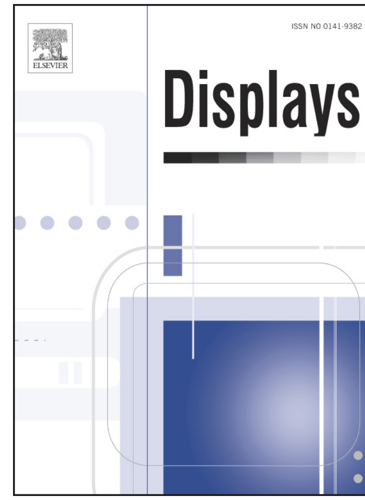
Accepted Manuscript



The influence of color grouping on users' visual search behavior and preferences

Rafał Michalski

PII: S0141-9382(14)00038-9
DOI: http://dx.doi.org/10.1016/j.displa.2014.05.007
Reference: DISPLA 1688

To appear in: *Displays*

Received Date: 11 July 2013
Revised Date: 5 May 2014
Accepted Date: 8 May 2014

Please cite this article as: R. Michalski, The influence of color grouping on users' visual search behavior and preferences, *Displays* (2014), doi: http://dx.doi.org/10.1016/j.displa.2014.05.007

The influence of color grouping on users' visual search behavior and preferences

RAFAŁ MICHALSKI†

†Institute of Organization and Management (I23),
Faculty of Computer Science and Management (W8)
Wrocław University of Technology,
27 Wybrzeże Wyspiańskiego,
50-370 Wrocław, POLAND
e-mail: Rafal.Michalski@pwr.wroc.pl

Corresponding author:

Dr. Rafał Michalski

Institute of Organization and Management (I23),
Faculty of Computer Science and Management (W8)
Wrocław University of Technology,
27 Wybrzeże Wyspiańskiego,
50-370 Wrocław, POLAND
e-mail: Rafal.Michalski@pwr.edu.pl
RafalMichalski.com
phone: +48 71 348 50 50
fax: +48 71 320 34 32

2014

Abstract

The article investigates how the various colors and color pairs used as grouping factors affect the visual search process and direct manipulation activities in the context of toolbar-like graphical panels. Red, green, and blue colors having the same perceptual distance in the CIELab space are used. The results demonstrate significant influence of the examined color-related factors on the speed and accuracy. The color preattentive property depends strongly on the grouping pattern: layouts with smaller colored areas were operated worse than panels divided into larger parts. Meaningful differences were also observed between panels with single and two colored backgrounds.

Preferences were examined by pairwise comparisons before and after performing the 'search and select' tasks. Subjective judgments were significantly differentiated by the toolbar background color pattern both prior to the performance tasks and after them. The initial relative weights structure changed decidedly after the performance experience being more consistent with the 'search and select' results. The location factor was irrelevant for the speed and accuracy as well as for preferences. Objective and subjective findings are compared and discussed. Linear regression models showing the preference structure change and the relationship between mean acquisition times and mean preference weights are developed and discussed.

Keywords

color grouping; display design; visual search; preattentive processing; Analytic Hierarchy Process preferences; direct manipulation

1. Introduction

Despite much technological progress in the contemporary interactive devices including various types of personal computers, laptops, palmtops, tablets or smart phones, a great portion of the communication between the human and a digital device is still based on a simple interactive styles invented many years ago. New ways of exchanging information with interactive systems are being used along with classical ones. Probably, the most widespread type of interaction currently employed in various IT solutions is the direct manipulation proposed by Shneiderman (1982, 1983) already in the beginning of 80s of the previous millennium. One of the crucial components of this approach is the so called 'point and click' method which requires from users to search for a desired graphical object, point it, and confirm the selection. Naturally, to apply this method in practice, various technical solutions have been developed including, for instance, light pens, digitizers, joysticks, arrow keys, track balls, touch screens and, of course, computer mice (Greenstein and Arnaut, 1988). Among some latest developments in this area there are devices used in the video game consoles, namely the Xbox 360 Kinect (Microsoft, 2012a), PlayStation 3®Move Motion Controller (Sony, 2012), and Wii U (Nintendo, 2012) or solutions used for manipulating virtual objects in a three dimensional space (Hand, 1997) including special gloves (e.g. CyberWorld, 2012), and recently developed gloveless sensors (Microsoft, 2012b).

Although through years there has been a significant change in technological aspects of various pointing devices, the idea of the search and select method is always similar: find the object, select it, and approve the choice. The target objects to be searched for within the graphical user interface are very often grouped and presented together forming the so called toolbars that usually take a rectangular shape. This research concerns those kinds of graphical structures with various colored features operated by means of the 'search and confirm' interaction style technique. Taking into account recommendations included in international standards (ISO 9126–1, 1998 and ISO 9241–11, 1998), two different perspectives were applied to investigate how a user interacts with toolbars having diverse color properties: the objective and the subjective one. The former included the efficiency and effectiveness evaluations whereas the latter regarded subjects' preferences toward stimuli.

The next two subsections provide the concise review of research related to the subject of the present study. The first one encompasses investigations of the 'search and select' method while the other presents studies concerned with a role of a color in visual search. In the last subsection, main goals of the current investigation are specified.

1.1. Studies of the point and click tasks

Studies related to the graphical characteristics of a target object in the human-computer interaction field fall generally into three categories (Michalski et al., 2006). The first one deals with the movement time needed to click a given object by means of a pointing device while the target is constantly visible to the subject. This was first studied by Fitts and Peterson (Fitts, 1954; Fitts and Peterson, 1964) and later applied by MacKenzie (1991, 1992) in the HCI field. Plamondon and Alimi (1997) together with Seow (2005) provide more information and comprehensive reviews of studies in this area. In the second group, participants search the interface for a graphical target without selecting and confirming their choice by clicking on the item found. Thus, papers in this trend focuse on cognitive aspects of the visual search performed in the HCI context. Probably the first attempt to do research in this field was made by Backs and colleagues (1987) who examined vertically and horizontally oriented menus. In the third field, approaches from the previous two were applied simultaneously, and the experimental tasks involved both the visual search and visually controlled motor activities aimed at selecting the object searched for. The studies in this group were possibly initiated by Deininger (1960), who investigated different configurations of numerical telephone keypads. The further developments in this area before 1999 were reviewed in detail by Kroemer (2001). A more detailed description of the presented trends is provided in the work of Michalski et al. (2006).

1.2. Research regarding the color in visual search

The vision process is extremely important in our day-to-day life since about 80% of information retrieved from the surrounding world comes through our eyeballs (Net Industries, 2012). Therefore, it is not surprising that the nature of a human being visual system has been extensively studied in various fields of science including biology, physiology, neurophysiology or psychology. The comprehensive review of the achievements made in this domain was prepared for instance by Findlay and Gilchrist (2003).

The search for different objects is one of the most important visual tasks people are performing uncountable number of times during their usual activities. The scientists have developed various models describing how the process of finding target items in the visual field is being carried out. Among the most popular approaches are the *spotlight* proposal provided by Posner and collegues (1980) and the *zoom lens* model introduced by Eriksen and St. James (1986). According to Posner et al. (1980) the visual activity during the objects search resembles a moving spotlight which covers only relatively small portion of the whole visual field at a given time. While the area is being processed, the next saccade is prepared to transfer the attention and the foveal vision to a different location. After jumping to a next spot, the whole process repeats. The competitive model of Eriksen and St. James (1986) assumes different behavior. In this case, the visual search is conducted by gradually narrowing down the analyzed region similarly to zooming the pictures in cameras and cam-coders. The theory was later supported experimentally by Zelinsky and others (1997).

The presented models were complemented by ideas put forward in the *Feature Integration Theory* (Treisman, 1980; Treisman and Gelade, 1982) and *Guided Search* (Wolfe et al., 1989; Wolfe, 1994; Wolfe, 2007). Both of the approaches integrate to some extent the *spotlight* and *zoom lens* hypotheses. They assume that at first the field of view is explored generally. The process has a parallel nature thanks to some graphical features which can be processed almost at the same time. After this phase, the objects identified in the first step are investigated sequentially. The first stage is often called preattentive because it is believed that the process does not involve attention. Much effort and research has been devoted to identify attributes that facilitate the search process and make it more efficient. The results undoubtedly show that among the most important factors successfully used in the preattentive phase is a color. The detailed and comprehensive summary of the findings concerned with other preattentive features were presented by Wolfe and Horowitz (2004). Recent studies confirmed the existence of the preattentive mechanisms on a neurophysiological level. Bichot and colleagues (2001, 2005) examined a macaque monkey which visual system is very similar to human and registered the activation of different neuron layers in the primary visual cortex. They identified two separate phases assumed by the theoretical models.

In the domain of Human–Computer Interaction (HCI) the visual search along with the role of preattentive features are of great importance. Although there were very few studies related explicitly to the role of various preattentive features, there were some papers dealing with the factors that are considered to facilitate the preattentive procedure. For instance, the screen design recommendations proposed within well known Gestalt Theory (Borchers et al., 1995, 1996; Change et al., 2002) such as proximity, similarity or closure are possibly connected with the objects' attributes used in the early vision processing. A brief description of the studies pertaining to the role of color as an important aspect influencing the visual search behavior in various kinds of display is presented in next subsections.

1.2.1. Color related research pertaining to graphical displays

The color was identified as an important factor in the quality of the graphical information presentation already in the beginning of the 20^{th} century (Le Courrier du Livre, 1912). The research on the effect of color on visual search efficiency in the context of the display design started as early as in the 50s of the previous century. Christ (1975) reviewed 42 studies on the color role in the visual display search tasks published between 1952 and 1973. The results led to ambiguous conclusions showing a significant positive influence of the color usage in some circumstances and negative effect on visual search performance in some others. The importance of the various aspects of color usage was examined later under diverse conditions. For example, the work of Murch (1985) who specified good and bad combinations of the text and background colors or Matthews et al. (1989) who reported decidedly longer search times for green on black displays as compared to the red on black. Among the researches employing more modern displays one can also find numerous papers. For example, Bodrogi (2003) investigated the effect of a chromaticity contrast displayed on the multi-color graphical user interface, Wu and Yuan (2003) studied the influence of highlighting and text color on visual search performance while Buchner and Baumgartner (2007) analyzed the impact of text-background polarity, ambient illumination and color contrast. Recently, Wu et al. (2009) examined 18 various graphical user interfaces containing many color attributes used for choosing the desired sofa appearance. They showed, among other things, that the color grouping is one of the most important factors influencing the subjects' performance and satisfaction. Similar results confirming the importance of color grouping were obtained in the research of Wu et al. (2010) who analyzed a user's behavior in the context of choosing color properties of a mobile phone. The preattentive color features were also integrated into the most famous models of visual search, namely ACT-R (Fleetwood and Byrne, 2006) and EPIC (Halverson and Hornof, 2012).

1.2.2. Color properties in web site usability

Apart from the studies related with color properties for general displays there were many investigations devoted specifically for the visual search of the web pages. For instance, Schaik and Ling (2001) examined the effect of differences in background contrasts on visual search performance. Later, the same authors analyzed the role of text and background color factors in web sites visual search (Ling and Schaik, 2002). Similar studies in this regard were performed by Hall and Hanna (2004) or in a very systematically designed research reported by Humar et al. (2008). The colors of hyperlinks in web pages were in turn examined by Schaik and Ling (2003) as well as Pearson and Schaik (2003).

1.2.3. Color issues in visual search of sets of icons

A number of studies have been conducted to determine the impact of various color related features on the visual search for target graphical icons from a group of other icons. Näsänen et al. 2001 investigated the visual search process in four types of square arrays of icons (3×3, 5×5, 7×7, 10×10) representing Latin letters and numerals under five different contrasts. They reported the increase in search efficiency for increasing values of contrasts for all layouts. In the research described by Ojanpää and Näsänen (2003), the 7×7 arrays of Latin letters and digits of three different sizes were examined under various luminance and colour contrasts. The general conclusion was that a moderate or even high colour contrast might not be sufficient enough to increase the speed of visual perception, if the luminance contrast between foreground stimulus and background is small. Näsänen and Ojanpää (2003) researched, in turn, the influence of image contrast and sharpness factors on visual search for computer icons arranged in a 13×13 array. The obtained results showed first the increased search performance with increasing contrast or sharpness, and then stabilizing at their medium levels. The graphical icons properties were also used in the study of Fleetwood and Byrne (2002). Their experiments involved three independent factors, namely the set size including 6, 12, 18 and 24 items, border type (none, circular, square), and finally three types of icon quality in which color and shape features were used to make the pop-out (preattentive) effect. The results showed significant effect of icon set size and their quality on the average response times. The interaction between those two variables was also statistically meaningful whereas the border factor and its interactions were irrelevant.

The study of Michalski and Grobelny (2008) was explicitly concerned and fully devoted to the target icon background color attribute used as an early vision property in visual search and click tasks of various types of toolbars. The authors examined the toolbar background color pairs (red-green, red-blue, and green-blue), panel color pattern (random and chessboard-like) and the toolbar orientation (horizontal versus vertical). Although, the outcomes generally confirm the positive influence of the color on visual search performance, they also showed that vertical configurations decidedly decreased this effect. As to the pattern factor, the early vision gave better results for randomly dispersed colors as compared to systematic, chessboard-like patterns.

1.3. Statement of the problem and purpose of the study

The present study experimental setup requires both the visual search as well as the visually controlled motor activities which are employed to select the target object and confirm it by a mouse click and thus, can be located in the third trend described in section 1.1. It has already been proved that the classical Fitts's law was inappropriate for user's interface tasks that required meaningful cognitive load (e.g. Grobelny et al., 2005; Michalski et al., 2006). Additionally, the results of some studies show that similar areas of the human brain are activated during body parts movements and the vision process (Wurtz et al., 1982; Kustov and Robinson, 1996; Colby and Goldberg, 1999). This suggests that tasks that require both activities at the same time could give different results than analyzing them separately. Such an assumption was experimentally observed in the research conducted by Hoffmann and Lim (1997). Though the graphical properties of toolbars or dialogue windows were already researched, and there were some studies showing that grouping is an important factor in visual search of the icon sets (Goldberg and Kotval, 1999; Niemelä and Saarinen, 2000; Murata and Furukawa, 2005), only the research of Michalski and Grobelny (2008) entirely dealt with the various toolbar color patterns as a grouping factor in the usability of the search and select type of graphical user interface interaction.

Given the literature review which clearly shows the importance of a color feature as one of the crucial preattentive factors, a small number of researches in this area seem to be somewhat surprising. Though the color is proved to be an important early feature, little is known how grouping patterns influence the search performance and preferences. The work of Michalski and Grobelny (2008) showed that the color feature used as an early vision property does not work equally well in all circumstances. In light of these findings it is justified to continue the research in this domain to point out those solutions that are especially useful in speeding up the visual search in graphical user interfaces. This study extends the knowledge in this area by the following issues:

- First of all, it examines which of various patterns of color grouping in toolbars are better in terms of visual search efficiency, effectiveness and satisfaction.
- Takes advantage of CIE Lab color space (Robertson, 1977) for the selection of background colors in experimental design to control the human perceptual differences of the examined chromaticities.
- The research results regarding the toolbar locations are not consistent as they probably strongly depend on the experimental context (Michalski, 2011), thus the toolbar location factor was also included as an independent variable to verify whether it significantly influences the visual search given various toolbar color properties.
- Because the outcomes of the Michalski and Grobelny (2008) showed that vertical configurations of toolbars decidedly diminished the color preattentive effect, in this study only the horizontal structures were investigated.
- In this research both the objective and subjective measures are involved to get the fuller picture of the usability of toolbar operations. In order to increase the reliability of the subjective results, the consistency of the users' preferences was controlled by means of the Analytic Hierarchy Process (Saaty, 1977; 1980) approach.
- Since users may change their preferences towards examined layouts after gaining some experience (Michalski, 2011), the preferences are examined prior and post the experimental tasks.
- What is the relation between objective efficiency and effectiveness measures and subjective preferences while performing simple 'search and select' tasks in toolbars where icons are grouped by colors.

In next passages of this work, the carried out experiment is described in detail, the obtained results are presented and analyzed and the discussion of the findings along with the conclusions is provided.

2. Method

2.1. Participants

All of the 65 examinees that took part in this research were recruited from among the students of the Computer Science and Management Faculty at the Wroclaw University of Technology. They all have reported having normal or corrected to normal visual acuity. The sample consisted of 42 females (65%) and 23 males (35%) with the age from 19 to 27 years. Most of the participants (69%) usually spend between 2 to 5 hours a day using personal computers for various purposes.

2.2. Apparatus

The experiments were conducted by means of custom made software developed in a MS Visual Basic™ 6.0 environment. All of obtained during the examination information was being saved in a relational MS Access™ database. The study took place in Wroclaw University of Technology teaching laboratories on identical desktop computers. The workstations were equipped with the same optical computer mice and 17" CRT type monitors with the resolution set at 1024 by 768 pixels. The classical Windows XP operating system color scheme was applied, and default computer mouse parameters were used.

2.3. Stimuli

The present research was focused on the color properties of typical toolbars used in graphical interfaces of modern software. The analyzed structures consisted of 36 square buttons of the same size often employed in current computer operating systems. The width and height of the single target equaled 330 twips (twentieth of an inch point) which for the used screen resolution corresponds to 22 pixels, 6 mm and 0°41' of visual angle. In order to minimize the mental effort of processing the meaning of the graphical target, 26 Latin alphabet characters and ten Arabic numbers were employed. They were displayed on toolbar items in 12 pts, bolded Times New Roman font type. All of the toolbars were arranged horizontally and comprised two rows and 18 columns. The examined panels were differentiated according to three factors: *Target background color* (TBC), *Toolbar color pattern* (TCP), and *Toolbar location on the screen* (TL).

Target background colors. This factor was represented in six levels. The colors employed corresponded to one of the color systems commonly used in technology, which is based on three hues: red, green, and blue. The mixture of the same basic three colors is used to produce any other color by cones in the human retina. The colors were applied to icons' background for panels in pairs: red-blue (RB), red-green (RG), and green-blue (GB) which results in three levels. Additionally, for comparison purposes, the single colorful toolbars were examined, which gave additional three levels: red (R), green (G), and blue (B). The colors were chosen in such a way that the differences in psychological perception between every two colors were similar. The lightness of the color in the CIE Lab color space for every color used in this study was set to the value of the grey color lightness used in the Windows classical color scheme. The distance between each color in the CIE Lab color space (Robertson, 1977) was set to 40 and the exact specification of the colors in various color systems is given in Table 1.

Table 1.

Toolbar color pattern. Three different patterns were used. The panels were divided into equal in size rectangular groups of neighboring buttons with the same background color. For a given panel, two out of three colors (excluding the grey one) were applied. Each section consisted of two rows and a specific number of columns, different for every factor level. The first toolbar included two blocks of nine columns in every color (02×09), the second type – six rectangles (three for each color) of three column clusters placed alternately (06×03), and finally the pattern of 18 groups (nine for every color) containing only one column (18×01), which were also arranged in a similar way.

Toolbar location on the screen. All of the investigated structures appeared either in the upper left (L) or upper right (R) corner of the screen. However, to compensate for faster acquisition of targets situated at screen borders (Farris et al., 2002, 2006; Jones et al., 2005), the toolbars were moved away from the vertical and horizontal screen edges. The used gap of 18 pixels is the same as a typical height of the top title bar in various graphical operating systems.

The examples of toolbars investigated in the present study are given in Figure 1 (a, b, c).

Figure 1.

2.4. Dependent measures

Two general types of dependent measures were applied in this research. The first set included pointing task performance measures while the second one was concerned with subjective assessment of the tested toolbars. The employed software registered times necessary for finding the target and clicking it with a mouse as well as the wrong selections being made by subjects. The perceived users' usability of the examined toolbars was evaluated by means of their pairwise comparisons. The subjects' priorities were obtained by applying the Analytic Hierarchy Process (AHP) framework developed by Thomas Saaty (1977, 1980) and widely used in many areas – the reviews of applications are provided by Zahedi (1986) and Ho (2007). The method is based on computing the principal eigenvector of the matrix containing comparisons and additionally allows for approximating the level of participant's consistency in comparing the objects.

The algorithm for finding the relative weights was implemented in the supporting software. At first the symmetric and reciprocal pairwise comparison matrix ($\boldsymbol{C}$) is constructed for every participant. This square matrix ($n \times n$, n – is the number of assessed variants) contains dominance values c_{ij} of variant i over j. The vector of relative preferences $\boldsymbol{v} = [v_i, ..., v_n]^T$ towards the examined alternatives is computed from the following relationship:

$$\boldsymbol{C} \cdot \boldsymbol{v} = \lambda_{max} \cdot \boldsymbol{v},$$

In this approach the weights $\boldsymbol{v}$ take values between zero and one while the sum of all weights for a given participant equals one. The scalar value of λ_{max} is the maximal eigenvalue of the matrix $\boldsymbol{C}$, and is used to calculate consistency ratio values (CR) according to the following relationships:

$$CR = CI / RI,$$

$$CI = (\lambda_{max} - n) / (n - 1),$$

where CI is a Consistency Index, and RI - the Random Index depending on the number of alternatives being compared (Forman, 1990).

Both priority vectors and consistency ratios (CR) calculated within this framework were treated as dependent measures and further analyzed. The obtained for every participant hierarchy vector is normalized which means that the vector items – the weights – sum up to one. Bigger preferences toward toolbars correspond to bigger values of relative weights. The CR parameter results should be interpreted as follows: the bigger the CR value, the higher inconsistency was detected during the pairwise comparisons for a given participant.

2.5. Experimental design

Generally, all combinations of the independent variables levels produce 24 experimental conditions to be examined: [(*three target background color pairs*) × (*three toolbar color patterns*) + *three toolbars with single target background color*] × (*two toolbar location on the screen*) = (3 × 3 + 3) × 2 = 24. A mixed model including within as well as between subjects design was used in this investigation. Such a solution was mostly dictated by the number of pairwise comparisons required to assess the participants' preferences towards examined layouts. The application of only the within subjects design would result in too excessive number of necessary comparisons – namely $(24^2 - 24)/2 = 276$. Therefore, the toolbar background color pairs were treated between subjects, whereas the remaining two variables – within subjects. Such a design required three experimental groups each of which examined five different toolbars on both left and right upper corner of the screen. This allowed keeping the number of comparisons at an acceptable level of 45 per individual. The detailed assignment of experimental conditions to the groups is demonstrated in Table 2.

Table 2.

The selection of the color as the between subject factor was also supported by the Saaty's (1980) recommendation to assess the relative likings of objects having similar magnitude of differences. In light of the results provided by Wolfe and Horowitz (2004), the target object background colors were assumed to differentiate the users' attitudes the strongest.

2.6. Procedure

Before the investigation started, the subjects were given a short review of the forthcoming experimental procedure and were informed about the aim of the research. Before each stage of the examination, the users were instructed by the examiner in detail about the activities to be performed. Appropriate texts explaining what to do were provided by the supporting software and, additionally, the subjects were encouraged to ask for help if anything was unclear. The questions about personal data and computer literacy appeared at first. Then, pairwise comparisons of toolbars were performed to obtain information about users initial attitudes towards the examined layouts. The subjects were presented with a pair of toolbars and asked to specify their preferences by clicking one of the radio buttons situated below the toolbar images. The scale included five items *None*, *Somewhat better*, *Better*, *Far better*, and *Definitely better* (see Figure 3). The dimensions of all of the toolbars were exactly the same as during performing the search and click tasks. After choosing the more preferred option, the user clicked the *Next* button. Once all of the comparisons were performed, the subject clicked the button *Finish* which appeared in the last comparison. At this moment the gathered data was saved to the database and the next phase started.

The next stage included the repetitive 'search and click' tasks for the selected toolbars. The target icon was presented in the middle of a screen – at this moment the examined panel was invisible. After clicking the *START* button, the instructive window disappeared, and a randomly selected toolbar with randomly distributed letters and numbers was shown (Figure 2).

Figure 2.

When the user clicked any of the toolbar buttons, the next icon to be searched for was shown in a dialogue window. The target once demonstrated did not appear again during the block of ten trials for a given panel of buttons. The user was informed about the average completion times for the correctly selected icons and the number of mistakes when the examination of a given experimental condition was finished. Before the proper testing took place, the participants performed at least ten attemptive trials. The results of this part were not registered. The users performed ten trials for every examined layout so altogether they were to search for 100 icons in all of the ten experimental conditions. During the 'search and click' phase, the examiner continuously controlled whether participants sat directly in front of the computer screen, with the eyes at the same height as the top edge of the monitor and if the distance to the screen was about 50 centimeters.

In the final part of the examination, the participants were asked again to express their opinions by means of pairwise comparisons on which of the tested toolbars were better operated. In the preference analysis both before and after the 'search and select' tasks, the panels were presented graphically (Figure 3) at random by the utilized software.

Figure 3.

3. Results

The obtained in this study results are presented in this section. They have been generally divided into two groups. The first deals with the objective results pertaining to the 'search and click' tasks performed during the experiments whereas the other one demonstrates the outcomes of the users' preferences towards examined toolbars.

3.1. Search and click task results

In the next subsections, the 'search and click' tasks are first characterized by the basic descriptive statistics, and then the findings of the analysis of variance are described. The users' wrong selections during the trials were examined in the third subsection.

3.1.1. Descriptive statistics

In the conducted experiment, each of the 65 participants tested 10 different graphical structures by searching for 10 various target objects thus, altogether, the subjects performed 6500 individual search and select by a mouse click tasks for various examined toolbars. Among them, there were 151 incorrect selections. The calculated and further presented descriptive statistics and analyses of variance relate only to the proper 6349 selections. The basic descriptive statistics are demonstrated in Table 3. The central tendency measures are decidedly different from each other with the median smaller than the arithmetic average by almost 26%. Moreover, the big value of skewness (3.6) and the large kurtosis (21) indicate that the empirical probability density distribution for the acquisition times in this research is far from being Gaussian. The computed parameters suggest the distribution significantly left skewed and markedly less dispersed than the normal one.

Table 3.

Similar statistics were computed individually for all of the examined toolbars and were put together in Table 4. Although there are, naturally, differences, the parameters are in general qualitatively comparable to the overall descriptive statistics. Medians of the selection times for each panel are distinctly lower than their means, all of the skewness parameters are positive and the kurtoses - though they vary significantly – are much bigger than in the Gaussian distribution.

Table 4.

The toolbar divided into two red and blue areas which appeared on the left upper corner of the screen (L_RB_02×09) was operated the fastest by examinees with the average acquisition time of 1560 ms. The longest mean selection times were observed for the green single color panel located in the right upper corner (R_G) and amounted to 2319 ms. The relative difference between the best and the worst mean result equaled as much as 49%.

3.1.2. Analyses of variance

As it was shown in previous section of this article, the values of skewness, kurtosis as well as the significant discrepancy between the central tendency measures indicate that the experimental distribution of the dependent measure does not come from the normal distribution. Given the results reported by Michalski (2005) for similar experimental tasks, it has been assumed that the dependent variate has the inverse Gaussian (IG) distribution and the ANOVA available within the framework of Generalized Linear Models (GZLM; Nelder and Wedderburn, 1972) was applied to verify the significance of differences in means. The outcomes of the three way GZLM ANOVA presented in Table 5 show that the factors of *Toolbar background color pair*, and *Toolbar background color pattern* significantly affected the task execution times. The *Layout Location* on the screen effect along with the all of the interactions occurred to be statistically irrelevant.

Table 5.

The influence of the *Toolbar Background Color Pair* factor on the efficiency of toolbar operation is illustrated in Figure 4. The red-blue layouts happened to be searched for the fastest (1834ms) whereas the longest mean acquisition times (1959ms) were obtained for red-green graphical panels. The relative difference equalled 7%.

Figure 4.

The experimental design included apart from the background color pairs also toolbars with three uniform background colors, namely red, green, and blue. These experimental conditions however, could not be included in the initial three factorial GZLM ANOVA because none of the background color patterns could be applied. Obviously, the comparison between two-colored layouts and graphical structures with a single background color applied is very interesting and such an analysis was conducted. For this purpose the two way GZLM ANOVA including the *Toolbar location effect* and the *Toolbar background colors* factors was employed. The latter variable was studied on six levels which included all the three toolbar background color pairs together with the three toolbar types with one uniform color. The summary of the analysis is given in Table 6. The results show the insignificance of the *Toolbar location effect* and the interaction and the significance ($W = 20$, $p < 0.000001$) of the *Toobar background colors* variate.

Table 6.

The average selection times depending on the background color of examined layouts are demonstrated in Figure 5. One can visually assess looking at these results that the two-colored toolbars were generally better operated than single color panels. The relative difference between the shortest mean time (1834ms) and the longest one (2279ms) amounted to 24%.

Figure 5.

The formal verification of those differences was carried out by conducting a series of the post-hoc GLZM one way ANOVAs. The obtained results are put together in Table 7.

Table 7.

From the presented data regarding the factor levels discrepancies one can clearly notice that all the toolbars with a uniform background color were meaningfully slower operated than any of the toolbars with two different background colors. There was only one significant difference at the level of 0.1 observed among the panels with uniform background color: the times for green toolbars were considerably longer than those obtained for blue graphical panels. Among the two-colored panels the only irrelevant difference was between the red-green and green-blue toolbar types.

Figure 6 demonstrates mean selection times for the effect of *Toolbar background color pattern*. These results indicate that the layouts containing a smaller number of colored areas were operated decidedly faster than structures with multiple colored parts consisting of smaller number of targets. The shortest times (1709ms) were observed for the layouts divided only into two sections, while the longest ones (2176ms) for toolbars with as many as 18 rectangular groups. The relative difference between the worst and the best results amounted to 27%.

Figure 6.

The further post-hoc analysis of differences between the individual levels of the *Toolbar background color* variable (Table 8) showed additionally that the discrepancies were significant for all of the variable level pairs.

Table 8.

3.1.3. Wrong selections

The subjects mistakenly clicked 151 other than required target icons, which is 2.3% of all performed 'search and click' tasks. The applied nonparametric Chi-square tests showed significant differences for all of the analyzed factors (Table 9).

Table 9.

The toolbars that appeared in the right upper corner of the screen were operated more effective ($\chi^2 = 7.4$, $p = 0.0066$) with error rate equaled 1.8% than those on the left hand side (2.8%). The percentages of users' mistakes dependent on the *Toolbar background colors* factor ($\chi^2 = 16$, $p = 0.0056$) are given in Figure 7. The results suggest that participants made more errors while testing the toolbars with a single background color as compared to the two-colored ones. The biggest number of wrong selections was registered for the uniform red toolbars (3.8%), whereas the smallest number of mistakes was observed for the green-blue layouts (1.5%) which means that users in the former case committed almost two and a half more mistakes than in the latter group.

Figure 7.

In order to check the differences between individual factor levels, a series of additional χ^2 tests were carried out, and the results are put together in Table 10. They show none significant differences among the toolbars with a single background color. Similarly, no meaningful differences were observed among the group of two-colored toolbars. All of the statistically relevant discrepancies were noticed between the uniform colored toolbars and those with two colored background with only three exceptions, namely the green with the red-green and red-blue, and the blue one with the red-blue.

Table 10.

The effect of *Toolbar background color pattern* on error rates is illustrated in Figure 8. The graph demonstrates apparent discrepancy between the error rates registered for the patterns with two and six background color parts and the layouts containing none or 18 colored areas. The smallest error rate (1.2%) was observed for the 02×09 graphical structure while the biggest one (3.2%) for toolbars with no patterns.

Figure 8.

The results concerned with differences between the factor levels were provided in Table 11. They confirm the initial results observed in Figure 8. A series of χ^2 test showed no significant differences only between toolbars with patterns 02×09 and 06×03 as well as between 18×01 and no pattern. All the rest individual differences were statistically meaningful.

Table 11.

3.2. Relative likings results

This section contains the subjective results obtained within the framework of Analytic Hierarchy Process. At first, the consistency analysis is provided and next the outcomes of pairwise comparisons of the studied in this research layouts are demonstrated and thoroughly analyzed.

3.2.1. Comparisons' consistencies

In this investigations the subjects' consistency ratios varied considerably ranging from 0 up to even 0.436 with the average value of 0.0985 and standard deviation equaled 0.077. Some of the additional descriptive statistics for the examined experimental groups along with the number of subjects meeting the criteria of CR being lower than 0.1 are given in Table 12.

Table 12

A standard analysis of variance was applied to verify if the consistency ratios depended on the assignment to the experimental group and whether the *CR* values were different for pairwise comparisons before and after the 'search and select' tasks. The summary of this analysis is provided in Table 13.

Table 13.

The results show that the CR mean values did not differ considerably between the examined groups ($F = 0.64$, $p = 0.53$). The moment of comparison had a statistically meaningful impact on the average consistency ratios with greater mean consistency after the efficiency testing (0.086) as compared to the 0.11 obtained for the comparisons executed prior to performing the 'search and select' trials. However, the difference was important only at the 0.1 level ($F = 2.96$, $p = 0.088$). The interaction between those two factors occurred to be statistically irrelevant ($F = 0.066$, $p = 0.94$).

The further analyses of the users' preferences were conducted only for the subjects that provided coherent comparisons that is, those with CR values lower than 0.1. To check if the number of excluded subjects were dependent on the experimental group assignment and the moment of performing the preference analysis, the χ^2 test was employed. The results are put together in Table 14.

Table 14.

The demonstrated data showed no statistically significant influence of the subjects' assignments to experimental groups ($\chi^2 = 2.5$, $p = 0.29$) and the moment of making the comparisons ($\chi^2 = 0.54$, $p = 0.46$) on the number of excluded subjects.

3.2.2. Relative weights analysis

This part of the article presents the results analysis of the participants' relative likings of the tested graphical toolbars. They were computed according to the AHP methodology from pairwise comparisons performed twice: prior to the search and select tasks and directly after executing the trials. The outcomes do not include users for whom the calculated consistency ratios were lower than the 0.1 value recommended by Saaty (1980). The average relative weights together with the mean standard errors obtained for all of the experimental conditions are presented in Table 15.

Table 15.

The data from the table 15 were transformed to the hierarchy of preferences presented as ranks to facilitate the analysis. The outcome is put together in Table 16.

Table 16.

The analysis of the data gathered in Table 15 and 16 can lead to some general remarks. In almost all cases the best perceived were the patterns dividing the toolbars into six colored parts measured both before and after the search and click tasks. Only in three out of 12 cases the first rank was given to the pattern containing two uniform colored areas. The last places in the preference structure were occupied at all times by the single colored layouts. The graphical panels with the 18x01 pattern were usually in the middle of the preference hierarchy.

The six, two way analyses of variance were employed to verify whether the effects of the *Toolbar location on the screen* and the *Toolbar background pattern* meaningfully influenced the obtained relative likings prior to the toolbar efficiency testing and after it. The results are presented in Tables 17-19.

Table 17.

Table 18.

Table 19.

All of the ANOVAs applied for the relative weights provided similar results. The *Toolbar background color pattern* significantly differentiated the mean weights whereas the *Toolbar location on the screen* factor along with the interaction were statistically irrelevant both before and after the search and click trials. The statistically meaningful factor was graphically demonstrated in Figures 9-11, separately for given color pairs.

Figure 9.

The data from Figure 9 show the superiority of two-colored toolbars over single-colored ones. Moreover, one may notice that after the toolbar efficiency examination there was an increase in the preference weights for toolbars consisting of bigger areas of colored icons (RG_02×09 and RG_06×03) and the decrease in single-colored configurations together with the two-colored toolbar with the biggest number of small, uniformly colored parts (RG_18×01).

Figure 10.

The outcomes computed for the red-blue layouts (Figure 10) were generally similar to the results for the red-green toolbars with the higher weights attributed to the two-colored panels. Also in this case, after the search and click tasks, the rates for the single color toolbars along with the RG_18×01 panel were lower while the RB_02×09 and RB_06×03 configurations were better perceived than initially.

Figure 11.

The mean weights presented for the green-blue toolbars in Figure 11 are somewhat different that the results described above for the previous toolbar background colors. Generally, the toolbars with two target background colors were perceived better than the single colored layouts, before as well as after the test trials. After the search and click tasks, there was a large increase of the weight for the GB_02×09 pattern but the preference for the GB_06×03 toolbar decreased which was not observed for the red-green and red-blue panels of icons. However, the preference hierarchy observed after the toolbar efficiency examination resembles those obtained in previous two cases.

4. Discussion

The current study main goal was to investigate toolbar color properties in visual search for target items along with the pointing and confirming by a mouse click the identified icon. The obtained results showed a significant impact of the toolbar color pattern and applied color pairs on the efficiency, effectiveness as well as participants' preferences. The detailed discussion of the analyzed factors is given in the following subsections.

4.1. Objective measures

In the current study two types of objective dependent measures were applied, namely the target acquisition times and error rate. The 'search and select' times can be treated as an efficiency measure of the usability of this type of interaction whereas the number of mistaken selections could be used for assessing the effectiveness of the examined interface.

4.1.1. Search and select times

The findings related to the main effects investigated in this research clearly show that application of a colorful background as a grouping factor for target objects within a given set of icons decidedly improves the visual search performance as compared to the single colored toolbars. The obtained acquisition times were compared with the single colored grey toolbars used in a study of Michalski et al. (2006) by means of a series of one way GLZM ANOVAs. The results of this analysis (Table 20) demonstrate significantly shorter search times for the toolbars with two background colors than for their grey counterparts. No meaningful differences were observed between those grey toolbars and the red, green, and blue single-colored layouts applied in this study.

Table 20.

This result supports the findings obtained in literature on general visual search performance and the outcomes of the Michalski and Grobelny (2008) research regarding the horizontal panels with randomly dispersed button background colors.

The detailed analysis of the applied in this experiment patterns indicate that the color preattentive properties of the examined toolbars gradually decrease while the number of colored groups of targets increases. For configurations containing 18 separate groups of two items having the same background color, the 'search and click' performance is comparable with the single colored layouts. It seems that the users change their visual search strategies from the zoom-lens model (Eriksen and St. James, 1986) which assumes the preattentive parallel search in the first phase to more *spotlight*-like (Posner et al., 1980), serial type of visual behavior. This fact seems to be in sharp contradiction with the results concerned with the strong color preattentive properties obtained in studies on general vision. On the other hand, it confirms the initial conclusion on considerable limitations of the color early vision role in the usability of toolbar operation drawn from the Michalski and Grobelny (2008) results. It is also in concordance with previous studies in the human factor field showing that the color usage may either work in favor or not for the users' performance (Christ, 1975).

Although the perceptual color differences were controlled by selecting the toolbar background colors equally distant from each other in the CIE Lab color space, the results indicate significant differences among the used color pairs. The selection times were the shortest for red-blue combination (1834ms), next there were green-blue toolbars, and the longest mean times were observed for the red-green pair. Similar results were reported by Michalski and Grobelny (2008). For the toolbars with randomly distributed background colors the best, in terms of the shortest mean acquisition times, were red-blue panels (1908ms). The mean 'search and select' results for red-green and green-blue combinations were reversed in comparison with current investigation findings, however the medians for those two types of layouts were equal in the Michalski and Grobelny (2008) research. The difference could have been caused by not controlling the color perceptual differences in their study. Possibly, the obtained differences in acquisition times concerned with the use of different colors might be explained on the grounds of the findings obtained in research dealing with the impact of various colors on the human being physiology. The early research showed that the color hue influences some of the psychophysical parameters such as blood pressure, frequency of eye blinks, respiratory and heart rate (Gerard, 1958), skin conductance (Wilson, 1966), electrical brain responses (Gerard, 1958; Clynes and Kohn, 1968). The results generally show that the red stimuli are decidedly more arousing than blue or green colors. Clynes (1977), additionally, argues that the red color significantly stronger activates the human brain than other colors and this appears irrespective of other conditions. These suggestions seem to be supported by this study results. They explain shorter results obtained for toolbars containing red background colors both in single and two colored panels even if the perceptual color differences are similar.

In this study, the distance between the START button and the target icon were not controlled in order simulate more real conditions of searching and selecting icons. It is well known that the time of pointing the target, according to Fitts's law (Fitts, 1954; Fitts and Peterson, 1964), depend strongly on the amplitude of movement, however in our study the task accomplishment times depended also on the cognitive effort necessary to find the appropriate icon. In the current experiment one may assess the approximate time necessary solely for the mouse movement to check its share in the whole experimental task. The Fitts's *ID* (Index of Difficulty) proposed for mouse movement activities by Welford (1960) takes the following form $ID = \log_2(A / W + 0.5)$, where A denotes the movement amplitude while W is the width of a target. If we take the Fitts's equation parameters provided by MacKenzie (1992) for computer tasks ($MT = 12.8 + 94.7 \cdot ID$) and calculate the ID for the worst ($ID = 5.01$) and the best ($ID = 3.78$) scenario in terms of the distance covered by the mouse, the predicted movement time would amount to 487ms and 371ms respectively. Thus, the difference was equaled merely 116ms. In relation to mean acquisition times obtained experimentally in this research (2042ms), this seems to be of minor importance especially if we take into account that the visual search and the mouse movement activity may overlap.

As for the toolbar location variable, the efficiency results occurred to be statistically irrelevant. None of the left top or right top locations of the examined layouts was better operated. This finding is consistent with the results provided by Michalski et al. (2006) for similar horizontal configurations with grey toolbar background colors and with the outcomes presented in the article of Michalski (2011) for vertical layouts. Also none of the interaction of the location variable with others studied in the present investigation was statistically important. It can be concluded that the visual processing of the examined colored toolbars along with the necessary mouse movement does not depend on the toolbar location on the screen. In some other studies where the target location on the screen was examined the results were inconsistent. Campbell and Maglio (1999) as well as Schaik and Ling (2001) showed the significant impact of the target location while Pearson and Schaik (2003) and Kalbach and Bosenick (2003) did not observed any meaningful impact. In these studies, however, the experimental setups differed considerably from the present study as they examined the search for hyperlinks in web pages.

4.1.2. Committed errors

The overall number of mistakenly selected targets was at a level of 2.3% which did not decidedly diverge from the previous similar 'search and select' studies' results. Grobelny et al. (2005) report less than 2% for the small- and middle-size icon sets, for big icons there were no errors at all. In the work of Michalski et al., (2006) the mean wrong selections for all experimental conditions amounted to 1.7% of all trials, in the research Michalski and Grobelny (2008) only 1.25% whereas the error rate in Michalski (2011) experiments equaled 1.5%. In studies that had significantly different experimental procedures the incorrect selections percentages were very diverse. In the web page visual link identification research of Schaik and Ling (2001) the mean percentage of targets incorrectly identified amounted to 1.6%. In a subsequent investigation by these authors (Ling and Schaik, 2002) the mean error rate for all experimental conditions was much bigger (10.6%). However, in this case the experimental task was more difficult and involved more information processing than in their previous research. The error rate for the Schaik and Ling (2003) study including information retrieval and pointing tasks varied from 1% to as much as 29% depending on the minimum number of pages to visit in order to find answer. In the experiments of Pearson and Schaik (2003) which included the 'search and select' procedure, the overall wrong selections rate amounted to 12%.

In the current study, all of the examined factors, namely *Toolbar panel location*, *Toolbar background colors* and *Toolbar background color pattern* had a significant impact on error rates (Table 9). Toolbars presented in the right upper corner of the screen were operated with significantly better accuracy than those appearing in the left upper corner. This outcome is in contradiction with the results provided in Michalski et al. (2006) and Michalski (2011) regarding similar toolbars where the location effect on the error percentages was irrelevant. The differences possibly result from that the color related factors were not included in these experiments. Thus, it seems that the location impact on the effectiveness might possibly be connected with different processing of the set of icons with colored backgrounds. Similarly, Pearson and Schaik (2003) showed no significant impact of the target position on the search effectiveness both for the visual and interactive search tasks. The presented results are also different from those reported in the work of Schaik and Ling (2001). They demonstrated significantly better accuracy if the searched item was on the left hand side of the screen than on the right one, which is opposite to the outcome of this study. In this case however, the users searched for word targets in a menu of a mock web page instead of simple icons and responded by pressing the appropriate key on the keyboard.

The significant impact of *Toolbar background colors* on the error rates obtained in this study conflict the results obtained in the work of Michalski and Grobelny (2008). Although the experimental procedures were similar, they did not register any meaningful influence of this factor. This could probably be explained by the fact that they did not control the differences in the color hues. Michalski and Grobelny (2008) did not notice any impact of *Toolbar background color pattern* on the number of mistakes, but their pattern conditions differed considerably from this research – they analyzed chessboard-like and randomly arranged patterns.

There were also some research where the number of mistakes depended on the colors employed. Ling and Schaik (2002) examined six combinations of text and background colors in navigation frames, namely black text on white background, blue on white (default), blue on yellow, yellow on blue, red on green and green on red. The color combination effect was significant and further pairwise comparisons showed better accuracy for the green-red, blue-yellow and yellow-blue conditions as compared to the blue on white condition. However, one should bear in mind that in this investigation they did not included the visually controlled goal movement, and the background colors were applied to all experimental conditions without producing the grouping effect.

The results for the interactive search tasks given in the article of Pearson and Schaik (2003) seem to be partly consistent with the current research results. They reported meaningfully better accuracy for the blue links at the right of the screen than for the red ones. However, for the visual search tasks performed in their study the link color effect (blue, red) was insignificant. Schaik and Ling (2003) analyzing the 'search and click' tasks for blue and black hyperlinks in the navigation menu of an intranet web page, also reported the meaningful impact of hyperlink colors (black and blue) on the accuracy of task performance. Blue links produced more correct answers than black ones. Although, the studies of Pearson and Schaik (2003) and Schaik and Ling (2003) are generally in concordance with this research results, one should remember that they manipulated the text hyperlink color, while here the toolbar icons with various background colors were used.

Additionally, when the *Toolbar background colors* factor from the current study is concerned, one may notice that generally toolbars with a single background color were operated generally slower than toolbars with two background colors having at the same time generally bigger error rates (compare Figures 7 and 5). However, is seems reasonable to analyze those two toolbar groups separately. If we take into account the performance and error rate only for the two-colored layouts, the best operated red-blue toolbars had the biggest mean error rate. And similarly, comparing toolbars with only single background colors, the fastest operated toolbars with red backgrounds had the biggest error rate. Apparently, subjects were more aroused by the properties of the colors and the influence of the color on their physiology resulted in executing tasks faster but unfortunately with lower accuracy. This supports the observed speed-accuracy trade-off phenomenon reported in earlier goal directed rapid movement studies (Plamondon and Alimi, 1997; Dean et al., 2007).

Surprisingly, the speed-accuracy trade-off was absent in the *Toolbar background color pattern* effect. In this case, the shorter acquisition times corresponded to the lower number of committed mistakes (compare Figures 8 and 6). Thus, the higher the preattentive effect was and as a consequence bigger efficiency, the effectiveness was also better. It seems that the grouping effect related to background colors was so strong that the speed-accuracy trade-off occurred to be irrelevant.

4.2. Subjective measures

4.2.1. Consistency ratios

The application of the AHP technique allowed for assessing the reliability of the obtained pairwise comparison results by calculating the consistency ratio values and excluding results which are not acceptable. The percentage of exclusions in this study amounted to 35% which is decidedly smaller to the value of 59% presented by Michalski (2011), who applied the same procedure for assessing comparisons' consistencies of grey vertical toolbars. Since the consistency ratio might also be related to the subjects' difficulty of making subjective judgments due to the similarity of the compared variants, it can be supposed that expressing the preferences towards colored toolbars is generally easier than towards grey toolbars. The exclusion rate depended neither on the experimental group nor on the moment of performing comparisons which is consistent with the results from the Michalski (2011) study.

The analysis of the *CR* values revealed no differences between the experimental groups, which means that the level of judgment concordance did not depended on the applied color pair. The consistency depended, however, on the moment of making the comparisons with lower mean *CR* values after the search and click tasks than prior to them ($p = 0.088$). Such a result demonstrates that the experience gained during the experimental task helps the subjects to clarify their structure of preference hierarchy and thus improve their consistency. Such an effect was not observed in the work of Michalski (2011) which could be caused by a bigger similarity of the presented experimental conditions in their work as compared to this study.

4.2.2. Preference weights

The mean relative likings computed for all experimental groups were strongly differentiated by the *Toolbar background color pattern* factor both prior to and post search and click tasks whereas the *Toolbar location on the screen* and the interaction between those two variables were insignificant in all cases.

Although the impact of the *Toolbar background color pattern* panel layout on the users' preferences was meaningful both before and after accomplishing the tasks, the preference structure changed decidedly. Michalski (2011) demonstrated that people are able to change their preferences quite fast after gaining some experience in operating grey vertical toolbars. The present study observes the same effect for the toolbars with various colored background patterns. To see how the preference hierarchy has changed, the formal models were made separately for the results obtained prior and after the performance tasks. The general regression model based on the dummy variables made for the relative weights obtained prior to the 'search and click' tasks takes the following form:

$$\textit{Mean preference weights prior} = 0.0746 + 0.0654 \bullet D_{06\times03} + 0.0309 \bullet D_{02\times09\_or\_18\times01},$$

where the binary $D_{06\times03}$ variable takes the value one for the 06×03 pattern type of toolbar and zero elsewhere, while the $D_{02\times09\_or\_18\times01}$ is one for 02×09 and for 18×01 pattern toolbars and zero for the remaining cases. The coefficient of determination amounts to $R^2 = 92.5\%$ (adjusted 91%), the model is significant ($F = 74$, $p < 0.000001$) and all of the parameters are significantly different from zero ($\alpha = 0.0001$). The observed and predicted mean relative weights for all experimental groups obtained before performance tasks are presented in Figure 12.

Figure 12.

Analyzing the model presented in Figure 12 one may notice a clear general pattern of the users' preference structure in all of the experimental groups. Single colored toolbars are the least preferred and the level of relative weights for them is consistent at similar level. Decidedly the most preferred toolbar type in all cases is the two colored one with the 06×03 pattern. The remaining two patterns (02×09 and 18×01) are comparable and significantly better in participants' opinions than the single colored toolbars but also distinctively worse than the 06×03 pattern panels.

The model for preferences expressed after the experimental task performance has this formula:

$$\textbf{\textit{Mean preference weights post}} = = 0.0721 + 0.0697 \bullet \boldsymbol{D_{02\times09\_or\_06\times03}}$$

where the binary $\boldsymbol{D_{02\times09\_or\_06\times03}}$ variable takes the value one for 02×09 and for 18×01 pattern toolbars and zero for the other panel types. The determination coefficient is slightly lower than in the previous model: $R^2 = 90.5\%$ (adjusted 90%). The model is also significant ($F = 124$, $p < 0.000001$) and both of the parameters are significantly different from zero ($\alpha < 0.000001$). The observed and predicted values for all of the experimental groups are illustrated in Figure 13.

Figure 13.

Comparing the second model with the first one, one can clearly see that the search and click experience caused a major change in the users' preferences. The single colored toolbars were rated only slightly worse as compared with the previous results, but the two-colored panels with the 18×01 pattern obtained substantially lower weights, comparable with those assigned to the single colored panels. Another noteworthy change took place for the 02×09 pattern layouts. Those changes apparently were brought about by the significantly shorter selection times registered for the 02×09 than for 18×01 pattern toolbars. Thus, those changes in preferences concerned initially underestimated (02×09) and overestimated (18×01) types of layouts and demonstrated that people altered their first impression according to the experience they gained. Although this general tendency was observed, from Figure 13 it can be noticed that only results for the green-blue panels (Group 3) fully reflects the efficiency based order of examined toolbars (compare with Figure 6). In the case of red-green layouts the preferences towards the 02×09 and 06×03 patterns were at almost exactly the same level while the 02×09 pattern toolbars were significantly better in terms of the average selection times. For the red-blue panels, in turn, the ratings of those two types of layouts were totally opposite in comparison with their efficiency results. It seems that in these two groups people were so strongly convinced that the 06×03 pattern panels are the best that even significantly worse objective results could not change their initial judgments. This somewhat surprising finding requires further exploration.

Differences between single color toolbar preferences were statistically irrelevant for all of the three experimental groups both before and after the search and click tasks. The Fisher LSD post hoc tests for groups gave the following results: $G1_{red\text{-}green}$: $p_{before} = 0.73$, $p_{after} = 0.98$; $G2_{red\text{-}blue}$: $p_{before} = 0.79$, $p_{after} = 0.92$; $G3_{green\text{-}blue}$: $p_{before} = 0.93$, $p_{after} = 0.40$. This outcome is in contrast to the results from the general psychology where preferences towards red, green and blue colors differ significantly. For instance Granger (1955) and Guilford and Smith (1959) obtained the following hierarchy of preferences blue ≻ green ≻ purple ≻ red ≻ yellow. Helson and Lansford (1970) reported similar order: blue ≻ green ≻ red ≻ purple ≻ yellow. This inconsistency may be attributed to the observation that people could have different preferences about colors depending on the object to which the color is applied (Taft, 1997; Schloss et al., 2012). Thus, the color preference order obtained in the previous studies without a specific context might not apply to the toolbar background color preferences investigated in the current research.

The lack of influence of the *Toolbar location on the screen* on the subjects' likings in the present research conflicts the result obtained by Michalski (2011), where this effect was significant before the performance examination took place. This contradiction could possibly be explained by the inclusion of color in the present study, which probably much stronger influences the user preferences than target position factor. On the other hand this result might be somewhat surprising if we become aware that the human brain hemispheres are to some degree specialized: the left one is more 'logical' while the right one is more 'emotional'. Having regard that the right visual field of view is projected to the left hemisphere and the left visual field of view to the right hemisphere, and that the users have some preferences towards colors, the toolbar location could have had an impact on the performance and preferences.

In order to obtain clearer view on the relation between the objective average acquisition times and the preferences, simple linear regressions were employed. The idea of computing these regressions is aimed at verifying whether the impact of the examined factors on the objective measure was the same as for the subjective one. If the preferences were shaped in the same way as the objective acquisition times, the regression would be close to linear. In other words, the bigger R squared value, the more the subjective preferences follow these objective ones. Prior to the 'search and click' tasks the regression formula takes the form of:

Mean preference weights prior* = 0.254 – 0.0000753 • *Mean selection times

with the R^2 = 51% (adjusted 47%), the model is significant (F = 13, p = 0.0029) and both parameters significantly different from zero (α = 0.005). After the performance examination, the regression is as follows:

Mean preference weights post* = 0.378 – 0.000136 • *Mean selection times

with R^2 = 82% (adjusted 81%); F = 59, p = 0.0000034 and all parameters significantly different from zero (α = 0.005). The graphical illustration of this approach is reported in Figures 14 and 15.

Figure 14.

Figure 15.

Both models present negative slopes which means that users ranked higher the faster operated toolbars, however after experimental tasks the slope is steeper which indicates stronger relationship. Moreover, examinees' preferences were generally decidedly closer to the objective results after the experience they gained as the R square raised from 51% before to as much as 82% after the 'search and click' performance. The preferences after the experimental trials were to a higher extent influenced by the objectively better variants. These analyses also show that users' preferences before the experimental tasks do not reflect the users' real performance during the search and click tasks. It seems that users are not able to correctly foresee how they would perform in this type of experimental task and their judgments are probably influenced by other factors. This finding is generally in concordance with the metaanalysis results of the relation between preferences and performance carried out by Nielsen and Levy (1994). They concluded that though there exists a positive correlation, one should not neglect the users' subjective opinions since they might be quite often different from the objective measures. This latter case is especially visible in the current study before the performance evaluation.

4.3. Theoretical and practical implications

The current research has several implications that extend our knowledge about human visual behavior and offer useful and specific guidelines for designing various graphical interfaces. The presented findings allow for understanding how preattentive features work together with goal directed movements. The outcomes also contribute to learning how the human visual system is coping with various types of color grouping. It was commonly assumed that the color is a strong preattentive factor, however as far as the author is aware nobody has shown that the search performance may differ considerably given various color patterns. Moreover, the experimental results contribute to the existing body of knowledge by providing some more insights into fundamental visual search involving background colors. These outcomes might be useful for developing or improving existing models of human visual systems or comprehensive cognitive models such as ACT-R, EPIC or EMMA (e.g. Anderson et al, 1997; Kieras and Meyer, 1997; Salvucci 2001; Byrne, 2001; Fleetwood & Byrne, 2006). They may also constitute a starting point for preparing further research in other specific environments.

Generally, the grouping, one of the Gestalt psychology principles (Borchers, 1996; Change et al. 2002), has already been applied for facilitating and organizing the human perception and is broadly used in designing various types of interfaces (e.g. Goldberg and Kotval 1999). The provided in the current research results revealing details about how various types of grouping influence the human performance and preferences are of significant interests for practitioners from various fields dealing with almost for any type of interactive systems based on visual interfaces.

More specifically, the present study outcomes demonstrated that the use of color for grouping purposes markedly speeds up tasks performance. This should incline the designers to use different background colors for enhancing usability of interactions involving graphical structures similar to toolbars. The red-blue pair seems to be working the best, but all two-color layouts are much better operated than single color backgrounds. While choosing the single background arrangements the colors used did not affected the efficiency, therefore, the decision which one to apply should result from other premises such as users' preferences. A next general recommendation for practitioners is concerned with a way the grouping should be applied. In light of the presented findings, grouping works worse if there are smaller groups of uniform areas. This should also be taken into account while making design decisions all the more that bigger uniform areas are better not only in terms of operation speed but also in accuracy. In cases where the error rate is of the greatest importance one should not forget about the speed accuracy trade-off effect. When the focus is both on speed and accuracy, it might be reasonable to improve the efficiency by applying colors and decrease the possible error rate by other means such as increasing the size of the target object or apply appropriate color based grouping which does not exhibit the speed accuracy trade-off effect.

Other practical implications come from the users' subjective responses to the examined factors. First, the designers should be aware that peoples may be wrong while assessing their future performance and, secondly, their preferences may be quite easily changed after allowing users to test proposed solutions. As a result of this, interactive systems designers should familiarize potential customers with products and make design changes after taking into account both the objective and subjective factors.

There are currently available technical possibilities of taking advantage of this study results in newly developed applications. In many visual programming languages one may specify background colors for the toolbar items – for example this study research software takes advantage of such a property.

Software developers may additionally consider allowing users to freely group the buttons by means of changing the button's background colors, after the software product is compiled and distributed, e.g. by adding appropriate properties to the application settings module.

Although toolbars are usually associated with classical stand-alone software, similar graphical structures may be used in various other circumstances such as mobile devices interfaces, electronic kiosks, web sites, automatic tellers machines, visually based controllers used in manufacturing, recently also in cars or other vehicles. The current study results may especially reflect real life situations where: (1) users either meet the graphical interface for the first time or (2) the arrangement of options consist of so many items that it is impossible to remember them or process them in the working (short term) memory or (3) time between consecutive uses is long enough to forget what the layout of target items.

4.4. Limitations and future research

There are a few limitations that should be considered while interpreting the results of this study. Almost all of the participants were young and experienced computer users, so the 'search and click' results may not be applicable to people less skilled in manipulating the cursor with a computer mouse. Second, only comparatively small, square shaped buttons with letters and digits were used in this study and it is not certain whether the results could be extended to target objects having other shapes and more complex icons. This may also be a worthwhile further investigation topic. Moreover, one should bear in mind that the findings come from fully controlled experiments, and though the design is much closer to the human-computer interaction reality, it might still require additional research in a more ecologically valid environment.

In the current study, the sex differences were neither controlled nor examined, but there are some studies showing significance of this factor both in the 'search and click' tasks (Michalski, 2008) as well as regarding the color preferences (e.g. Helson and Lansford, 1970; Ou et al., 2004). On the other hand, other investigations gave no evidence for the importance of the gender effect with regard to color preference (Granger, 1955; Guilford and Smith, 1959) or the acquisition times (Michalski et al., 2006). Although the previous findings are inconsistent, it seems to be interesting to explore this topic in future works. Some scientists also shown that color preferences might result from culture differences (e.g. Saito, 1996) so this could also be verified in subsequent visual display design studies.

There were some works showing the influence of various color combinations on subjects' preferences (e.g. Ou et al., 2004; Humar et al., 2008; Wu et al., 2009; Schloss and Palmer, 2011). In the present study experiments, however, only the preferences towards the toolbar location and background pattern were compared because the applied AHP technique allows exclusively for relative analysis within a specific group. Further research may directly examine preferences towards various color combinations in one experimental group which will enable the comparison of people perception regarding two (or more) colored graphical structures.

The inclusion in future research other background patterns along with additional colors varied not only in hue but also in other colorimetric properties will also be interesting. Additionally, more research is needed to verify to what extent, the observed in this study change in preferences is stable over time.

5. Conclusions

The main objective of this paper was to explore toolbar color properties in simple direct manipulation tasks involving both the goal directed movement as well as a search process. Unlike prior studies in this regard, the current investigation examined colors having the same perceptual differences. Much attention was also given to the role of various background color patterns used as grouping factors.

Generally, the obtained findings showed very strong and positive influence of the color grouping effect both on the speed of finding and selecting the target as well as for the accuracy of performing the tasks. The colored panels were significantly better operated than single colored suggesting possible change in subjects' visual search behavior and taking advantage of preattentive properties of the two colored toolbars.

As to the users' perception of the examined layouts, their preferences were meaningfully differentiated by background color patterns and decidedly changed after gaining some experience with operating them. The alteration of the subjective opinions corresponded to the efficiency and effectiveness of the explored experimental conditions, which was illustrated by formal models.

The demonstrated in this study results seem to be very interesting, in some areas even surprising. The use of color as an effective grouping factor was advocated by the Gestalt theorists, and later repeated in some studies. However, the comparatively small number of experimental research dealt with the colors influence on performance and preferences in the context of the graphical interface design. As it can be seen from this work and from the previous study of Michalski and Grobelny (2008), the findings obtained on the grounds of general vision research are not always fully applicable in the design of graphical interfaces of interactive systems. Furthermore, the color preferences are also strongly influenced by the type of the object to which they are applied (Taft, 1996; Schloss et al., 2012). These arguments should encourage scientists to continue studies in this field.

Although the research was focused on toolbars, the use of easily recognized icons and typical 'search and click' tasks along with the extensive exploration of both objective and subjective data may be useful in the analysis of other graphical structures in the context of a user interface design.

Acknowledgements

I would like to thank Ewelina Cuper for her assistance in collecting the empirical data and I express my great appreciation to all of the students who participated in this study. I am also very grateful to Jerzy Grobelny for his scientific support and stimulating discussions about this research. I appreciate the suggestions and comments of the anonymous reviewers which substantially contributed to improving the quality of the paper. The work was partially financially supported by National Science Center grant no. 2011/03/B/ST8/06238.

References

Anderson, J.R., Matessa, M. and Lebiere, C., 1997, ACT-R: a theory of higher level cognition and its relation to visual attention. Human–Computer Interaction, 12, 439–462. doi: 10.1207/s15327051hci1204_5

Backs, R., Walrath, L. & Hancock, A. (1987). Comparison of horizontal and vertical menu formats. In 31st Annual Meeting of the Human Factors Society (Santa Monica: Human Factors and Ergonomics Society), 715–717.

Bichot, N. P., Rao, S. C., & Schall, J. D. (2001). Continuous processing in macaque frontal cortex during visual search. Neuropsychologia, 39(9), 972–982. doi:10.1016/S0028-3932(01)00022-7

Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. Science, 308(5721), 529–534. doi:10.1126/science.1109676

Bodrogi, P. (2003). Chromaticity contrast in visual search on the multi-colour user interface. Displays, 24(1), 39–48. doi:10.1016/S0141-9382(02)00070-7

Borchers, J., Deussen, O., Klingert, A., & Knörzer, C. (1996). Layout rules for graphical web documents. Computers & Graphics, 20(3), 415–426. doi:10.1016/0097-8493(96)00011-8

Borchers, J., Deussen, O., Knörzer, C. (1995). Getting it across: layout issues for kiosk systems. SIGCHI Bulletin, 27(4), 68–74. doi:10.1145/214132.214170

Buchner, A., & Baumgartner, N. (2007). Text – background polarity affects performance irrespective of ambient illumination and colour contrast. Ergonomics, 50(7), 1036–1063. doi:10.1080/00140130701306413

Byrne, M.D., 2001, ACT-R/PM and menu selection: applying a cognitive architecture to HCI. International Journal of Human–Computer Studies, 55, 41–84. doi: 10.1006/ijhc.2001.0469.

Campbell, C. S., & Maglio, P. P. (1999). Facilitating navigation in information spaces: Road-signs on the World Wide Web. International Journal of Human-Computer Studies, 50(4), 309–327. doi:10.1006/ijhc.1998.0246

Change, D., Dooley, L. & Tuovinen, J. E., (2002). Gestalt theory in visual screen design. A new look at an old subject. In Proceedings of the seventh world conference on computers in education: Australian topics - Volume 8 (Darlinghurst, Australia: Australian Computer Society, Inc.), pp. 5–12.

Christ, R. E. (1975). Review and analysis of color coding research for visual displays. Human Factors, 17(6), 542–570. doi:10.1177/001872087501700602

Clynes M. & Kohn M. (1968). Recognition of visual stimuli from the electrical response of the brain, In: Kline NS, E. Laska E, editors. Computers and Electronic Devices in Psychiatry. New York: Grune and Stratton.

Clynes, M. (1977). Sentics: the touch of emotions. Garden City, N.Y.: Anchor Press.

Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. Annual Review of Neuroscience, 22(1), 319–349. doi:10.1146/annurev.neuro.22.1.319

CyberWorld (2012). Virtual Reality Products - Data Gloves, accessed on 12.12.2012.

Dean, M., Wu, S.-W., & Maloney, L. T. (2007). Trading off speed and accuracy in rapid, goal-directed movements. Journal of Vision, 7(5). doi:10.1167/7.5.10

Deininger, R. L. (1960). Human factors engineering studies of the design and use of pushbutton telephone sets. *Bell System Technical Journal*, XXXIX, 995–1012.

Eriksen, C. W., & James, J. D. S. (1986). Visual attention within and around the field of focal attention: A zoom lens model. Perception & Psychophysics, 40(4), 225–240. doi:10.3758/BF03211502

Farris, J. S., Johnson, B. R., & Jones, K. S. (2006). Width guidelines for rectangular objects with penetrable and impenetrable borders. Behaviour & Information Technology, 25(1), 83–90. doi:10.1080/01449290500102086

Farris, J. S., Jones, K. S., & Anders, B. A. (2002). Factors affecting the usefulness of impenetrable interface element borders. Human Factors, 44(4), 578–591. doi:10.1518/0018720024496971

Findlay, J. M. & Gilchrist, I. D. (2003). Active vision. The psychology of looking and seeing. New York: Oxford University Press.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47(6), 381–391. doi:10.1037/h0055392

Fitts, P. M., & Peterson, J. R. (1964). Information capacity of discrete motor responses. Journal of Experimental Psychology, 67(2), 103–112. doi:10.1037/h0045689

Fleetwood, M. D., & Byrne, M. D. (2002). Modeling icon search in ACT-R/PM. Cognitive Systems Research, 3(1), 25–33. doi:10.1016/S1389-0417(01)00041-9

Fleetwood, M. D., & Byrne, M. D. (2006). Modeling the visual search of displays: a revised ACT-R model of icon search based on eye-tracking data. Human–Computer Interaction, 21(2), 153–197. doi:10.1207/s15327051hci2102_1

Forman, E.H., 1990. Random indices for incomplete pairwise comparison matrices. European Journal of Operational Research, 48(1), 153–155. doi:10.1016/0377-2217(90)90072-J

Gerard, R. M. (1958). Differential effects of colored lights on psychophysical functions. Unpublished doctorial dissertation, University of California, Los Angeles, California.

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. International Journal of Industrial Ergonomics, 24(6), 631–645. doi:10.1016/S0169-8141(98)00068-7

Granger, G. W. (1955). An experimental study of colour preferences. The Journal of General Psychology, 52(1), 3–20. doi:10.1080/00221309.1955.9918340

Greenstein, J. S. & Arnaut, L. Y. (1988). Input devices. In Handbook of Human-Computer Interaction, M. Helander (Ed.) (Amsterdam: Elsevier Science Publishers B.V), 485–519.

Grobelny, J., Karwowski, W., & Drury, C. (2005). Usability of graphical icons in the design of human-computer interfaces. International Journal of Human-Computer Interaction, 18(2), 167–182. doi:10.1207/s15327590ijhc1802_3

Guilford, J. P., & Smith, P. C. (1959). A System of color-preferences. The American Journal of Psychology, 72(4), 487. doi:10.2307/1419491

Hall, R. H., & Hanna, P. (2004). The impact of web page text-background colour combinations on readability, retention, aesthetics and behavioural intention. Behaviour & Information Technology, 23(3), 183–195. doi:10.1080/01449290410001669932

Halverson, T., & Hornof, A. J. (2012). A computational model of "active vision" for visual search in human–computer interaction. Human–Computer Interaction, 26(4), 285–314. doi:10.1080/07370024.2011.625237

Hand, C. (1997). A Survey of 3D interaction techniques. Computer Graphics Forum, 16(5), 269–281. doi:10.1111/1467-8659.00194

Helson, H., & Lansford, T. (1970). The role of spectral energy of source and background color in the pleasantness of object colors. Applied optics, 9(7), 1513–1562. doi:10.1364/AO.9.001513

Ho, W. (2008). Integrated analytic hierarchy process and its applications – A literature review. European Journal of Operational Research, 186(1), 211–228. doi:10.1016/j.ejor.2007.01.004

Hoffmann, E. R., & Lim, J. T. A. (1997). Concurrent manual-decision tasks. Ergonomics, 40(3), 293–318. doi:10.1080/001401397188161

Humar, I., Gradisar, M., & Turk, T. (2008). The impact of color combinations on the legibility of a Web page text presented on CRT displays. International Journal of Industrial Ergonomics, 38(11–12), 885–899. doi:10.1016/j.ergon.2008.03.004

ISO 9126–1 (1998). Software product quality, Part 1: Quality model, International Standard.

ISO 9241–11 (1998). Ergonomic requirements for office work with visual display terminals (VDTs), Part 11: Guidance on usability, International Standard.

Jones, K. S., Johnson, B. R., & Farris, J. S. (2005). GUI objects with impenetrable borders: Instruction (not practice) makes perfect. International Journal of Human-Computer Studies, 62(6), 687–712. doi:10.1016/j.ijhcs.2005.02.002

Kalbach, J. & Bosenick, T. (2003). Web page layout: A comparison between left- and right-justified site navigation menus. Journal of Digital Information, 4, Article No. 153, 2003-04-28. Retrieved from http://journals.tdl.org/jodi/article/view/94/93 on February 12, 2011.

Kieras, D.E. & Meyer, D.E., 1997, An overview of the EPIC architecture for cognition and performance with application to human–computer interaction. Human–Computer Interaction, 12, 391–438. doi: 10.1207/s15327051hci1204_4

Kroemer, K. H. E. (2001). Keyboards and keying An annotated bibliography of the literature from 1878 to 1999. Universal Access in the Information Society, 1(2), 99–160. doi:10.1007/s102090100012

Kustov, A. A., & Lee Robinson, D. (1996). Shared neural control of attentional shifts and eye movements. Nature, 384(6604), 74–77. doi:10.1038/384074a0

Le Courrier du Livre (1912). Lisibilite des affiches en couleurs, Sheldons Limited House, Cosmos, p. 255.

Ling, J., & Van Schaik, P. (2002). The effect of text and background colour on visual search of Web pages. Displays, 23(5), 223–230. doi:10.1016/S0141-9382(02)00041-0

MacKenzie, I. S. (1991). Fitts' law as a performance model in human-computer interaction. PhD thesis, University of Toronto.

MacKenzie, I. S. (1992). Fitts' law as a research and design tool in human-computer interaction. Hum.-Comput. Interact., 7(1), 91–139. doi:10.1207/s15327051hci0701_3

Matthews, M. L., Lovasik, J. V., & Mertins, K. (1989). Visual performance and subjective discomfort in prolonged viewing of chromatic displays. Human Factors, 31(3), 259–271. doi:10.1177/001872088903100302

Michalski, R. (2005). Komputerowe wspomaganie badań jakości ergonomicznej oprogramowania (Computer-aided research of software ergonomic quality). PhD thesis, Wroclaw University of Technology.

Michalski, R. (2008). The effects of panel location, target size, and gender on efficiency in simple direct manipulation tasks, Human Computer Interaction, Ioannis Pavlidis (Ed.), ISBN: 978-953-7619-19-0, InTech, Available from: http://www.intechopen.com/books/human_computer_interaction/the_effects_of_panel_location__target_size__and_gender_on_efficiency_in_simple_direct_manipulation_t

Michalski, R. (2011). Examining users' preferences towards vertical graphical toolbars in simple search and point tasks. Computers in Human Behavior, 27(6), 2308–2321. doi:10.1016/j.chb.2011.07.010

Michalski, R., & Grobelny, J. (2008). The role of colour preattentive processing in human–computer interaction task efficiency: A preliminary study. International Journal of Industrial Ergonomics, 38(3–4), 321–332. doi:10.1016/j.ergon.2007.11.002

Michalski, R., Grobelny, J., & Karwowski, W. (2006). The effects of graphical interface design characteristics on human–computer interaction task efficiency. International Journal of Industrial Ergonomics, 36(11), 959–977. doi:10.1016/j.ergon.2006.06.013

Microsoft (2012a). Kinect - Xbox.com, http://www.xbox.com/en-US/kinect, accessed on 12.12.2012.

Microsoft (2012b). Freehand 3D computer interaction without gloves, http://research.microsoft.com /apps/video/dl.aspx?id=173838, accessed on 12.12.2012.

Murata, A. & Furukawa, N. (2005). Relationships among display features, eye movement characteristics, and reaction time in visual search. *Human Factors*, **47(3)**, 598–612. doi:10.1518/001872005774860032

Murch, G. M. (1985). Colour graphics—blessing or ballyhoo? Computer Graphics Forum, 4(2), 127–135. doi:10.1111/j.1467-8659.1985.tb00202.x

Näsänen, R. & Ojanpää, H. (2003). Effect of image contrast and sharpness on visual search for computer icons. Displays, 24(3), 137–144. doi:10.1016/j.displa.2003.09.003

Näsänen, R., Ojanpää, H., & Kojo, I. (2001). Effect of stimulus contrast on performance and eye movements in visual search. Vision Research, 41(14), 1817–1824. doi:10.1016/S0042-6989(01)00056-6

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized Linear Models. Journal of the Royal Statistical Society. Series A (General), 135(3), 370-384. doi:10.2307/2344614

Net Industries (2012). Eye - Evolution Of The Eye, Anatomy And Function Of The Human Eye, Retina, Optic Nerve, http://science.jrank.org/pages/2655/Eye.html, accessed on 13.12.2012.

Nielsen, J., & Levy, J. (1994). Measuring usability: preference vs. performance. Communications of the ACM, 37(4), 66–75. doi:10.1145/175276.175282

Niemelä, M., & Saarinen, J. (2000). Visual search for grouped versus ungrouped icons in a computer interface. Human Factors, 42(4), 630–635. doi:10.1518/001872000779697999

Nintendo (2012). Wii U Official Site at Nintendo, http://www.nintendo.com/wiiu, accessed on 12.12.2012.

Ojanpää, H., & Näsänen, R. (2003). Effects of luminance and colour contrast on the search of information on display devices. Displays, 24(4–5), 167–178. doi:10.1016/j.displa.2004.01.003

Ou, L.-C., Luo, M. R., Woodcock, A., & Wright, A. (2004). A study of colour emotion and colour preference. Part II: Colour emotions for two-colour combinations. Color Research & Application, 29(4), 292–298. doi:10.1002/col.20024

Pearson, R. & Schaik, P. (2003). The effect of spatial layout of and link color in web pages on performance in a visual search task and an interactive search task. International Journal of Human-Computer Studies, 59(3), 327–353. doi:10.1016/S1071-5819(03)00045-4

Plamondon, R., & Alimi, A. M. (1997). Speed/accuracy trade-offs in target-directed movements. Behavioral and Brain Sciences, 20(2), 279–303; discussion 303–349. doi:10.1017/S0140525X97001441

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. Journal of Experimental Psychology: General, 109(2), 160–174. doi:10.1037/0096-3445.109.2.160

Robertson, A.R. (1977). The CIE 1976 color-difference formulae. Color Research and Application, 2, 7-11

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology, 15(3), 234–281. doi:10.1016/0022-2496(77)90033-5

Saaty, T. L. (1980). The analytic hierarchy process. New York: McGraw-Hill.

Saito, M. (1996). A comparative study of color preferences in Japan, China and Indonesia, with emphasis on the preference for white. Perceptual and Motor Skills, 83(1), 115–128. doi:10.2466/pms.1996.83.1.115

Salvucci, D.D. (2001). An integrated model of eye movements and visual encoding. Cognitive Systems Research, 1(4), 201–220. doi:10.1016/S1389-0417(00)00015-2

Schaik, P. & Ling, J. (2001). The effects of frame layout and differential background contrast on visual search performance in web pages. Interacting with Computers, 13(5), 513–525. doi:10.1016/S0953-5438(00)00054-0

Schaik, P., & Ling, J. (2003). The effect of link colour on information retrieval in educational intranet use. Computers in Human Behavior, 19(5), 553–564. doi:10.1016/S0747-5632(03)00004-9

Schloss, K. B., & Palmer, S. E. (2011). Aesthetic response to color combinations: preference, harmony, and similarity. Attention, Perception, & Psychophysics, 73(2), 551–571. doi:10.3758/s13414-010-0027-0

Schloss, K. B., Strauss, E. D., & Palmer, S. E. (2012). Object color preferences. Color Research & Application. doi:10.1002/col.21756

Seow, S. C., (2005). Information theoretic models of HCI: a comparison of the Hick-Hyman law and Fitts' law. Human Computer Interaction, 20(3), 315–352. doi:10.1207/s15327051hci2003_3

Shneiderman, B. (1982). The future of interactive systems and the emergence of direct manipulation. Behaviour & Information Technology, 1(3), 237–256. doi:10.1080/01449298208914450

Shneiderman, B. (1983). Direct manipulation: A step beyond programming languages. Computer, 16(8), 57 –69. doi:10.1109/MC.1983.1654471

Sony (2012). PlayStation®Move Motion Controller - PlayStation®3 Move Info, Games & Updates, http://us.playstation.com/ps3/playstation-move/, accessed on 12.12.2012.

Taft, C. (1997). Color meaning and context: Comparisons of semantic ratings of colors on samples and objects. Color Research & Application, 22(1), 40–50. doi:10.1002/(SICI)1520-6378(199702)22:1<40::AID-COL7>3.0.CO;2-4

Treisman, A. M. (1982). Perceptual grouping and attention in visual search for features and for objects. Journal of Experimental Psychology: Human Perception and Performance, 8(2), 194–214. doi:10.1037/0096-1523.8.2.194

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, 12(1), 97–136. doi:10.1016/0010-0285(80)90005-5

Welford, A. T. (1960). The measurement of sensory-motor performance: survey and reappraisal of twelve years' progress. Ergonomics, 3(3), 189–230. doi:10.1080/00140136008930484

Wilson, G. D. (1966). Arousal properties of red versus green. Perceptual and Motor Skills, 23(3), 947–949. doi:10.2466/pms.1966.23.3.947

Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. Psychonomic Bulletin & Review, 1(2), 202–238. doi:10.3758/BF03200774

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? Nature Reviews Neuroscience, 5(6), 495–501. doi:10.1038/nrn1411

Wolfe, J. M., 2007, Guided Search 4.0: Current Progress with a model of visual search. In W. Gray (Ed.), Integrated Models of Cognitive Systems (pp. 99-119). New York: Oxford.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. Journal of Experimental Psychology: Human Perception and Performance, 15(3), 419–433. doi:10.1037/0096-1523.15.3.419

Wu, F.-G., & Chen, C.-Y. (2009). Effects of color display, color name, color formation and color alignment on the screen layout usability of customized product color combinations. International Journal of Industrial Ergonomics, 39(5), 655–666. doi:10.1016/j.ergon.2009.04.004

Wu, F.-G., Chen, C.-Y., Lee, Y.-J., & Chen, R. (2010). Effects of color sample display and color sample grouping on screen layout usability for customized product color selection. Computers in Human Behavior, 26(1), 51–60. doi:10.1016/j.chb.2009.08.003

Wu, J.-H., & Yuan, Y. (2003). Improving searching and reading performance: the effect of highlighting and text color coding. Information & Management, 40(7), 617–637. doi:10.1016/S0378-7206(02)00091-5

Wurtz, R. H., Goldberg, M. E. & Robinson, D. L. (1982). Brain mechanisms of visual attention. *Scientific American*, 246, 124–135.

Zahedi, F. (1986). The Analytic Hierarchy Process—A survey of the method and its applications. Interfaces, 16(4), 96–108. doi:10.1287/inte.16.4.96

Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. Psychological Science, 8(6), 448–453. doi:10.1111/j.1467-9280.1997.tb00459.x

Tables

Table 1. Colors used in toolbars.

Color name	Color sample	CIE Lab	Decimal RGB	Hexadecimal RGB	MS Visual Basic palette
Grey		(84, 0, 5)	(213, 209, 200)	#D5 D1 C8	#C8 D1 D5
Red		(84, 23, 9)	(255, 193, 193)	#FF C1 C2	#C2 C1 FF
Green		(84, -11, 30)	(205, 215, 150)	#CD D7 96	#96 D7 CD
Blue		(84, -12, -9)	(175, 217, 226)	#AF D9 E2	#E2 D9 AF

Table 2. The assignment of the 24 experimental conditions: [(*three target background color pairs*) × (*three toolbar color patterns*) + *three toolbars with single target background color*] × (*two toolbar location on the screen*) to groups. Every group of participants investigated 10 toolbar variants involving two out of three examined colors: [(*one target background color pair*) × (*three toolbar color patterns*) + *two toolbars with single target background color*] × (*two toolbar location on the screen*) = (1 × 3 + 2) × 2 = 10.

No	Toolbar background colors	Target location	Toolbar color pattern	Toolbar Symbol	Group 1 (No)	Group 2 (No)	Group 3 (No)
1.	Red and Green	Left upper corner	02×09	L_RG_02×09	× (1)		
2.			06×03	L_RG_06×03	× (2)		
3.			18×01	L_RG_18×01	× (3)		
4.		Right upper corner	02×09	R_RG_02×09	× (4)		
5.			06×03	R_RG_06×03	× (5)		
6.			18×01	R_RG_18×01	× (6)		
7.	Red and Blue	Left upper corner	02×09	L_RB_02×09		× (1)	
8.			06×03	L_RB_06×03		× (2)	
9.			18×01	L_RB_18×01		× (3)	
10.		Right upper corner	02×09	R_RB_02×09		× (4)	
11.			06×03	R_RB_06×03		× (5)	
12.			18×01	R_RB_18×01		× (6)	
13.	Green and Blue	Left upper corner	02×09	L_GB_02×09			× (1)
14.			06×03	L_GB_06×03			× (2)
15.			18×01	L_GB_18×01			× (3)
16.		Right upper corner	02×09	R_GB_02×09			× (4)
17.			06×03	R_GB_06×03			× (5)
18.			18×01	R_GB_18×01			× (6)
19.	Red	Left upper corner	—	L_R	× (7)	× (7)	
22.		Right upper corner	—	R_R	× (8)	× (8)	
20.	Green	Left upper corner	—	L_G	× (9)		× (7)
23.		Right upper corner	—	R_G	× (10)		× (8)
21.	Blue	Left upper corner	—	L_B		× (9)	× (9)
24.		Right upper corner	—	R_B		× (10)	× (10)

Table 3. Overall descriptive statistics for all correctly performed 'search and click' tasks.

Parameter	Value
Valid cases	6349
Means	
Arithmetic	2042 [ms]
Geometric	1753 [ms]
Harmonic	1566 [ms]
Median	1625 [ms]
Minimum	594 [ms]
Maximum	20 110 [ms]
Variance	2 100 182 $[ms]^2$
Standard Deviation	1449 [ms]

Mean Standard Error	18.2 [ms]
Skewness	3.6
Kurtosis	21

Table 4. Individual descriptive statistics of selection times for all of the investigated toolbars

No	Toolbar symbol	N*	Mean [ms]	MSE** [ms]	Median [ms]	Min [ms]	Max [ms]	SD*** [ms]	Skew-ness	Kur-tosis
1.	L_RG_02×09	207	1853	91	1453	735	9984	1308	3.5	16
2.	L_RG_06×03	207	1944	82	1640	719	8563	1178	2.6	8.7
3.	L_RG_18×01	203	2256	118	1766	672	12 360	1679	2.7	9.6
4.	L_RB_02×09	195	**1560**	64	1312	734	8687	899	4.1	25
5.	L_RB_06×03	195	1829	73	1563	703	8000	1019	3.1	12
6.	L_RB_18×01	197	2080	88	1750	765	10 015	1230	2.3	8.8
7.	L_GB_02×09	237	1753	81	1390	719	10 922	1251	3.9	21
8.	L_GB_06×03	237	1769	89	1484	672	16 578	1368	6.5	61
9.	L_GB_18×01	230	2280	116	1680	625	10 672	1761	2.3	6
10.	L_R	391	2219	79	1813	750	16 734	1564	3.6	22
11.	L_G	435	2239	76	1782	625	15 547	1585	3	15
12.	L_B	424	2228	77	1828	594	15 625	1576	3.4	18
13.	R_RG_02×09	208	1688	59	1414	750	5250	847	1.9	3.9
14.	R_RG_06×03	207	1773	69	1516	719	9031	986	3.1	16
15.	R_RG_18×01	206	2248	123	1750	703	11 781	1761	2.9	9.6
16.	R_RB_02×09	197	1678	65	1422	719	6469	914	2.5	8.7
17.	R_RB_06×03	196	1824	85	1562	750	10 984	1189	4.3	26
18.	R_RB_18×01	195	2033	117	1625	719	16 500	1637	4.7	33
19.	R_GB_02×09	240	1707	73	1328	594	7422	1129	2.5	7.2
20.	R_GB_06×03	240	1789	69	1578	750	11 859	1076	4.7	36
21.	R_GB_18×01	234	2139	123	1633	719	20 110	1875	5.3	41
22.	R_R	398	2239	71	1852	734	11 016	1410	2.4	7.9
23.	R_G	441	**2319**	81	1844	625	14 828	1694	3.2	15
24.	R_B	429	2292	77	1875	672	13 297	1592	3	13

* N – number of valid cases
** MSE – Mean Standard Error
*** SD – Standard Deviation

Table 5. GLZM Analysis of variance results for the selections times (excluding single color toolbars)

Factor	df	Wald statistics	p
Toolbar location (TL)	1	1.003	0.32
Toolbar background color pair (TBP)	2	9.6	* 0.0081
Toolbar background color pattern (TCP)	2	123	*< 0.000001
TL × TBP	2	1.9	0.38
TL × TCP	2	0.53	0.77
TBP × TCP	4	3.9	0.42
TL × TBP × TCP	4	5.4	0.25

* $p < 0.01$

Table 6. GLZM Analysis of variance results for the selections times depending on the *Toolbar location* and *Toolbar background color*

Factor	df	Wald statistics	p
Toolbar Location (TL)	1	0.0058	0.94
Toolbar Background Colors (TBC)	5	20	*< 0.000001
TL × TBC	5	0.72	0.61

* $p < 0.000001$

Table 7. Series of GLZM one way Anovas for the *Toolbar background color* levels. (*Wald* statistics, *p* values in brackets)

	RG	RB	GB	R	G	B
RG	×	9.9 (**0.0016)	1.9 (0.17)	27 (***< 0.00001)	37 (***< 0.00001)	33 (***< 0.00001)
RB		×	3.2 (*0.072)	65 (***< 0.00001)	80 (***< 0.00001)	75 (***< 0.00001)
GB			×	41 (***< 0.00001)	54 (***< 0.00001)	49 (***< 0.00001)
R				×	0.63 (0.43)	0.25 (0.62)
G					×	0.090 (*0.076)
B						×

* p < 0.1
** p < 0.005
*** p < 0.00001

Table 8. Series of GLZM one way Anovas for the *Toolbar backgroung color pattern* effect (*Wald* statistics, *p* values in brackets)

	02×09	06×03	18×01
02×09	×	11 (*0.0009)	128 (**< 0.00001)
06×03		×	72 (**< 0.00001)
18×01			×

* p < 0.001
** p < 0.00001

Table 9. Error results analysis by means of the χ^2 tests.

Factor	df	χ^2	p
Toolbar Location (TL)	1	7.4	* 0.0066
Toolbar Background Colors (TBC)	5	16	* 0.0056
Toolbar Color Pattern (TCP)	3	21	** 0.00013

* p < 0.01
** p < 0.005

Table 10. The effect of *Toolbar background colors* on errors. A series of χ^2 tests.

	RG	RB	GB	R	G	B
RG	×	0.54	0.65	* 0.0040	0.14	* 0.044
RB		×	0.28	* 0.023	0.38	0.16
GB			×	* 0.00067	** 0.054	* 0.012
R				×	0.19	0.42
G					×	0.61
B						×

* p < 0.05
** p < 0.1

Table 11. The effect of *Toolbar background color pattern* on errors. A series of χ^2 tests.

	02×09	06×03	18×01	None
02×09	×	0.73	** 0.0072	** 0.00030
06×03		×	* 0.018	** 0.0010
18×01			×	0.43
None				×

* p < 0.05
** p < 0.01

Table 12. Basic descriptive characteristics of *CR* values for all experimental conditions

Group No (Toolbar Background Colors)	Moment of Comparison	Mean CR	Median	*MSE	**SD	Min	Max	***N	No. of subjects with CR < 0.1
G1	Before	0.10	0.068	0.017	0.076	0.028	0.33	21	13
(RG)	After	0.081	0.061	0.014	0.064	0.015	0.29	21	17
G2	Before	0.11	0.082	0.021	0.094	0.017	0.44	20	13
(RB)	After	0.084	0.053	0.014	0.061	0.026	0.24	20	14
G3	Before	0.12	0.085	0.021	0.10	0.030	0.43	24	14
(GB)	After	0.094	0.096	0.0095	0.046	0.000	0.17	24	13

* MSE – Mean Standard Error
** SD – Standard Deviation
*** N – Number of valid cases

Table 13. Anova results regarding the effects of experimental group number and the moment of comparison on consistency ratios.

Factor	SS	df	MS	F	p	η^2
Group No	0.0076	2	0.0038	0.64	0.53	0.010
Moment of comparisons (MOC)	0.018	1	0.018	2.96	*0.088	0.023
Group No × MOC	0.00078	2	0.00039	0.066	0.94	0.0010
Error	0.74	124	0.0060			0.97
Total	0.76					

* $p < 0.1$; df – degrees of freedom; SS – sum of squares; MS – mean squares; η^2 – Eta squared

Table 14. The effects of the group number and the moment of comparisons on the number of subjects excluded from the AHP weight analysis.

Factor	df	χ^2	p
Group No	2	2.5	0.29
Moment of comparisons (MOC)	1	0.54	0.46

Table 15. Mean AHP weights (preferences) for all experimental conditions (mean standard error in brackets).

Toolbar location on the screen	Toolbar color pattern	Toolbar background colors: Group 1 (Red-Green) Prior	Group 1 (Red-Green) Post	Group 2 (Red-Blue) Prior	Group 2 (Red-Blue) Post	Group 2 (Green-Blue) Prior	Group 2 (Green-Blue) Post
Left upper corner	**02×09**	0.110 (0.0133)	**^0.140 (0.0152)**	0.121 (0.0107)	0.124 (0.0133)	0.096 (0.0072)	**^0.153 (0.0175)**
	06×03	**^0.123 (0.0126)**	0.137 (0.0118)	**^0.141 (0.0163)**	**^0.153 (0.0162)**	**^0.160 (0.0165)**	0.124 (0.0118)
	18×01	0.109 (0.0171)	0.091 (0.0163)	0.094 (0.0206)	0.083 (0.0133)	0.113 (0.0159)	0.070 (0.0094)
	Single color 1	0.078 (0.0109)	***0.061 (0.0064)**	***0.071 (0.0084)**	***0.064 (0.0097)**	0.073 (0.0101)	***0.065 (0.0118)**
	Single color 2	***0.076 (0.0138)**	0.063 (0.0068)	0.078 (0.0132)	0.069 (0.0109)	***0.071 (0.0116)**	0.084 (0.0131)
Right upper corner	**02×09**	0.104 (0.0100)	0.140 (0.0147)	0.114 (0.0132)	0.125 (0.0137)	0.098 (0.0078)	**^^0.165 (0.0215)**
	06×03	**^^0.133 (0.0157)**	**^^0.146 (0.0155)**	**^^0.143 (0.0160)**	**^^0.163 (0.0170)**	**^^0.140 (0.0135)**	0.132 (0.0114)
	18×01	0.109 (0.0150)	0.092 (0.0157)	0.091 (0.0167)	0.084 (0.0118)	0.107 (0.0172)	0.069 (0.0094)
	Single color 1	****0.073 (0.0093)**	0.066 (0.0086)	****0.0735 (0.0084)**	0.068 (0.0101)	****0.070 (0.0090)**	****0.068 (0.0119)**
	Single color 2	0.084 (0.0155)	****0.063 (0.0092)**	0.0736 (0.0092)	****0.067 (0.0107)**	0.074 (0.0116)	0.071 (0.0116)

* The lowest mean weight value among toolbars in the left upper corner
** The lowest mean weight value among toolbars in the right upper corner
^ The highest mean weight value among toolbars in the left upper corner
^^ The highest mean weight value among toolbars in the right upper corner

Table 16. The hierarchy of preferences based on the AHP relative weights. One denotes the highest preference.

Toolbar location on the screen	Toolbar color pattern	Toolbar background colors					
		Group 1 (Red-Green)		Group 2 (Red-Blue)		Group 2 (Green-Blue)	
		Prior	Post	Prior	Post	Prior	Post
Left upper corner	02×09	2	**1**	2	2	3	**1**
	06×03	**1**	2	**1**	**1**	**1**	2
	18×01	3	3	3	3	2	4
	Single col. 1	4	**5**	**5**	**5**	4	**5**
	Single col. 2	**5**	4	4	4	**5**	3
Right upper corner	02×09	3	2	2	2	3	**1**
	06×03	**1**	**1**	**1**	**1**	**1**	2
	18×01	2	3	3	3	2	4
	Single col. 1	**5**	4	**5**	4	**5**	**5**
	Single col. 2	4	**5**	4	**5**	4	3

Table 17. Anova analyses of preferences before and after the search and click tasks for the first experimental group testing the red-green layouts.

Factor (Group 1: red and green)	Prior						Post					
	SS	df	MS	F	p	η^2	SS	df	MS	F	p	η^2
Toolbar location on the screen (TL)	0.000061	1	0.000061	0.025	0.87	0.00018	0.00031	1	0.00031	0.11	0.74	0.00048
Toolbar background color pattern (TCP)	0.050	4	0.013	5.3	*0.00061	0.15	0.21	4	0.052	19	*<0.000001	0.33
TL × TCP	0.0015	4	0.00038	0.16	0.96	0.0044	0.00049	4	0.00012	0.046	0.996	0.00077
Error	0.29	120	0.0024			0.85	0.43	160	0.0027			0.67
Total	0.34						0.64					

* p < 0.001, df – degrees of freedom; SS - sum of squares; MS - mean squares; η^2 – Eta squared

Table 18. Anova analyses of preferences before and after the search and click tasks for the second experimental group testing the red-blue layouts.

Factor (Group 2: red and blue)	Prior						Post					
	SS	df	MS	F	p	η^2	SS	df	MS	F	p	η^2
Toolbar location on the screen (TL)	0.00012	1	0.00012	0.046	0.83	0.00030	0.00027	1	0.00027	0.12	0.73	0.00057
Toolbar background color pattern (TCP)	0.091	4	0.023	9.1	*0.000002	0.23	0.18	4	0.045	19	*<0.000001	0.37
TL × TCP	0.00044	4	0.00011	0.045	0.996	0.00114	0.00069	4	0.00017	0.075	0.99	0.0014
Error	0.30	120	0.0025			0.77	0.30	130	0.0023			0.63
Total	0.39						0.48					

* p < 0. 00001, df – degrees of freedom; SS - sum of squares; MS - mean squares; η^2 – Eta squared

Table 19. Anova analyses of preferences before and after the search and click tasks for the third experimental group testing the green-blue layouts.

Factor (Group 3: green and blue)	Prior						Post					
	SS	df	MS	F	p	η^2	SS	df	MS	F	p	η^2
Toolbar location on the screen (TL)	0.00085	1	0.00085	0.39	0.53	0.0021	0.00015	1	0.00015	0.063	0.80	0.0003
Toolbar background color pattern (TCP)	0.12	4	0.029	13	*<0.000001	0.29	0.18	4	0.044	19	*<0.000001	0.38
TL × TCP	0.0023	4	0.00058	0.26	0.90	0.0057	0.0024	4	0.00060	0.25	0.91	0.0052
Error	0.28	130	0.0022			0.70	0.28	120	0.0023			0.61
Total	0.40						0.46					

* p < 0. 00001, df – degrees of freedom; SS - sum of squares; MS - mean squares; η^2 – Eta squared

Table 20. Comparisons of the mean acquisition times obtained in this study toolbars with gray toolbars used in the work of Michalski et al. 2006. Series of GLZM Anovas (*Wald* statistics, *p* values in brackets).

	RG	RB	GB	R	G	B
Grey	58 (*< 0.00001)	118 (*< 0.00001)	84 (*< 0.00001)	1.8 (0.18)	0.23 (0.63)	0.67 (0.41)

* $p < 0.000001$

Figures

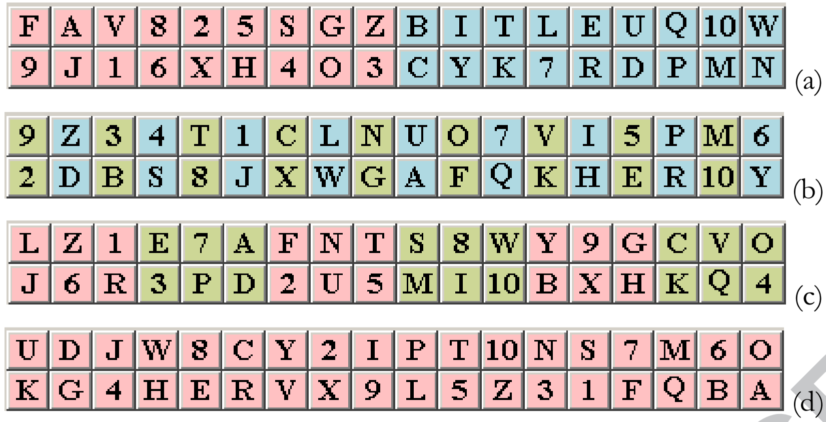


Figure 1. Exemplary toolbars used in experiments: (a) the red–blue panel with a pattern of two blocks (RB_02×09), (b) the green–blue panel including 18 groups (GB_18×01), (c) the red–green panel containing six sections (RG_06×03), (d) the toolbar with single background color.

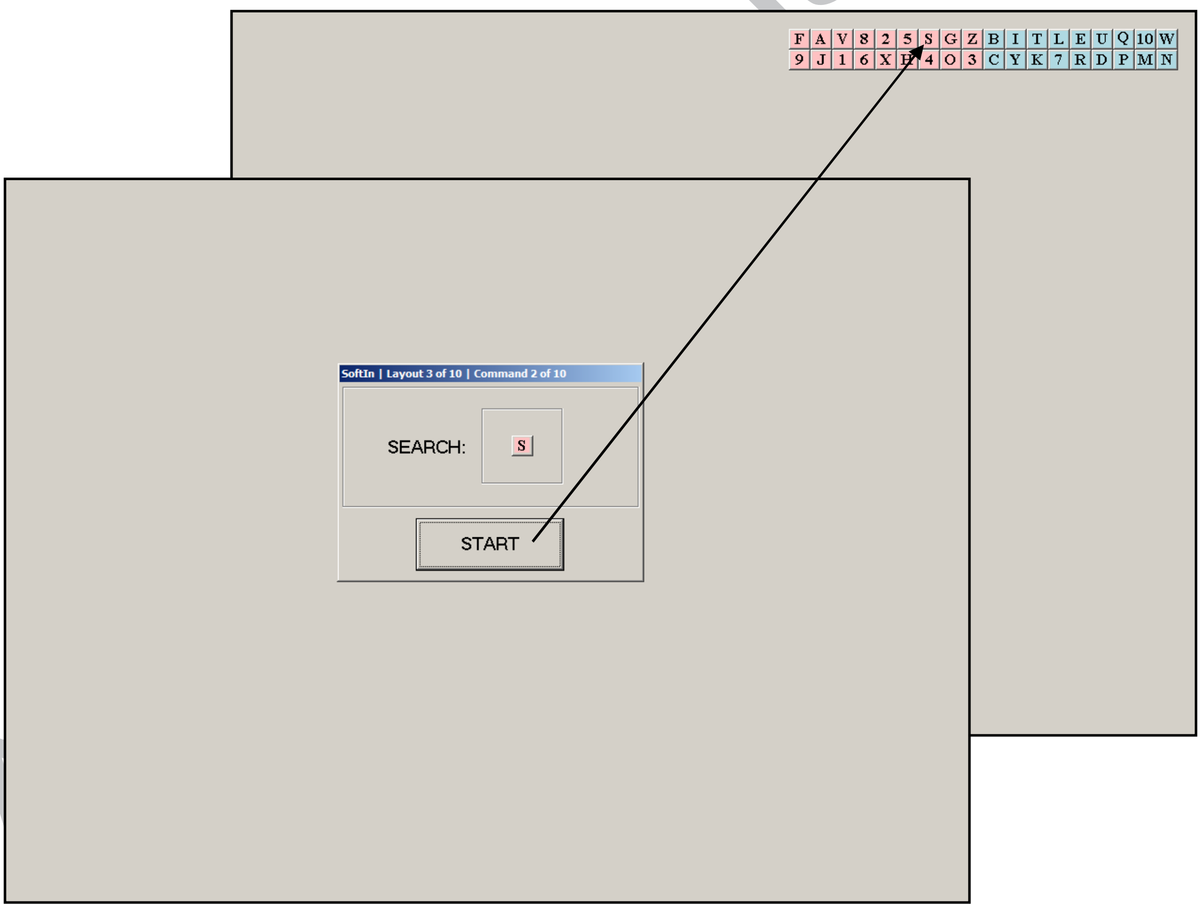


Figure 2. The illustration of the individual search and click task performed by the participants

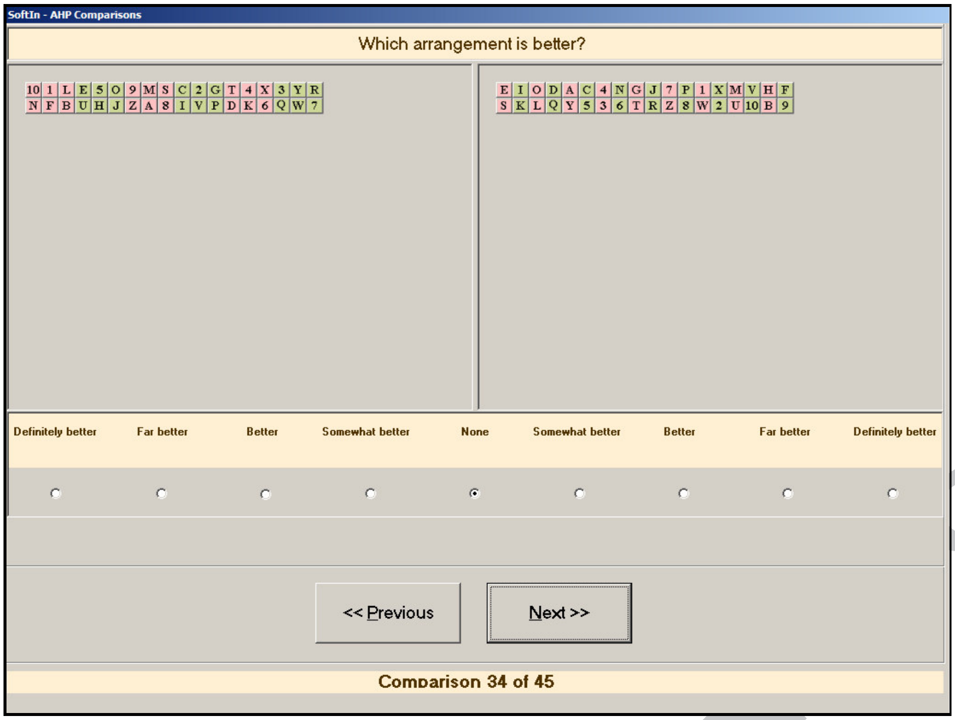


Figure 3. One of the software windows used for making comparisons

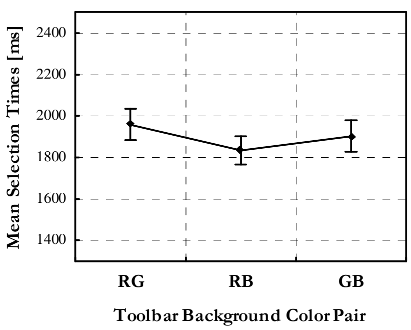


Figure 4. Effect of *Toolbar background color pairs* on mean selection times ($W = 9.6$, $p < 0.0081$). Whiskers denote mean standard errors.

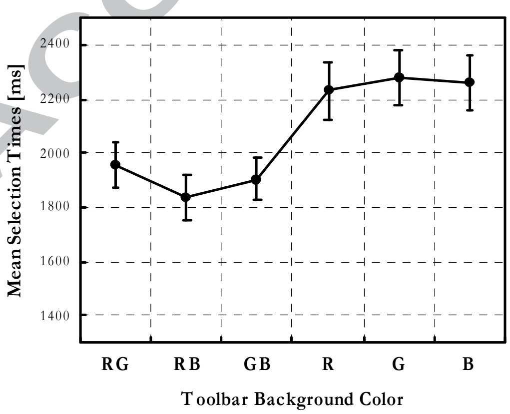


Figure 5. Effect of *Toolbar background color* on mean selection times. ($W = 20$, $p < 0.000001$). Whiskers denote mean standard errors.

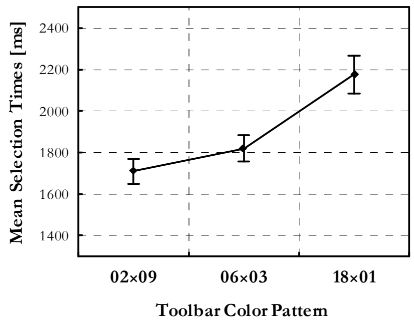


Figure 6. Effect of *Toolbar background color pattern* on mean acquisition times ($W = 123$, $p < 0.000001$). Whiskers denote mean standard errors.

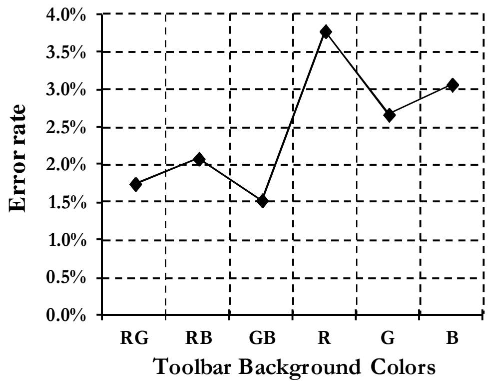


Figure 7. Error rates depending on the *Toolbar background colors* factor

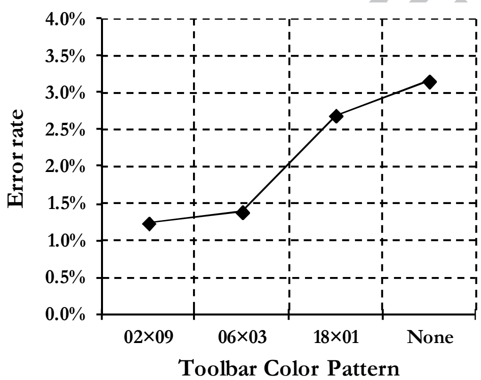


Figure 8. The effect of *Toolbar Color Pattern* on error rates

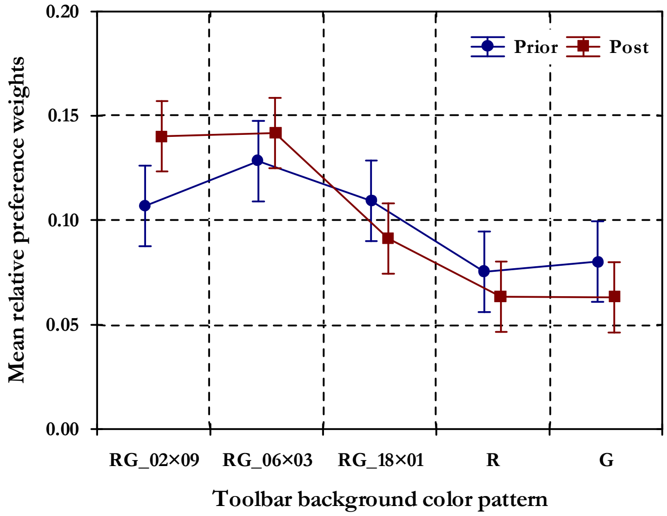


Figure 9. The mean preference weights for red and green layouts depending on *Toolbar background color pattern* prior and post the experimental trials (whiskers denote mean standard errors).

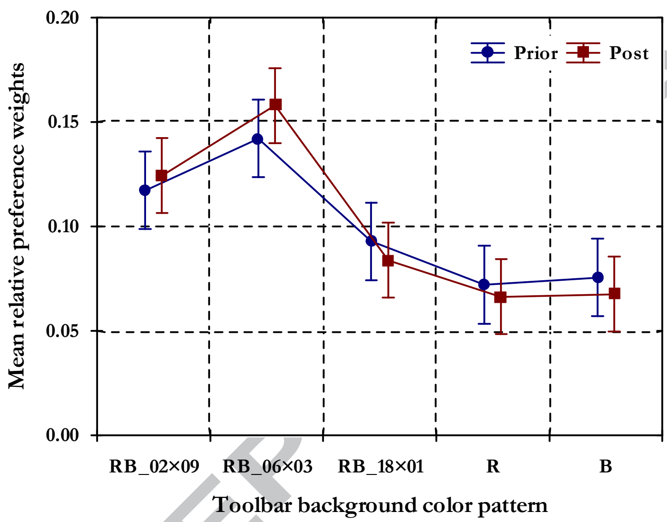


Figure 10. The mean preference weights for red and blue layouts depending on *Toolbar background color pattern* prior and post the experimental trials (whiskers denote mean standard errors).

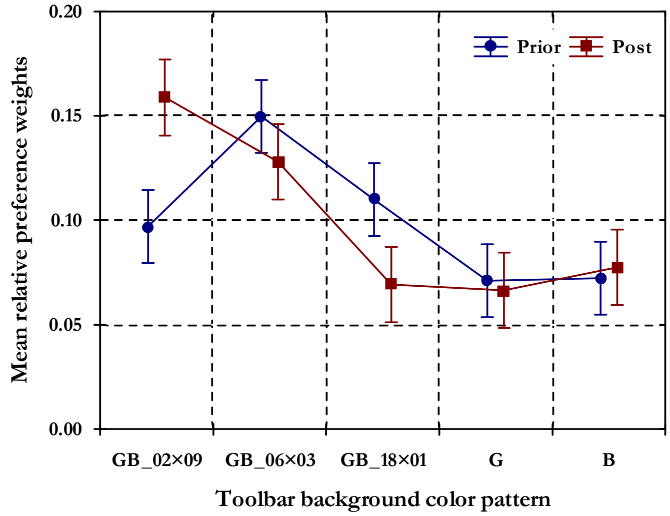


Figure 11. The mean preference weights for green and blue layouts depending on toolbar background color pattern prior and post the experimental trials (whiskers denote mean standard errors).

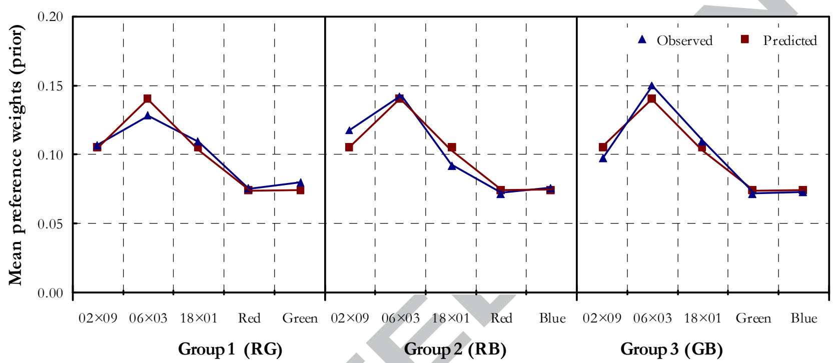


Figure 12. The observed and predicted mean relative weights for all experimental groups obtained before performance tasks. (*Mean preference weights prior*= 0.0746 + 0.0654 • $D_{06\times03}$ + 0.0309 • $D_{02\times09\_or\_18\times01}$, R^2 = 92.5%; $R^2_{adjusted}$ = 91%; F = 74, p < 0.000001)

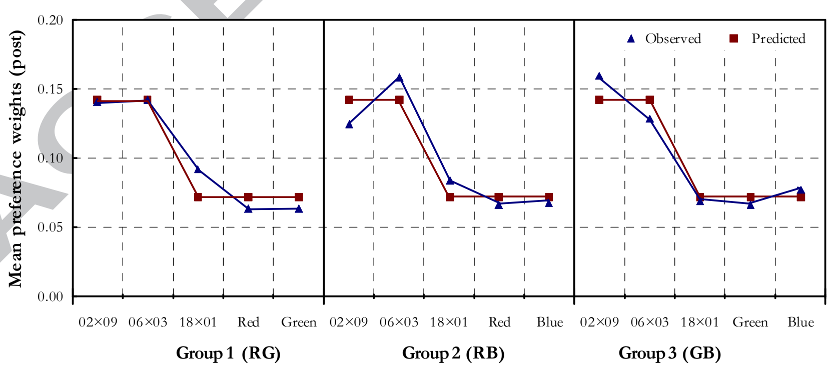


Figure 13. The observed and predicted mean relative weights for all experimental groups obtained after performance tasks. (*Mean preference weights post* = 0.0721 + 0.0697 • $D_{02\times09\_or\_06\times03}$, R^2 = 90.5%; $R^2_{adjusted}$ = 90%; F = 124, p < 0.000001)

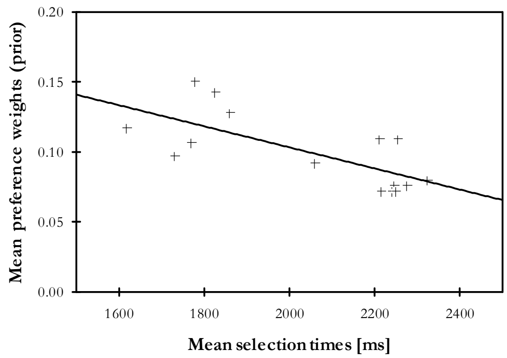


Figure 14. Linear regression of the mean preference weights obtained prior to search and click tasks as a function of mean selection times. (*Mean preference weights prior* = 0.254 – 0.0000753 • *Mean selection times*; R^2 = 51%; R^2_{adjusted} = 47%; F = 13, p = 0.0029)

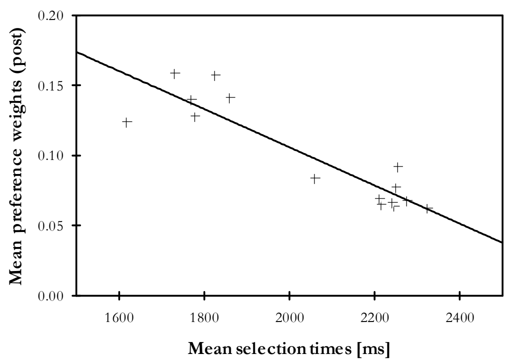


Figure 15. Linear regression of the mean preference weights obtained after the search and click tasks as a function of mean selection times. (*Mean preference weights post* = 0.378 – 0.000136 • *Mean selection times*; R^2 = 82%; R^2_{adjusted} = 81%; F = 59, p = 0.0000034)

Highlights

> Toolbars with background colors having the same perceptual distance are examined. > Smaller, uniformly colored areas are operated worse in terms of speed and accuracy than those divided into larger parts. >Preferences strongly depend on the toolbar color pattern and decidedly change after the performance tasks. > Regression models of preference change and relationship between speed and preferences are developed and discussed.