear terms. Therefore, one simplified model could be termed superimposition without interactions. This hypothetic models gains some further support by an analysis of the distribution of individual affective appraisals (cf. e.g., Franz, 2005, pp. 172-175) that usually approximate a normal distribution. One popular explanation of the often recurring Gaussian distributions of real world phenomena basically interpets them as sums of multiple independent underlying factors.

On the other hand, some empirical results also point in a different direction. Most prominently, the theories and observations of Berlyne (1972) on interdependencies of emotional dimensions have lead to the famous inverted U-shaped conceptual relation between arousal and valence, indicating a primary importance of absolute levels at least for valence. Furthermore, Franz et al. (2005b) found indications that in case of an extreme variance single environmental factors (e.g., area) may determine the overall response alone. This finding gains support by two further observations: First, in colloquial language architectural spaces are often characterized by using just single outstanding features (Franz, von der Heyde, & Bülthoff, 2002). The second argument arises from the notion that central aspects of emotion can be essentially understood as immediate and highly integrated evaluative responses to a stimulus within a given context (Pessoa, 2005). While such integrative responses should normally consider multiple factors at the same time, in extreme cases the consideration of just one dominating factor appears to be more reasonable than a mediation by further secondary aspects for behavior control. In sum, it seems well arguable to expect a majority of linear superimpositions; in certain cases, however, the overall response may mediated or determined by single factors.

Objectives

Overall, the study aimed at an exemplary investigation of the interaction of multiple factors from different design aspects contributing to the experience of architectural space. More specifically, the hypothesis was raised and tested that within the normal ranges of variance overall affective responses can be seen as a superimposition of independent individual factors. For this exploratory purpose, the interaction of spatial properties and surface properties was selected, because of their ubiquity and the ease to manipulate them independently.

Method

Experimental design. The study consisted of an integrative web experiment and two well-controlled comparative conditions. In order to detect the hypothesized superimposition of factors in the integrative main experiment, a number of 64 scenes seemed to be appropriate for a quasi-orthogonal variation of factors and a detection of presumably four medium to large effects. This large number of scenes did not allow for a within-subject design, because experimental sessions of about two hours had inevitably lead to strong global order effects due to fatigue. Consequently, the complete set had to be distributed over several participants.

Since absolute rating scores tend to be adjusted to the sample (Russell, 1988), no fixed subsamples were used, but individual random subsamples were presented. Due to the large number of stimuli and necessary participants, the study was designed as an internet-based experiment (Reips, 2002; van Veen, Bülthoff, & Givaty, 1999). Since the variables of interest addressed in this experiment were the general directions and relative magnitudes of relations, a likely moderate increase of noise in the data due to the loosely controlled presentation conditions was seen as acceptable. The controlled randomization algorithm made sure that all scenes were rated by a similar number of participants. In addition, systematic biases induced by uncalibrated monitors of the participants were thus counteracted.

For the exploratory purposes of this study, the semantic differential as simple and widely used method for the approximative quantification of emotional responses was chosen (Osgood, Suci, & Tannenbaum, 1957; Heise, 1970). The combination of pairs of oppositional adjectives with a seven step Likert scale appears to be well suitable to capture the three primary dimensions (pleasingness corresp. to valence, arousal, and dominance) of emotion as occurring in affective appraisals (Mehrabian & Russell, 1974; Russell, 1988) (cf. Table 1). In addition, ratings of openness and spaciousness were collected in order to detect possible interactions between room color and experienced room dimensions.

Dimension	Category	English low extreme	English high extreme	German low extreme	German high extreme
valence	pleasingness	unpleasant	pleasant	unangenehm	angenehm
	beauty	ugly	beautiful	hässlich	schön
arousal	excitement	calming	exciting	beruhigend	aufregend
	interestingness	interesting	boring	langweilig	interessant
dominance	obtrusiveness	inobtrusive	obtrusive	zurückhaltend	aufdringlich
	gravity	light	oppressive	leicht	drückend
spatiality	spaciousness	narrow	spacious	eng	weit
	enclosure	open	enclosed	offen	geschlossen

Table 1: Semantic differential rating categories used in the experiment.

Procedure. Before the experiment, participants were briefed by an introductory web site on the general purpose of the experiment and its operation. Afterwards, personal and technical data were collected on a voluntary basis using a web-based questionnaire form. Then the participants proceeded to the actual experiment. On the left-hand screen side, eight semantic differential scales were displayed in random order. The actual stimuli were presented on the central and right half of the screen at a resolution of 640x480 pixels, they were displayed as panorama images using a small Java applet (PTViewer by Helmut Dersch, see http://webuser.fh-furtwangen.de/~dersch/). Subjects could freely choose their gaze direct by dragging the mouse in the display window. The experiment instructions explicitly advised them to familiarize themselves with the interface and to examine the complete room carefully before rating the scenes. Regarding the ratings, subjects were instructed to respond quickly and to adhere to their first impression in case of indecisiveness. Only after rating all eight categories, subjects could proceed to the next scene. Each participant was asked to rate a subsample of 23 scenes, yet, due to the design as web experiment, they were free to terminate the experiment at any time before. A complete experimental session took around 45 minutes.

Stimuli. The experiment was based on 64 vacant normal rectangular rooms (see Figure 1). The stimuli were automatically generated by the SceneGen tool (Franz, 2005, p. 95). Architectural elements were varied in a balanced manner in realistic ranges. The wall color was systematically varied by dividing the HSV color space in 16 equidistant hue bands and using 4 saturation and brightness levels. The 64 colors were randomly assigned to the spaces, leading to very low correlations between color and spatial descriptors ($r^2_{max} < .08$). Ceilings were uniformly covered by a white ingrain texture, the floor was textured using naturally wooden square parquet. The indoor scenes were surrounded by an urban background texture.

Figure 1: Screenshots of a 16 out of 64 scenes used in the main experiment showing the variability as regards architectural elements, dimensions, and color.

Analysis. From the scenes characteristic values were derived based on the numbers and dimensions of the architectural elements (floor, wall, windows, and doors). The wall color was analyzed in two color spaces using the phenomenologically oriented HSV and the psychophysics-based Lαβ model. In order to detect correlations with individual color hues, HSV hue values were linearized between complementary colors in steps of 45 degrees. Initially, 44 descriptor variables were calculated. A factor analysis addressing internal correlations between descriptors reduced this number to 18 (7 spatial, 11 color descriptors) that were much less interrelated ($r^2_{mean} = .08$).
The statistical analysis was primarily based on multivariate linear regressions between semantic differential ratings averaged over all participants and characteristic values derived from the scenes. The applied regression algorithm tested for sets of significant regressor variables using a forward strategy, always the variable rendering the strongest significant ($p<.05$) partial correlation was added to the regressor set. Additionally, nonlinear relations were evaluated qualitatively via fitting square regression functions. Since no significant non-linear interrelations became apparent in the main experiment, the following section solely reports the results of the linear analysis.

Results

Main experiment. Altogether, 91 participants (37 female, 51 male, 3 anonymous) contributed to the web experiment. Each scene was assessed at least eight times, on average 11.8 independent ratings contributed to get the characteristic values. The variance within the ratings for each scene was fairly constant over all rating categories, the standard deviation was around 1.4 and clearly below an equal distribution level of 2.16 on a 7 step Likert scale, indicating mainly unimodal rating distributions and therefore meaningful mean values. The one-way analysis of variance (ANOVA) ascribed between 17% (beauty) and 41% (obtrusiveness) of the differences in the individual ratings to the differences in the stimuli.

None of the recorded differential factors (gender, academic background/profession, personal identity) was able to significantly explain further variance in the ratings.
Generally, the factors entered in the multivariate linear regression explained between 43% (pleasingness) and 86% (spaciousness) variance of the averaged ratings (see Table 2). For all rating categories except of interestingness, factors derived from the room geometry as well as factors describing the wall color significantly contributed to the overall result. Often recurring color-based regressors were Lαβ luminance and the saturation of the Lαβ α dimension (i.e. the intensity of red/green hues). From the spatial descriptors, mainly openness related descriptors and room proportions contributed significantly to the explained variance.

Rating Category	Explained variance R^2	Regressor variables and individual correlations
Integrative main experiment		
pleasingness	.43	saturation Lαβ α(r=-.66, p<.001), wall openness (r=.41, p<.001)
beauty	.54	saturation Lαβ α (r=-.52, p<.001), wall openness (r=.47, p<.001)
excitement	.78	saturation Lαβ α (r=.77, p<.001), Lαβ luminance (r=-.65, p<.001), room proportion length/width (r=.18, p=.16), HSV hue red to green (r=.3, p=.01), HSV saturation (r=.73, p<.001)
interestingness	.45	HSV saturation (r=.64, p<.001), HSV hue red to green (r=.23, p=.07)
obtrusiveness	.83	saturation Lαβ α (r=.81, p<.001), HSV saturation (r=.80, p<.001), HSV hue violet to yellow-green (r=.26, p=.04), room proportion length/width (r=.13, p=.31)
gravity	.75	Lαβ luminance (r=-.59, p<.001), wall openness (r=-.52, p<.001), room area (r=-.48, p<.001), saturation Lαβ α (r=.58, p<.001), balustrade height (r=.39, p<.01), HSV hue cyan to orange (r=-.06, p=.66)
spaciousness	.86	room area (r=.83, p<.001), wall openness (r=.55, p<.001), room proportion length/width (r=-.40, p<.01), Lαβ luminance (r=.27, p=.03), balustrade height (r=-.32, p=.01)
enclosure	.70	wall openness (r=-.66, p<.001), room area (r=-.55, p<.001), Lαβ luminance (r=-.38, p<.01), balustrade height (r=.54, p<.001)
Comparative study on spatial dimensions		
pleasingness	.72	openness (r=.84, p<.001)
beauty	.75	openness (r=.87, p<.001)
excitement	.70	balustrade height (r=-.75, p=.001), openness (r=.63, p<.01)
interestingness	.79	openness (r=.76, p=.001), balustrade height (r=-.69, p<.01)
spaciousness	.91	room area (r=.84, p<.001), balustrade height (r=-.28, p=.29), room proportion length/width (r=-.35, p=.19)
enclosure	.82	openness (r=-.83, p<.001), balustrade height (r=.61, p=.01)
Comparative study on room colors		
pleasingness	.30	HSV saturation (r=-.54, p=.001)
beauty	.00	-
excitement	0.85	HSV saturation (r=.90, p<.001), HSV hue red to green (r=.18, p=.34)
interestingness	0.89	HSV saturation (r=.92, p<.001), HSV hue violet to yellow-green (r=.19, p=.30)
obtrusiveness	0.91	HSV saturation (r=.93, p<.001), HSV hue red to green (r=.18, p=.33)
gravity	0.64	Lαβ luminance (r=-.80, p<.001)

Table 2: Results multivariate regression analysis of the three experiments.

Comparative experiment varying solely spatial dimensions. In this precedent exploratory study (described in detail in Franz et al., 2005a) quantitative relations between the experience of architectural indoor spaces and quantities and dimensions of their architectural elements were explored. Unlike the main study, vacant indoor spaces uniformly showing white walls were used as stimuli and no dominance related affective appraisals were collected.
The experimental design and procedure was very similar to the main study, 16 participants (8 female / 8 male) rated 16 virtual rooms using the semantic differential scaling technique. The laboratory experiment used as display device a spherical wide-angle projection system (Elumens VisionStation™) offering a resolution of 1024x768 pixels. The geometrical field of view (FOV) matched the physical FOV of about 130x90 degrees. A comparative analysis

between studies by Franz et al. (2005b) and Wiener & Franz (2005) suggested that directions and magnitudes of correlations are comparable between these two setups.
The one-way ANOVA ascribed between 16% (excitement) and 63% (spaciousness) of the variance in the individual ratings directly to the differences between the stimuli. Generally, significant and strong individual linear correlations were found for all rating dimensions. Altogether, four widely independent linear factors (openness, room proportion length/width, room area, and balustrade height) turned out to be effective for statistically explaining a major share of the observed variances in the mean ratings ($.70 < R^2 < .91$, cf. Table 2). Furthermore, a quadratic regression analysis revealed considerable non-linear relations between valence ratings and room proportions and showed maxima near to the golden section (e.g., beauty over room length/width maximum at ratio 1.7, beauty over room width/height maximum at ratio 1.5).
Comparative study varying solely colors. The second comparative study tested for influences of room color on affective responses in a constant spatial environment under laboratory experimental conditions. Despite the different levels of control, similar relations between color dimensions and affective appraisals as in the main experiment were expected. In order to test this hypothesis, the wall color of one virtual room having the average features and dimensions of the scene set of the main experiment was systematically varied, whereas shape, features, and dimensions were constant.
An equally distributed subset of 32 colors of the main study was selected and used as wall color of the virtual indoor scenes. The stimuli were presented in a psychophysics laboratory offering constant and completely controlled laboratory conditions on a 21" CRT screen whose geometry and color rendering was carefully adjusted. Spectroscopic measurements allowed to translate the HSV values of the wall colors into CIE XYZ chromaticity values that were the basis for the analysis in the perceptually oriented Lαβ color space. The restriction to 32 scenes allowed a within-subject design, all 18 participants (9 female, 9 male) rated the complete scene set using the same six rating categories for valence, arousal, and dominance as the main experiment. A complete experimental session took about 40 minutes.
The one-way ANOVA attributed between 9% (beauty) and 60% (obtrusiveness) of the variance in the individual ratings to the stimuli. Analogous to this considerable difference, the multivariate regression analysis (Table 2) rendered an uneven result between the rating categories. Whereas the analysis was capable of explaining a large share of variance of the arousal and dominance related averaged ratings ($R^2>.64$, highest individual correlations to HSV saturation, resp. Lαβ luminance), only a minor fraction of the mean valence ratings could be ascribed to the entered linear regressors. In particular no significant correlation between mean beauty ratings and any of the HSV and Lαβ based descriptors of the main study could be found, although beauty ratings were strongly correlated to pleasingness ratings ($r=.81$, $p<.001$). An exploration of correlations to further color descriptors found also significant correlations of moderate magnitude to variables derived from the RGB color space (beauty - RGB green saturation $r=-.39$, $p=.025$, beauty – RGB red intensity $r=-.43$, $p=.01$).

Discussion

Comparison main experiment - baseline condition color. The most surprising result when comparing the outcomes of the main study to the comparative color experiment is the apparent discrepancy as regards beauty ratings. One main reason for the lack of corresponding regressor variables seems to be the low proportion of variance of beauty ratings attributable to the differences in the stimuli. Since a regression analysis of each participant individually did not render convincing indications for systematic differences between them, it seems most probable that aesthetic judgments of colors are either less consistent than other affective appraisals or more dependent on the individual context (e.g., contrast to the precedent stimulus). Apart from this notable discrepancy, the results of the comparative color experiment almost completely corresponded to the main study. The directions and magnitudes of correlations between scene features and ratings were uniformly very similar ($r=.90$, $p<.001$, cf. Figure 2). As only further systematic difference between the studies, the effect sizes of color variations were significantly higher in the single factor study. This can be well explained by the absence of spatial variations as further sources of variance.

Comparison main experiment - baseline condition spatial dimensions. Also the comparison between the integrative experiment and the single factor study on spatial dimensions rendered mainly corresponding results in most rating categories. Yet in contrast to the single factor study, no evidence for nonlinear relations to room proportions could be found in the integrative experiment. It remains unsolved whether this relation was just widely masked by the various additional factors, or whether the finding in the single-factor study was mainly an artifact, or even a peculiarity of monochrome white rooms. Furthermore, the comparative data analysis revealed two at first glance surprising differences regarding openness and arousal: In the single factor study, openness related measurands such as wall openness ratio and balustrade height were highly correlated with rated interestingness, whereas correlations were close to zero in the integrative study. Since in the latter the level of explained variance with respect to interestingness was comparatively low, this suggested that some important feature was not adequately captured by the scene descriptions. Therefore, the scenes were sorted according to interestingness in order to tentatively identify some not yet formalized patterns. Yet besides color properties, differences with respect to openings seemed indeed to be a main distinguishing feature. However, this inference also evoked the notion that absolute openness was not the actually decisive factor but rather the quality of the vista into the urban context, which depended mainly on the positions of the windows. Hence, accurate predictions of likely influences of openings therefore probably also require a consideration of the qualities of the surroundings, whose formalization is definitely more complex than a simple local openness value.

Second, as regards correlations between openness and rated excitement, even opposite tendencies became apparent (single factor study strong positive correlation $r=.63$, $p<.01$, integrative study negative tendency $r=-.20$, $p=.12$). One tentative interpretation of this arose from the inspection of the scene sets as a whole: As compared to the vacant white indoor walls, the moderately colored and variable urban background scene was indeed relatively exciting, whereas in relation to an intensively colored room it was rather calm. Therefore, both the absolute arousingness of the scene as well as the general arousal level induced by the experiment as a whole could be underlying factors influencing this difference.

Implications on the integration of design aspects. Generally, the vast majority of the individual observations fitted well to the hypothesized linear superimposition of multiple factors: In all tested experiential dimensions significant linear terms based on color as well as on spatial properties were found in the main study. It is particularly remarkable that wall color luminance apparently also significantly affected the ratings of spaciousness and enclosure. In direct comparison, color properties appeared to be tendentially more influential on the ratings than spatial properties in this study. Instead of assuming a stronger potency if color, this may be rather explained by the different degree of variability of the two design aspects: While room features were varied within usual ranges, colors were systematically varied with respect to the HSV color space, thereby including several values that are rather unusual for indoor environments.

In addition to this predominant pattern of correspondences between the conditions, also the observed differences appear to be informative as regards the overall formation of affective responses to architectural environments. In order to proceed from qualitative statements on general tendencies to actual quantitative predictions, the observations on arousal and openness might indeed indicate that actual baseline levels need to be taken into account. Affective appraisals of stimulus aspects may be both related to the direct perceptual context as well as to the more general experiential background (Russell & Lanius, 1984). Although this is purely speculative, also the comparatively modest success in explaining valence ratings may be tentatively ascribed to this context dependency: The findings of Berlyne (1972) suggest a close relation between the absolute arousal level and the relative valence of stimulus aspects. If arousal levels are not sufficiently operationalized or widely constant, an exact predication of valence responses gets therefore more difficult.

Conclusions

The presented study investigated individual and interactive effects of spatial properties and colors on affective responses to architectural environments, in order to predict their likely conjoint contribution to the general indoor environmental quality. Taken together, the main experiment and its comparative baseline conditions covered appraisals by 125 participants, 112 indoor scenes, and a notable breadth of factors. Despite the exploratory character of the study and analysis, a considerable proportion of variance could be ascribed to objectively measurable color descriptors and spatial properties. Most statistically inferred relations had direct correspondences between the single factor studies and the integrative web experiment, supporting both the reliability of methods and the likely validity of findings obtained in exploratory single factor studies. Among the investigated properties, in particular color saturation levels and openness measurands turned out to be primary factors affecting emotional responses. While the majority of individual observations suggest a simple superimposition of individual factor as most probable general case, also the detected differences between the conditions appear instructive, indicating that reliable real world predictions may require a proper operationalization of baseline levels and context, factors which clearly should be addressed in more detail in future studies.

All in all, the study strongly supports the notion that systematic investigations of affective qualities of indoor environments are feasible and have the potential to finally result in qualified predictions. Continuing this line of applied research will hopefully help to design buildings that comprehensively respond to human needs, and therefore contributes to the creation of successful and sustainable indoor environments.

Acknowledgements

This work was supported by the Max Planck Institute for Biological Cybernetics, Tübingen. The author expresses his thankfulness to Professor Dr. Heinrich H. Bülthoff for his general support. Many thanks also to Dr. Harald Teufel for providing an invaluable software tool that allowed an automatic calculation of the correct RGB to XYZ color conversion matrices and gamma factors based on a series of spectroscopic measurements.

References

Appleton, J. (1988). Prospects and refuges revisited. In *Environmental aesthetics: Theory, research, and application* (pp. 27–44). New York: Cambridge University Press.

Bailey, R. J., Grimm, C. M., & Davoli, C. (2006). The real effect of warm-cool colors. Tech. rep. WUCSE-2006-17, Department of Computer Science And Engineering -Washington University in St. Louis.

Berlyne, D. E. (1972). *Aesthetics and Psychobiology*. New York: Appleton.

Bischof, W., Bullinger-Naber, M., Kruppa, B., Müller, B. H., & Schwab, R. (Eds.). (2004). *Expositionen und gesundheitliche Beeinträchtigungen in Bürogebäuden [Exposures and health impairments in office buildings]*. Stuttgart, Germany: Fraunhofer IRB Verlag.

Franz, G. (2005). *An empirical approach to the experience of architectural space.* Ph.D. thesis, Bauhaus-Universität, Weimar, Germany. Available online at http://www.kyb.mpg.de/publications/pdfs/pdf3464.pdf.

Franz, G., von der Heyde, M., & Bülthoff, H. H. (2002). Evaluating architectural interiors with everyday terms. In H. H. Bülthoff, K. R. Gegenfurtner, H. A. Mallot, & R. Ulrich (Eds.), *Beiträge zur 5. Tübinger Wahrnehmungskonferenz*, p. 165 Kirchentellinsfurt, Germany. Knirsch-Verlag.

Franz, G., von der Heyde, M., & Bülthoff, H. H. (2005a). An empirical approach to the experience of architectural space in virtual reality: Exploring relations between features and affective appraisals of rectangular indoor spaces. *Automation in Construction*, **14**(2), 165–172.

Franz, G., von der Heyde, M., & Bülthoff, H. H. (2005b). Predicting experiential qualities of architecture by its spatial properties. In B. Martens & A. G. Keul (Eds.), *Designing Social*

Innovation: Planning, Building, Evaluating, pp. 157–166 Cambridge, MA. Hogrefe and Huber.
Frieling, H. (1974). *Farbe im Raum: Angewandte Farbenpsychologie [Color in space: Applied color psychology]*. München, Germany: Callwey Verlag.
Giedion, S. (1941). *Space, Time and Architecture*. Cambridge, MA: Harvard University Press.
von Goethe, J. W. (1810). *Zur Farbenlehre [Color doctrine]*. Tübingen, Germany: Cotta.
Heise, D. R. (1970). The semantic differential and attitude research. In G. F. Summers (Ed.), *Attitude Measurement* (pp. 235–253). Chicago: Rand McNally.
Joedicke, J. (1985). *Raum und Form in der Architektur : Über den behutsamen Umgang mit der Vergangenheit [Space and Form in Architecture].* Stuttgart, Germany: Kraemer.
Küller, R. (2001). The architectural psychology box of infinite knowledge. In *Aesthetics, well-being and health: essays within architecture and environmental aesthetics* (pp. 129–142). Aldershot, UK: Ashgate.
Marmot, A. F., Eley, J., Stafford, M., Stansfeld, S. A.,Warwick, E., & Marmot, M. G. (2006). Building health: an epidemiological study of "sick building syndrome" in the Whitehall II study. *Occupational and Environmental Medicine*, **63**(4), 283–289.
Mehrabian, A. & Russell, J. A. (1974). *An Approach to Environmental Psychology*. Cambridge, MA: MIT Press.
Mendell, M. J. & Fisk, W. J. (2006). Is health in office buildings related only to psychosocial factors? Electronic letter to *Occupational and Environmental Medicine*, available online at http://oem.bmjjournals.com/cgi/eletters/63/4/283#308.
Nasar, J. L. (1988). Perception and evaluation of residential street scenes. In *Environmental aesthetics: Theory, research, and application* (pp. 275–289). New York: Cambridge University Press.
Nemcsics, A. (1993). *Colour Dynamics: Environmental Colour Design*. London, UK: Ellis Horwood.
Newman, O. (1996). *Creating Defensible Space.* Washington: US Department of Housing and Urban Development.
Osgood, C., Suci, G., & Tannenbaum, P. (1957). *The Measurement of Meaning.* Urbana, IL: University of Illinois Press.
Pessoa, L. (2005). To what extent are emotional visual stimuli processed without attention and awareness?. *Current opinion in neurobiology,* **15**(2), 188–196.
Reips, U. D. (2002). Standards for internet-based experimenting. *Experimental psychology*, **49**(4), 243–256.
Russell, J. A. (1988). Affective appraisals of environments. In *Environmental aesthetics: Theory, research, and application* (pp. 120–132). New York: Cambridge University Press.
Russell, J. A. & Lanius, U. F. (1984). Adaptation level and the affective appraisal of environments. *Journal of Environmental Psychology*, **4**, 119–135.
Russell, J. A. & Snodgrass, J. (1987). Emotion and the environment. In *Handbook of Environmental Psychology*, (pp. 245–280). New York: John Wiley and Sons.
Stamps, A. E. (2000). *Psychology and the aesthetics of the built environment*. Boston, MA: Kluwer.
Stamps, A. E. (2005). Isovists, Enclosure, and permeability theory. *Environment and Planning B: Planning and Design*, **32**, 735–762.
van Veen, H. A. H. C., Bülthoff, H. H., & Givaty, G. (1999). Psychophysical experiments on the internet. In H. H. Bülthoff, M. Fahle, K. R. Gegenfurtner, & H. A. Mallot (Eds.), *Beiträge zur 2. Tübinger Wahrnehmungskonferenz*, p. 43, Kirchentellinsfurt, Germany. Knirsch.
Wiener, J. M. & Franz, G. (2005). Isovists as a means to predict spatial experience and behavior. *Lecture Notes in Computer Science,* **3343,** 42–57.
Yerkes, R. M. & Dodson, J. D. (1908). The Relation of Strength of Stimulus to Rapidity of Habit-Formation. *Journal of Comparative Neurology and Psychology*, **18**, 459–482.
Zapf, D. (1991). *Wohlbefinden: Theorie-Empirie-Diagnostik [Well-being: Theory, Empirics, Diagnostics]*, (pp. 227–244). Weinheim, Germany: Juventa.