Mental Models for Online Shops -

The Influence of Location Typicality on Usability over Time

Master's Thesis of:

Philippe Chresta, B.Sc.

Faculty of Psychology

Department of Cognitive Psychology and Methodology

University of Basel

November 2011

Thesis Supervisors:

Alexandre Nicolas Tuch, M.Sc.

Department of Cognitive Psychology and Methodology, University of Basel

Prof. Dr. Klaus Opwis

Department of Cognitive Psychology and Methodology, University of Basel

Abstract

There are endless possibilities to place web-objects in websites. The location of web-objects is especially important in online shops, where user orientation, perceived usability and performance could lead to either sales increase, or worse, sales decrease. Previous studies showed that there is a general consent among users about typical and atypical placement of web-objects indicating robust mental models. Two versions of an online shop (one with atypical web-object locations and one with typical web-object locations) were used in this online study with n = 53 participants in order to investigate the influence of location typicality on usability, orientation and user performance. Results show, that user performance is independent from location typicality while perceived orientation and usability is better in the shop version designed according to users expectations. This effect is robust over time. Therefore one may assume that users seem to adapt quickly to unexpected web-object locations and are able to improve their performance over time. Websites corresponding with users mental models lead to more satisfaction with the usability, better orientation, and an improvement for the amount of clicks to execute tasks.

Mental Models for Online Shops –
The Influence of Location Typicality on Usability over Time

The number of websites in the Internet is enormous. An estimated 11.45 billion websites exist, increasing every day[1]. After the average Internet user has overcome the first obstacle, namely to find a requested website, the second challenge arrives. Every website looks different; there is a wide variety of interfaces with different layouts, different levels of complexity, varying functions, and many more features that can potentially overwhelm a user. As Otter and Johnson (2000) state: “Being disorientated or lost is one of the fundamental difficulties which user experience when trying to navigate within hypertext systems.” (p. 3).

Therefore, a fast and effective orientation on a website can decide about its success or lead to frustration which, in turn, can cause users to leave the website and browse for an existing alternative (Oulasvirta, Kärkkäinen, & Laarni, 2005). According to Vila and Kuster (2011) 60% of the time spent in online shops, users do not find the information they are looking for. More than 83% of Internet users leave a website if they feel that it is too complicated to find a product or service and approximately 58% of visitors who experience usability problems will not return to the website. The average e-commerce site, so the authors, could increase its sales by 100% if it had improved usability which causes an estimated sales loss of $24 billion[2] (Vila & Kuster, 2011).

Attentional mechanisms to improve orientation

In order to deal with information flood, people are equipped with attentional mechanisms that sort through this constant flood of sensory input (Sobel, Gerrie, Poole, & Kane, 2007). Two interacting and cooperating mechanisms, top-down and bottom-up processes, guide the users’ attention and help them to visually recognize objects. The

[1] Estimated by http://www.worldwidewebsize.com/ on November 9 2011
[2] Amount valued for the United States of America

bottom-up mechanisms are thought to work on raw sensory input. Attention is shifted rapidly, involuntarily and spontaneously towards potential important items in the actual visual field (Kim, Dey, Lee, & Forlizzi, 2011). Within bottom-up mechanisms, the concept of "pop-out" is important (Baldassi & Burr, 2004). Pop-out is the drawing of attention to an object, which is somehow remarkable. Design factors as color, movement and brightness can cause pop-out effects (Wolfe, 2001).

The top-down mechanisms, in contrast, conclude knowledge to select objects most likely to help. Attention, therefore, is goal-driven and cognitive strategies are used to shift visual attention (Connor, Egeth, & Yantis, 2004). According to Sobel, et al. (2007) top-down and bottom-up mechanisms often interact: Top-down mechanisms underlie the ability to control the focus of attention, therefore, users are able to ignore distractors that activate bottom-up mechanisms. Web designers, therefore, can direct the users' attention to desired locations by triggering bottom-up processes, for example with a pop-out effect, or they can place web-objects at locations where users would expect them to be in order to trigger top-down mechanisms (Roth, Schmutz, Pauwels, Bargas-Avila, & Opwis, 2010).

Through repeated interaction, users have built *mental models* of websites that reflect the knowledge they gained through experience (Thatcher & Greyling, 1998). Mental models are top down mechanisms that include available functionality, effects of functions and the locations of those functions (Hui, Partridge, & Boutilier, 2009). The term *mental model* is used in psychology to denote a person's mental representation of (concrete or abstract) objects (Johnson-Laird, 1983). In order to differentiate mental models from similar concepts used in the field of Human Computer Interaction (HCI), as for example system conceptual models and system images, Norman (1983) referred to mental models specifically as users' mental representations of the system they work with. The ideal case would be, if the system designed (or in our case a website), would correspond with the user's mental model. Users anticipate where web-objects will appear on their screen and thus this knowledge can be used to improve their performance (Vaughan & Dillon, 2006). Moreover, mental models so Norman (1983), facilitate comprehension and lead to an improvement in usability.

In this study we use the term *mental model* in its HCI sense, relating to software / website usability – a mental model is a user's representation of a website. Bernard (2002) claimed that an essential need for the construction of a website is to know the typical mental models of the users for the characteristic *location* of web-objects. This knowledge, so Bernard (2002) aids the website's accessibility and this, in turn, facilitates more accurate and faster information retrieval, as well as greater satisfaction with the websites. Roth, et al. (2010) were able to show that users generally agree on the locations of many web-objects and that these mental models are not influenced by demographic factors like gender and web expertise. In a further study Roth, Tuch, Mekler, Bargas-Avila, & Opwis (2011) investigated the influence of mental models on orientation and their results indicate that people use mental models subconsciously for orientation. The concept of mental models and their influence of users' behavior have been researched in several studies presented in the following paragraph.

Related literature

A study conducted by Zona Research Inc. (2001) reported that participants having difficulties to locate products caused them to give up looking for these products. Therefore Bernard (2002) investigated the user's expectations of the location of common web-objects such as the shopping cart, the login/register button and the help button. He asked participants to place cards representing different web-objects in an empty browser window, where they expected them to be located and found a general consensus among participants on location for most of the objects. Bernard (2002) concludes that an important improvement for online companies, namely to keep potential customers on their website searching for merchandise, would be to place key web-objects, as for example the shopping cart or the help button, where they would be expected to be located by a majority of the users. Subsequently Oulasvirta et al. (2005) examined how expectations of prototypical locations influenced the user's search behavior. This was tested in an eye-tracking study with 15 websites, of which each had three different layouts differing in the location of the navigation (left, right, both sides of the page). The participants' task was to find target links as quickly as possible. Results showed strong left-preference for the links. Therefore the users, following their expectations, first looked to the left side of the presented websites, independently of where the actual target link was presented. After the search task, the (participants') memory of locations of two specific links was tested. Memory task results showed that the links were better remembered if they were placed according to the user's expectations (mental models). Vaughan and Dillon (2006) examined the impact of online newspapers designed according to genre-conforming designs and genre-violating designs on user performance and usability over time. The genre-violating newspaper website differed from the conforming one in several ways, i.e. the navigation bar was on the right side, the background color was different, the headers were smaller, the layout was different and the stories were categorized differently. Performance was measured with the amount of time used to complete information-seeking tasks, and usability was assessed with a self-developed 5-item Likert-scale questionnaire. These questions were designed to capture the subject's general level of satisfaction with the web newspaper by addressing the degree to which they liked it, their level of comfort, whether it was a fun experience and whether it reflected their model of a "web newspaper". Vaughan and

Dillon (2006) found that the group in the genre-conforming design performed faster in the information-seeking tasks than the group in the genre-violating design but they did not find significant between-group differences concerning satisfaction. Another finding of their study was that participants were able to improve their performance in the genre-violating condition over time, but did not reach performances equal to those in the genre-conforming condition. Santa-Maria and Dyson (2008) investigated what happens to user performance and disorientation when visual conventions of a genre are violated. They constructed conventional and conventional-violating web forums to compare user performance over time. They found no significant between-subjects differences. What they did find, however, was a gradual improvement. Both groups were able to improve their task performances over time, and finally the performance between conditions leveled out. Roth et al. (2011) tested whether web-objects that are located according to users' expectations are found faster than uncommonly placed objects using an eye tracking study with three different website types (online shops, online newspaper and company web pages). The results showed that participants needed twice as many fixations when the web-objects were located atypically. On online shop pages, the login area and search field, followed by the shopping cart, seem to be especially sensitive to location typicality. In sum these studies indicate that users seem to have certain expectations where to find different web-objects i.e. mental models which trigger top-down mechanisms.

Aim of the present study

All the studies mentioned before were conducted in laboratory settings and cannot be fully generalized to the real world. With the intention to extend previous findings and to reach a high external validity, we conducted our study in a more natural setting. All the participants were intended to take part in the study using their own computer in a familiar environment. Further shortcomings of previous studies are, that either screenshots of web shops were analyzed: Oulasvirta et al. (2005); Roth et al. (2011) or users' tasks were very simple and furthermore tasks often only consisted in searching web-objects without real interaction: Oulasvirta et al. (2005); Vaughan and Dillon (2006); Santa-Maria and Dyson (2008); Roth et al. (2011). Hence, in order to recreate an as real shopping situation as possible, we intended to program a fully functional online shop containing several interactive web-objects, allowing users a real interaction. As noted before, Bernard (2002) pointed out the importance of the location of web-objects. In contrast to Vaughan and Dillon's (2006) study, where besides the location of web-objects various other factors were manipulated, we examined whether the mere location of chosen web-objects has an influence on user performance, perceived usability and orientation. Hornbæck (2006) noted that there is a lack of studies measuring subjective user experience over time. Further, Santa-Maria and Dyson (2008) stated that time is an important factor to consider because results of the authors study showed that disorientation and user performance change over time. Therefore this study also considers the impact of location typicality on users' performance, perceived orientation and usability over time. Based on the theory above we proposed the following hypotheses:

H1. *Location typicality*. High location typicality of web-objects results in a better user performance and higher ratings on perceived orientation and usability than low location typicality.

H2. *Time x location typicality*. The participants exposed to low location typicality improve significantly more in user performance, perceived orientation and usability ratings over time than the participants exposed to high location typicality.

Method

Design

The experiment had a 2 x 2 mixed design to study the influence of different locations of web-objects on user performance and perceived usability over time. The unrelated independent variable was the *location typicality* of web-objects, with two levels (low vs. high) and the related measures independent variable was *time* with two levels (session 1 vs. session 2). The sessions took place one week apart. The dependent variables were perceived usability and user performance (number of clicks to task completion and task completion time). The experiment was designed as an online experiment.

Participants

The participants were recruited from the database of the Department of Psychology at the University of Basel, which contains the data of people interested in psychological studies. As an incentive, an iPhone 4 was raffled between all participants who participated in the experiment. In total, 153 participants started the experiment of which 38 dropped out during the first session. From the remaining 115 participants, 69 started the second session, but only 58 completed it. In addition, five participants had to be excluded from the study sample: two participants were using Internet Explorer (IE) 7 and three participants completed the study by using mobile devices. IE 7 and mobile devices were not supported by the applied data tracking method. The final sample consisted of 53 participants. This corresponds to a dropout rate of 34.64%.

The participants were randomly assigned to the high or low location typicality condition: 26 to the high location typicality group and 27 to the low location typicality group. There were 15 males (M = 29.7 years, SD = 10.6; range = 21-60) and 38 females (M = 27.8 years, SD = 7.0; range = 19-49). The participants ranked their computer skills, their Internet skills, their online shopping skills, and their website designing expertise on a scale from 1 to 100 (1 = no experience; 100 = expert). The average self-rated computer skill was 65.9 (*SD* = 18.1), the average self-rated Internet skill was 68.79 (*SD* = 19.9), the average self-rated shopping skill was 58.58 (*SD* = 23.7) and the average self-rated website designing experience was 19.74 (*SD* = 3.62). Computer skills, Internet skills, shopping skills and website designing experience were tested for possible differences between the two groups. Independent-samples t-tests were conducted to compare different skill sets in the high location typicality condition and the low location typicality condition. As shown in Table 1, no significant differences between the two conditions were found.

Table 1

Mean values for the skill sets distribution for both conditions compared with an independent t-test

	high (*n* = 26)	low (*n* = 27)		
Location typicality	*M (SD)*	*M (SD)*	*T (51)*	*p*
Computer skills	67.1 (19.2)	64.8 (17.4)	.442	.660
Internet skills	67.2 (21.7)	70.3 (18.3)	-.555	.884
Shopping skills	63.5 (20.4)	53.8 (26.0)	1.509	.137
Website designing expertise	37.5 (29.6)	30.7 (24.4)	-.291	.773

Measurements

User performance was measured (using Javascript events) by the time needed to solve a given task (from loading the page to last click needed for solving the task) and by the number of clicks to task completion. The Web server on which the online shop was hosted used the UNIX time stamp[3] to record the page loading times and the times at which each interaction with web-objects in the online shop of the study were made by participants. Completion time for a given task was thus measured as the period between

[3] http://www.unixtimestamp.com

the server's delivery of the page and a participant's next click on a web-object. This method to measure performance time is precise and browser independent (Sheikh, Wegdam, & Van Sinderen, 2008). Clicks made outside the online shop or on non-clickable objects were ignored.

The subjective satisfaction with the performance at task level was measured with the After-Scenario Questionnaire (ASQ; Lewis, 1991, 1995). The ASQ is a short questionnaire (three 7-point scale items using 1 = strongly disagree; 7 = strongly agree). The items address three important aspects of user satisfaction with system usability: (1) ease of task completion, (2) time to complete a task, and (3) adequacy of support information. The overall ASQ score is the mean value of responses to these three items (Lewis, 2006). Because no help or support information was implemented in the examined online shop, the last ASQ question (3) was not used in this experiment. The internal consistency for the ASQ was excellent (Cronbach α = .939) for the present study.

The overall perceived usability of the online shop was measured with two validated questionnaires: a measuring instrument for the perceived orientation in online shops (WOOS; Yom & Wilhelm, 2004) and the System Usability Scale (SUS; Brooke, 1996). For the WOOS, seven questions about the structure, clarity, findability of products, orientation, and about the meaningful naming were answered by the participants on a five-point Likert scale (1 = strongly disagree to; 5 = strongly agree). The internal consistency for the WOOS was good (Cronbach α =.834). The SUS contains ten questions that are answered on a five-point Likert scale of strength of agreement (1 = strongly disagree; 5 = strongly agree). The questions cover a variety of aspects of system usability such as the need for support, training, and complexity, and thus have a high level of face validity for measuring usability of a system. The overall value of system usability, the SUS score, ranges from 0 to 100. The higher the SUS score, the higher is the perceived usability (Brooke, 1996). The internal consistency for the SUS was acceptable (Cronbach α = .782).

Materials

Online Shop. To examine the effects of location typicality as realistically as possible, a fully functional online-shop for clothing containing navigation, a login and logout function, a shopping cart, and a product display was implemented.

There were two versions of the shop, which differed in the location of the three examined web-objects (login/-out, shopping cart, navigation).

Preliminary study to identify the most typical and the most atypical location of web-objects. The goal of the preliminary study was to develop two versions of an online shop – namely one with high location typicality and one with low location typicality. According to Roth et al. (2010), the login area, the shopping cart, and the search field seem to be especially sensitive to location typicality. In our study, the search field was excluded because browsing tasks were used to examine location typicality effects. Hence, three key web-objects, (1) login/logout, (2) shopping cart and (3) navigation were selected for further examination. The shop version with high location typicality was constructed according to the data on location typicality available from the study of Roth et al. (2010). To determine the most atypical location for each web-object, the shop was divided into four parts as shown in the example Figure 1. Each web-object was then placed in one of the four parts, so that every combination of web-objects was given. This led to a final set of five different web shops (see Appendix A for all locations used in the preliminary study).

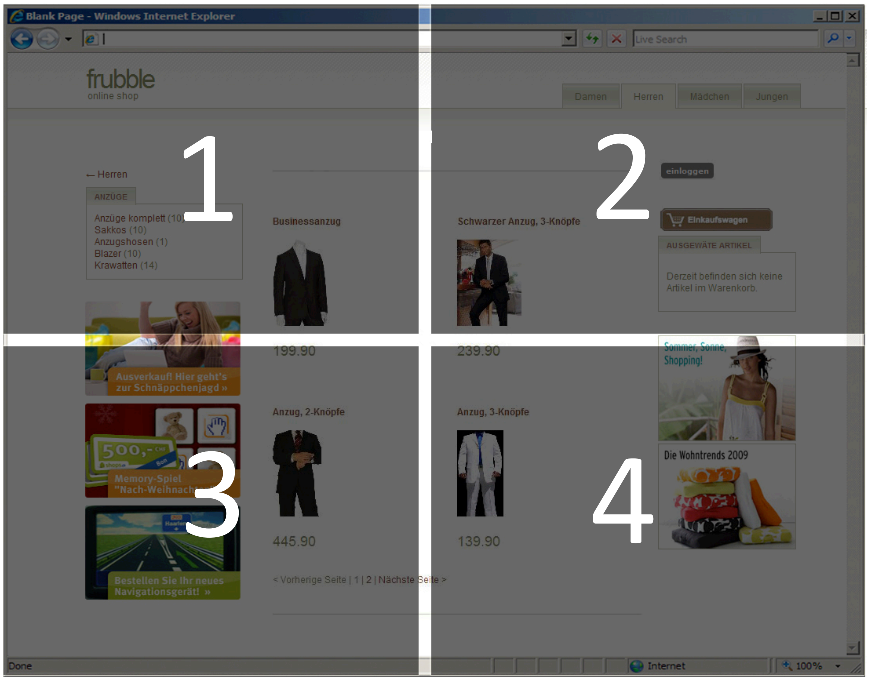


Figure 1. Possible locations for the web-objects in the preliminary study.

The screenshots of the five different shop versions (four with suspected low location typicality and one with suspected high location typicality) were presented in an online survey. 57 participants were asked to rate the location of all three key web-objects in each shop version on a scale from 1 to 100 (1= very atypical; 100 = very typical). The ratings for each location of each web-object are shown in Table 2.

Table 2

Mean ratings for locations of the three key web-objects

Web object	Location	Mean score	*SD*
Login/-out	Top left	49.1	28.7
	Top right	**80.1**[a]	21.9
	Bottom left	**11.3**[b]	18.2
	Bottom right	12.3	16.9
Shopping cart	Top left	43.7	30.0
	Top right	**86.9**[a]	18.6
	Bottom left	49.3	27.9
	Bottom right	**23.1**[b]	22.4
Navigation	Top left	**84.6**[a]	21.5
	Top right	45.0	31.9
	Bottom left	35.3	28.9
	Bottom right	**25.3**[b]	24.3

Note. Bold number: web-object was selected for the main study
[a] Location was chosen for the high location typicality shop
[b] Location was chosen for the low location typicality shop.

For the final study the locations with the highest scores were chosen for the high location typicality shop version and the locations with the lowest scores were chosen for the low location typicality shop version.

Online shop used in the main study. As found in the preliminary study, the three key web-objects were ultimately placed in a combination of a perceived high typical position or in a combination of a perceived atypical position. The screenshots of the final locations of these three web-objects in the two shop conditions can be seen in Figure 2.

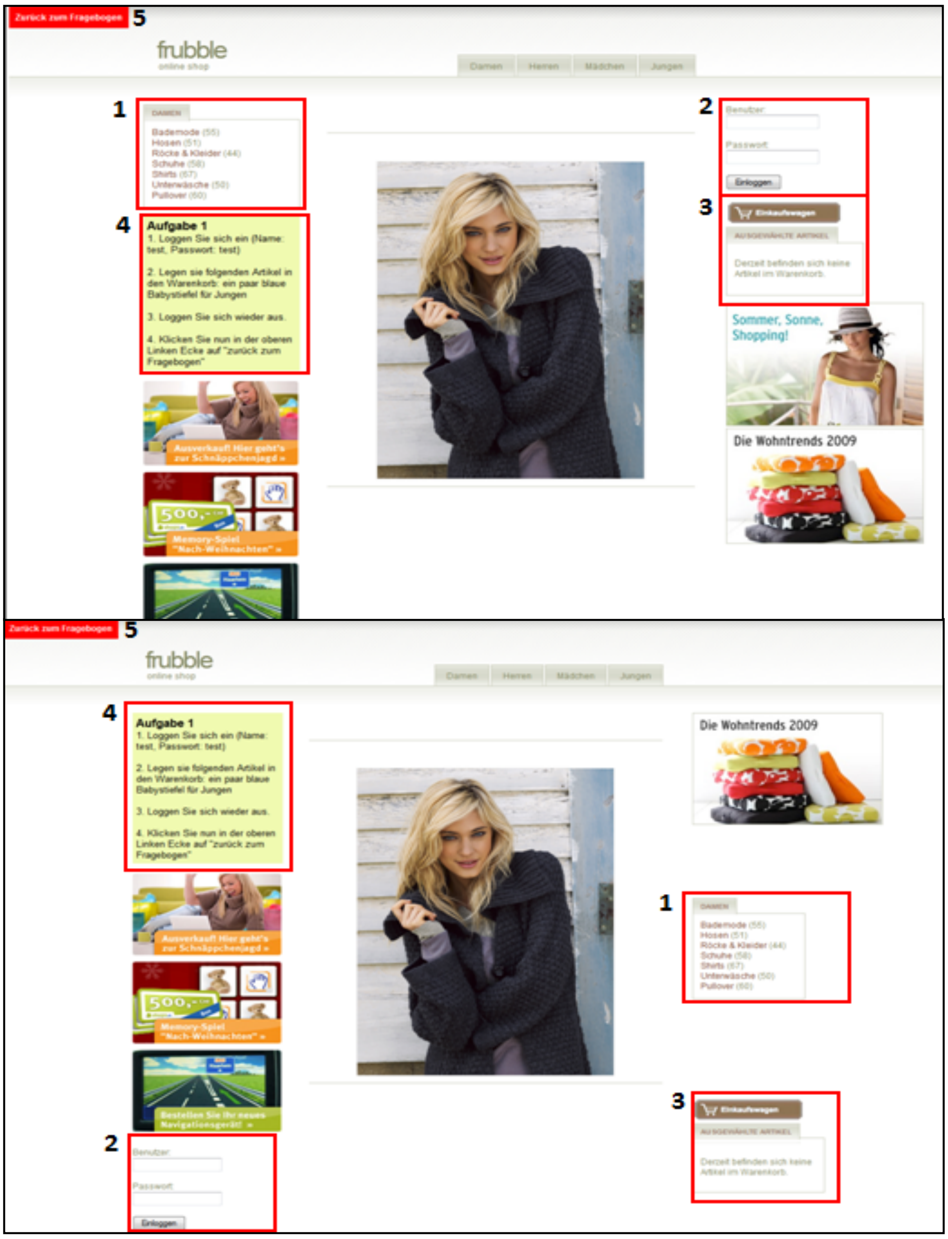


Figure 2. Start pages of the high (top image) and low (bottom image) location typicality shop including the task description. **1** = Navigation; **2** = Login/-out; **3** = Shopping cart; **4** = Task list; **5** = Link to questionnaires.

The three key web-objects, login/out, navigation and shopping cart, were designed to be fully integrated in the general design of the online shop to prevent eventual pop-out effects. In order to achieve a high comparability between the two shops, the web-objects were located at equal distances to the center of the screen.

The shop used in this study is similar to the one used in the study from Tuch, Roth, Hornbæk, Opwis and Bargas-Avila (2011). The shop had four categories at the top-level represented with tabs at the top of the page. Each main category had seven subcategories represented in form of a menu on the side of the page and each subcategory had again five to seven subcategories also represented in form of a menu on the side of the page. By clicking on a category of the lowest level, product items were displayed in the center of the page. Product items were organized in a 2 x 3 matrix, displaying a picture, the name and the price of each item. Clicking on a product item led to a detail page displaying a picture of the product item, the name, the price and a detailed product description. In total, the shop contained more than 1300 different product items. Moreover, a shopping cart listing all added product items and a login/-out function was available.

Tasks. Participants had to solve three tasks in the online shop. Each task consisted of three sub-tasks. For example, participants had to buy a specific product. This task consisted of following three sub-tasks: (a) log in, (b) browse for a specific product item and add it to the shopping cart, and (c) log out. Trying to avoid learning effects, the sessions contained not exactly the same sub-tasks but highly comparable ones. For example, instead of having to browse for a brown cashmere sweater, participants had to browse for a sweater called “Glacier-Blue”. The complete task list for both sessions can be found in Appendix B.

Acquisition of questionnaires. The questionnaires ASQ, SUS, and WOOS were collected online with the survey software Unipark Enterprise Feedback Suite 8.0[4].

Procedure

The online study was conducted in German language from June to July 2011. Participants were randomly assigned to one of the two conditions and then directed to the instruction page. Next, they had to provide demographical data and were then led to the instruction page for the first shopping task. With a click on a link on this instruction page, participants were led to the online shop where the task instructions were presented again on a banner on the left side of the page. After the completion of the first task, participants were asked to rate their task use-experience with the ASQ questionnaire. In the same manner, two more shopping tasks followed. After completing all three shopping tasks, participants were asked to evaluate their overall shopping experience (across all three tasks) on the SUS questionnaire. Also, they had to indicate on the WOOS questionnaire how oriented they felt.

One week after the first session, participants were reinvited via E-mail to the second session of the study. The procedure was the same as in session one. For an overview of the experimental procedure and the assessed measurements, see Figure 3.

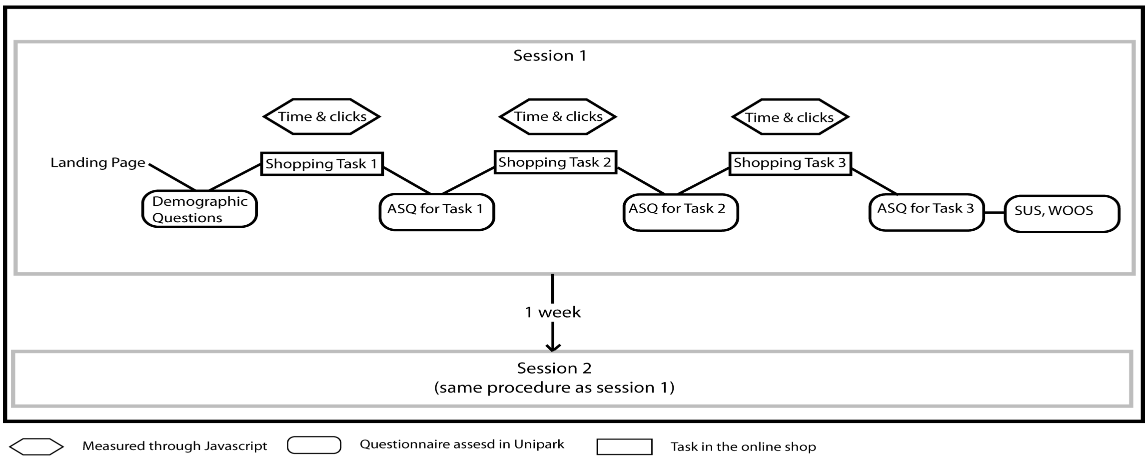


Figure 3. Experimental procedure.

[4]http://www.unipark.de

Results

The data were analyzed using SPSS 19 for Windows. All analyses are performed at the significance level of .05. All data were checked for normal distribution and whether they met the required conditions for the statistical tests. Time values were log-transformed to achieve normal distribution. Furthermore, the data were handled according to the approach of Tabachnick and Fidell (2006). Missing data for *time* and *clicks* were estimated through group mean substitution. Univariate outliners and missing data were analyzed and excluded. This exclusion is the reason for differing sample sizes within the statistical values.

Effects of location typicality on perceived usability and orientation

Independent samples t-tests were conducted to compare perceived usability in high and low location typicality conditions first. Table 3 provides a summary of these tests.

Table 3

Independent t-tests for perceived usability and user's performance

Location typicality	high *M (SD)*	low *M (SD)*	*t(df)*	*p*	*Effect size d*
Perceived Usability					
SUS	85.4 (6.7)	80.2 (9.9)	2.12 (46)	.020*	.614
WOOS[2]	4.3 (.57)	4.0 (.47)	1.74 (48)	.042	.469
User's performance					
Completion time(sec)[1]	243.7 (80.1)	253.7 (75.8)	-.45 (50)	.326	.128
Number of clicks[2]	51.3 (9.8)	57.0(8.5)	-2.05 (43)	.021*	.621

Note. *p* = one tailed; *medium effect; [1]The displayed values are not log-transformed; statistical tests are based on the log-transformed data; [2] Data not normal distributed, for better comparability t-test results are shown, Mann-Whitney U tests showed the same results.

As expected, the online shop usability was rated higher on the SUS in the high location typicality condition than in the low location typicality condition ($t(46) = 2.12$, $p = .020$) and the perceived orientation (WOOS) was significantly different for the two conditions ($t(48) = 1.74$, $p = .042$). Results show that ratings of orientation in the online shop were significantly higher among those participants in the high location typicality condition than among those in the low location typicality condition. Supporting our hypothesis, these results indicate that location typicality has an effect on perceived usability and orientation in online shops. Specifically, results suggest that when an online shop has its web-objects placed according to the expectations of a potential customer, said customer will be more satisfied with the usability of this online shop and will feel better oriented.

Effects of location typicality on user's performance

Independent samples t-tests were conducted to compare user's performance in high and low location typicality conditions. As shown in Table 3, there was a significant difference between the two conditions concerning the amount of clicks used to execute all three shopping tasks, ($t(43) = -2.05$, $p = .021$). Results confirmed our hypothesis and show that participants in the high location typicality condition needed significantly fewer clicks to execute the shopping tasks than those in the low location typicality condition. Contrary to our expectations, there was no significant difference between the two shopping conditions for the total amount of time used to execute all three shopping tasks ($t(50) = -.45$, $p = .326$).

Effects of location typicality on user's experience over time

In order to examine changes in usability depending on location typicality, a set of 2x2 repeated measures ANOVA with *session (1 vs. 2)* as within-subject factor and *location typicality* (low vs. high) as between-subject factor was calculated. An overview of results found over time is presented in Table 4. All descriptive data can be found in Appendix C.

Table 4

ANOVA for perceived usability, orientation and users' performance over time

	M (SD)	*M (SD)*	*F*	η_p^2	*p*
Location typicality	**high**	**low**			
Perceived orientation (WOOS)	4.27 (.62)	4.03 (.54)	(1,48) = 2.60	.051	.113
Perceived usability (SUS)	85.41 (7.66)	80.20 (10.86)	(1,46) = 4.48	.089*	.040
Completion time[1]	243.75 (84.50)	249.03 (86.12)	(1,50) = .20	.004	.653
Number of clicks[1]	51.31 (10.18)	57.06 (9.74)	(1,43) = .4.42	.093*	.041
Session (time)	**1**	**2**			
Perceived orientation (WOOS)	4.12 (.60)	4.17 (.56)	(1,48) = .63	.013	.431
Perceived usability (SUS)	82.02 (8.32)	83.54 (10.21)	(1,46) = 1.51	.032	.225
Completion time[1]	271.88 (91.96)	220.91 (78.66)	(1,50) = 21.40	.300**	.000
Number of clicks[1]	53.79 (10.04)	54.57 (9.58)	(1,43) = .487	.011	.489
Location typicality x time					
Perceived orientation (WOOS)	-	-	(1,48) = 1.75	.035	.254
Perceived usability (SUS)	-	-	(1,46) = 1.10	.024	.298
Completion time[1]	-	-	(1,50) = .90	.018	.347
Number of clicks[1]	-	-	(1,43) = .00	.000	.972

Note. * medium effect; ** large effect ; [1]the displayed values are not log-transformed; statistical tests are based on the log-transformed data.

Perceived usability and orientation over time. A repeated measures ANOVA with *session* (1 vs. 2) as within-subject factor and *location typicality* (high vs. low) as between-subject factor was performed to test whether the location of web-objects had an influence on perceived usability and orientation over time (see Figure 4 for an overview). Against our predictions, results indicate that there is no significant *location typicality x time* interaction for the SUS scores (F(1, 46) = 1.109, p = .298, η_p^2 = .024) nor for the WOOS scores (F(1, 48) = .1.751, p = .254, η_p^2 = .035) and there were no significant main effects for time for *perceived usability(SUS)* (F(1, 46) = 1.510, p = .225,

η_p^2 = .032) nor for time for *perceived orientation (WOOS)* (F(1, 48) = .630, p = .431, η_p^2 = .013).

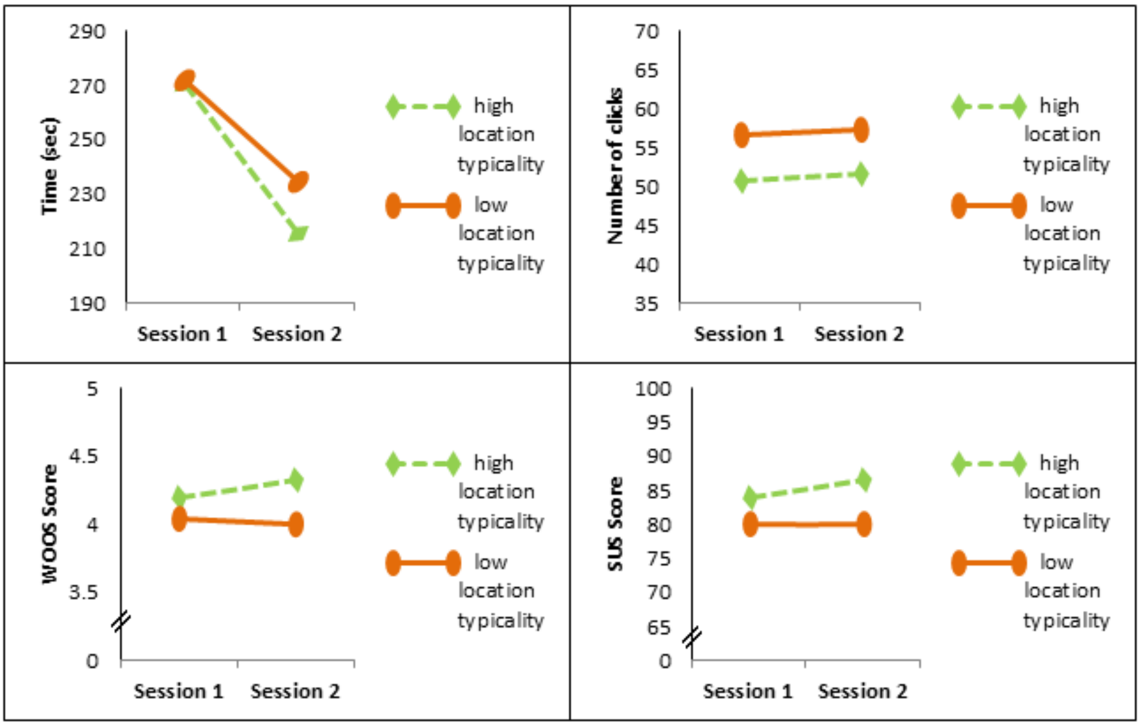


Figure 4. Scores for perceived usability (SUS), perceived orientation (WOOS), and users' performance (clicks and time) in Session 1 and Session 2.

Hence, according to our expectations, the ratings in the high location typicality condition for perceived usability and orientation remained stable over time, whereas contrary to our expectations, there was no improvement in the ratings for the low location typicality condition. Consequently, in both conditions, participants' perceived usability and orientation remained stable over time.

User performance over time. A repeated measures ANOVA with *session* (1 vs. 2) as within-subject factor and *location typicality* (high vs. low) as between-subject factor was performed to test whether the location of web-objects had an influence on users' performance over time. Against our predictions, results indicate that there is no significant *location typicality x time* interaction for the *completion time* (F(1, 50) = .902, p = .347, η_p^2 = .018) nor for *number of clicks* (F(1, 43) = .001, p = .972, η_p^2 = .000) and there was no significant main effect for the amount of clicks used to execute all three shopping tasks ($F(1, 43) = .487$, $p = .489$, $\eta_p^2 = .011$). However, a significant main effect was found for *completion time* used to execute the three shopping tasks $F(1, 50) = 21.40$, $p = .001$, $\eta_p^2 = .300$). As expected, the participants completed the shopping tasks in session 2 significantly faster than in session 1. These results suggest, that *time,* independent from location typicality, only affects users' performance: participants significantly improved in time to execute the shopping tasks from session 1 to session 2 (see Figure 4).

Effects at task level over time. To examine adaption to location typicality in more detail, i.e. on task level, multiple 2x2 repeated measures ANOVA with *tasks* (tasks in session 1 vs. tasks in session 2) as within-subject factor and *location typicality* (low vs. high) as between-subject factor were calculated. Dependent variables were the ASQ questionnaire, the amount of time and clicks to task completion. The same pattern of results (found on global level) for perceived usability, perceived orientation and users' performance was observed on task level: user performance over time improved for the time to complete the tasks while no change in perceived usability and orientation occurred over time (see Appendix D for the analysis on task level).

Discussion

The present study could show that the location typicality of web-objects has a statistically significant influence on perceived orientation, usability and user performance. Our results showed that web-objects located according to users' expectations, i.e. corresponding with mental models, result in higher perceived orientation and higher usability ratings than web-objects located at unexpected locations. For the user performance measurements, we found a between-group difference for the amount of clicks to execute the shopping tasks. Participants shopping in the high location typicality condition were able to perform the set of tasks with a significantly lower amount of clicks than the participants shopping in the low location typicality condition. However, we did not find an effect of location typicality of web-objects on the amount of time to execute the shopping tasks. Participants performed the tasks in the low location typicality condition in a comparable amount of time to the participants performing the tasks in the high location typicality condition.

Further, the results for the influence of time, i.e. repeated interaction, on users' performance, perceived orientation and usability suggest, that users only significantly improved in the time to execute the shopping tasks. Additionally, in our study, this effect is independent from the location typicality, meaning that participants in the low typicality condition improved at an equal level than the participants in the high location typicality condition over time. And even though participants were able to execute the shopping tasks faster in session 2, this had no impact on their perceived orientation and usability. Results showed no improvement over time for the perceived orientation and usability ratings, both remained stable over time.

Additionally, we were able to validate the mental models found by Roth et al. (2010) for online shops in a preliminary online study. The users asked in the preliminary study generally agreed about the typical locations for the login/-out (top right), the shopping cart (top right) and the navigation (top left). Additionally, we investigated the most atypical or uncommon locations for these three web-objects. And as well, users also expressed a general consent for the login/-out (bottom right), the shopping cart (bottom right) and the navigation (bottom right). Similar to the finding of Bernard (2002), mental models for online shops, therefore seem to be quite common and similar among users.

The findings of this study indicate that the mere location of web-objects seems to have an influence on perceived usability and orientation. The small manipulation of only the location of web-objects caused a change in participants' perception of the website. Users seem to feel better oriented and are more satisfied with the usability if web-objects are located at expected positions according to their mental models. If web-objects are not located at these expected locations, despite users are able to adapt their task execution behavior and execute the tasks faster after repeated interaction, their perceived orientation and usability does not improve over time. Therefore it is important to consider users' mental models in the decision process where to place web -objects in online shops in order to achieve a high satisfaction and orientation.

Our finding that violating users' mental models can leave users disoriented and influence perceived usability supports Normans' (1983) theory that designing systems according to users' mental models facilitates comprehension and leads to an improvement in usability, and replicates the findings of the studies from Santa-Maria and Dyson (2008), and Vaughan and Dillon (2006). But contrary to our expectations and past studies Oulasvirta et al. (2005); Vaughan and Dillon (2006); Roth et al. (2011), we found no significant differences for the total amount of time used to execute the shopping tasks. A possible explanation for this different finding is, that contrary to our study, the studies mentioned: (a) either manipulated, besides the location of web-objects, several more factors[5], (b) used only screenshots of websites to analyze user performance[6], (c) used tasks that consisted only of searching tasks and therefore did not include real interaction but just the mere tracking down of a web-object[7], or (d) were conducted in laboratory settings and therefore cannot be fully generalized to the real world[7].

[5] Vaughan and Dillon (2006); Santa-Maria and Dyson (2008)
[6] Oulasvirta et al. (2005) ; Roth et al. (2011)
[7] Oulasvirta et al. (2005) ; Vaughan and Dillon (2006) ; Santa-Maria and Dyson (2008); Roth et al. (2011)

However, our finding that time has an influence on user performance but does not affect perceived orientation and usability, is similar to the finding of Vaughan and Dillon (2006). According to Vaughan and Dillon (2006), it is a quite common phenomenon that satisfaction data fails to parallel behavioral data in the evaluation of software systems. In our case, this could be explained through a ceiling effect. Tasks were rather simple, so that they could be completed without great effort in both sessions. This, in turn, led to an already high performance satisfaction in session 1.

This findings practical implication applies especially for web designers. They have to keep in mind that although placing web-objects corresponding with users' mental models may be not as creative and innovative as violating conventions, but users seem to feel better oriented, are more satisfied with the usability and need fewer clicks to accomplish shopping tasks in conventional online shops. This, so one can speculate, could have an influence on users shopping behavior and therefore impact sales of an online shop.

There are some caveats to be aware of for our findings. First, a real online shop could be more complex. Not many distractor elements or ads were implemented in this study's online shop, so that the three web-objects were probably easily located although they were placed at unexpected locations. Tasks may have been done using bottom-up processes. Hence, the participants in the high location typicality condition had no measurable advantage using top-down processes to distinguish the web-objects. Second, we did not consider memory effects. For technical reasons, our tasks could not be presented in a randomized order. Therefore results found should be interpreted regarding this problem of eventual sequence effects. Finally, the lack of pressure to succeed and the fact that it was an online study, where the setting was not as strictly controlled as it would have been in a laboratory study, could have tempted the participants to explore the shop more than intended. This in turn could have influenced users' performance measurements.

Conclusion and future research

In this study we were able to show that the mere location of web-objects in online shops seems to influence users shopping experience in several ways. Atypical locations of web-objects lead to more disorientation, less perceived usability and result in a higher amount of clicks to accomplish shopping tasks than typical locations of web-objects. Viewed over time, users improve their performance independently from the location of web-objects. Further investigation is needed to determine whether people prefer conventional websites or whether they prefer to be surprised by unexpected designs. In relation with this, future research should also investigate carefully how perceived orientation, usability and user performance impact sales in order to consider the advantages and disadvantages of extraordinary locations for web-objects.

References

Baldassi, S., & Burr, D. C. (2004). “Pop-out” of targets modulated in luminance or colour: the effect of intrinsic and extrinsic uncertainty. *Vision research, 44*(12), 1227-1233.

Bernard, M. (2002). Examining user expectations for the location of common e-commerce web objects. *Usability News, 4*(1). Retrieved from http://www.gieger.com/site/research/webobject_placement.pdf

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189-194.

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: bottom-up versus top-down. *Current Biology, 14*(19), 850-852.

Dillon, A., & Vaughan, M. (1997). ‘It's the journey and the destination’: shape and the emergent property of genre in evaluating digital documents. *New Review of Hypermedia and Multimedia, 3*(1), 91-106.

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International journal of human-computer studies, 64*(2), 79-102.

Hui, B., Partridge, G., & Boutilier, C. (2009). *A probabilistic mental model for estimating disruption*. In: *Proceedings of the 14th international conference on intelligent user interfaces*. (pp. 288-296), Sanibel Island, Florida, FL: ACM

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*: Harvard Univ Press.

Kim, S. J., Dey, A. K., Lee, J., & Forlizzi, J. (2011). *Usability of car dashboard displays for elder drivers*. In: *Proceedings of the 2011 annual conference on Human factors in computing systems*, Vancouver, Canada, CA: ACM.

Lewis, J. R. (1991). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *SIGCHI Bulletin, 23*(1), 78-81.

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction, 7*(1), 57-78.

Lewis, J. R. (2006). Usability Testing. *Handbook of Human Factors and Ergonomics* (pp. 1275-1316): John Wiley & Sons, Inc.

Norman, D. (1983). Some Observations on MentaI ModeIs. *Mental models.* (pp. 7-14). Hillsdale, New Jersey: Lawrence Erlbaum Assicuates.

Otter, M., & Johnson, H. (2000). Lost in hyperspace: metrics and mental models. *Interacting with Computers, 13*(1), 1-40.

Oulasvirta, A., Kärkkäinen, L., & Laarni, J. (2005). Expectations and memory in link search. *Computers in Human Behavior, 21*(5), 773-789.

Roth, S., Tuch, A., Mekler, E., Bargas-Avila, J., & Opwis, K. (2011). You Never Get a Second Chance to Make a First Impression: Meet Your Users' Expectations Regarding Web Object Placement in Online Shops. In Pantano, E., & Timmermans, H. (Eds.), *Advanced Technologies Management for Retailing: Frameworks and Cases.* (pp. 221-235).

Roth, S., Schmutz, P., Pauwels, S., Bargas-Avila, J., & Opwis, K. (2010). Mental models for web objects: Where do users expect to find the most frequent objects in online shops, news portals, and company web pages? *Interacting with Computers, 22*(2), 140-152.

Santa-Maria, L., & Dyson, M. C. (2008). *The effect of violating visual conventions of a website on user performance and disorientation: how bad can it be?* In: Proceedings of the 26th annual ACM international conference on Design of communication, (pp. 47-54) Lisbon, Portugal, P: ACM.

Sheikh, K., Wegdam, M., & Van Sinderen, M. (2008). Quality-of-context and its use for protecting privacy in context aware systems. *Journal of Software, 3*(3), 83-93.

Sobel, K. V., Gerrie, M. P., Poole, B. J., & Kane, M. J. (2007). Individual differences in working memory capacity and visual search: The roles of top-down and bottom-up processing. *Psychonomic bulletin & review, 14*(5), 840-845.

Tabachnick, B., & Fidell, L. (2006). Cleaning up your act. *Screening data prior to analysis Using multivariate statistics* (pp. 56-110): Boston, MA: Allyn and Bacon.

Thatcher, A., & Greyling, M. (1998). Mental models of the Internet. *International Journal of Industrial Ergonomics, 22*(4-5), 299-305.

Tuch, A. N., Roth, S. P., Hornbaek, K., Opwis, K., Bargas-Avila, J. (2011). *Working Towards Understanding the Relation Between Usability and Aesthetics in HCI: Mediation by Affect.* Manuscript submitted for publication.

Unipark EFS Survey (Version 8.0) [Online-survey tool]. Cologne-Huerth, Germany. Retrieved from http://www.unipark.info/

Vaughan, M. W., & Dillon, A. (2006). Why structure and genre matter for users of digital information: A longitudinal experiment with readers of a web-based newspaper. *International journal of human-computer studies, 64*(6), 502-526.

Vila, N., & Kuster, I. (2011). Consumer feelings and behaviours towards well designed websites. *Information Management, 48*(4-5), 166-177.

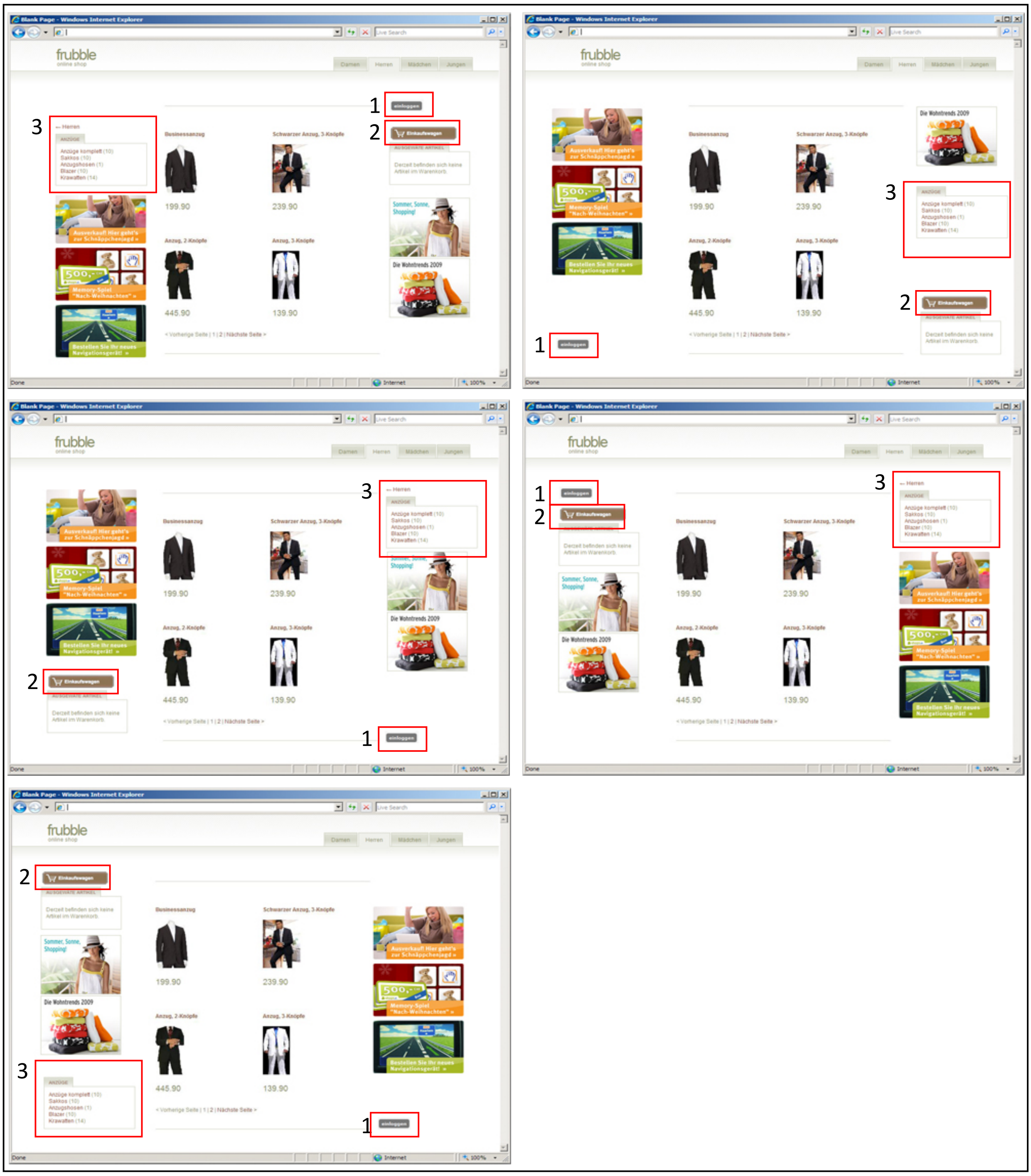
Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Attention, Perception, & Psychophysics, 63*(3), 381-389.

Yom, M., & Wilhelm, T. (2004). WOOS – Ein Messinstrument für die wahrgenommene Orientierung in Online Shops. [WOOS – A measuring instrument for the perceived orientation in online shops]. In R. Keil-Slawik, H. Selke & G. Szwillus (Eds.), *Mensch & Computer 2004: Allgegenwärtige Interaktion* (pp. 43-53). München: Oldenbourg Verlag.

Zona Research. (2001). The need for speed II. *Zona Market Bulletin*, 5, April 2001. Retrieved from http://www.keynote.com/downloads/Zona_Need_For_Speed.pdf

Appendix A

Final set of five different web shops for the preliminary study



Note. 1 = Login; 2 = Shopping cart; 3 = Navigation.

Appendix B

Tasks in both sessions

Session 1

1. a) Log in (Name: test, Password: test).
   b) Please look for a pair of blue baby boots for boys and add them to the shopping cart.
   c) Log out.

2. a) Log in again (Name: test, Password: test).
   b) Please look for a brown cashmere sweater for women and add it to the shopping cart.
   c) Check the total sum of all your purchases (viewable in the shopping cart) and remember it.

3. a) Please look for a pair of jeans called „Tourmaster" for men add them to the shopping cart.
   b) Remove the brown cashmere sweater for women from the shopping cart.
   c) Log out.

Session 2

1. a) Log in (Name: test, Password: test).
   b) Please look for a pair of brown boots for boys and add them to the shopping cart.
   c) Log out.

2.
   a) Log in again (Name: test, Password: test).
   b) Please look for the sweater called "Glacier Blue" for women and add it to the shopping cart.
   c) Check the total sum of all your purchases (viewable in the shopping cart) and remember it.

3. a) Please look for a brown pair of jeans called „Straight-Leg-Jeans" for men and add them to the shopping cart.
   b) Remove the "Glacier-Blue" sweater for women from the shopping cart.
   c) Log out.

Appendix C

Descriptive statistics

Descriptive statistics for low location typicality and high location typicality condition in session 1 and 2

Location typicality	High *M(SD)*		Low *M (SD)*	
Session	1	2	1	2
Perceived Usability				
SUS	84.06 (6.66)	86.77 (8.67)	80.10 (9.98)	80.31 (11.75)
WOOS	4.20 (.68)	4.34 (.56)	4.05 (.53)	4.01 (.56)
User's performance				
Time in seconds	271.32 (93.97)	216.19 (75.04)	272.44 (89.96)	225.63 (82.28)
Clicks	50.90 (10.64)	51.72 (9.72)	56.69 (9.44)	57.43 (10.05)

Note. * medium effect; the displayed values are not log-transformed; statistical tests are based on the log-transformed data.

Appendix D

ANOVA results on task level

ANOVA for performance measurements in tasks in both sessions

	M (*SD*) session 1		*M* (*SD*) session 2		*F*	η_p^2	*p*
Location Typicality	**High**	**Low**	**High**	**Low**			
Task 1							
Time in seconds	75.6 (39.2)[1]	74.1 (30.1)[1]	59.1 (24.8)[1]	72.3 (42.3)[1]	(1,50) = 4.543	.083*	.038
Clicks	17.6 (3.5)[2]	20.5 (4.7)[3]	18.1 (4.6)[2]	21.7 (6.6)[3]	(1,43) = .911	.021	.345
Task 2							
Time in seconds	110.2 (35.6)[1]	112.3 (43.3)[1]	80.1 (28.1)[1]	88.6 (38.9)[1]	(1,50) = 26.712	.348**	<.001
Clicks	16.5 (4.4)[2]	18.3 (4.1)[3]	16.9 (4.1)[2]	17.6 (3.1)[3]	(1,43) = .065	.002	.800
Task 3							
Time in seconds	85.4 (44.0)[1]	85.9 (35.7)[1]	76.9 (38.5)[1]	74.1 (29.0)[1]	(1,50) = 7.866	.136**	.007
Clicks	16.7 (4.4)[2]	17.8 (2.7)[3]	16.6 (3.8)[2]	18.0 (2.9)[3]	(1,43) = .023	.001	.880

Note. *medium effect; ** large effect; [1]n = 26; [2]n = 22; [3]n = 23

Wilcoxon's Matched-Pairs Test for perceived usability in tasks in both sessions

	n	*M* (*SD*) session 1	*M* (*SD*) session 2	Mean rank negative	Mean rank positive	Tied ranks	*z*	*p*
Task 1								
ASQ high[a]	22	5.5 (1.5)	6.0 (1.3)	7.6	8.2	7	-1.25	.104
ASQ low[b]	25	5.8 (1.05)	5.9 (1.1)	8.8	9.1	8	-.263	.396
Task 2								
ASQ high[a]	22	5.8 (1.1)	6.29 (.82)	5.1	7.1	10	-1.462	.073
ASQ low[b]	25	6.2 (.94)	6.12 (.93)	10.3	6.6	8	-.757	.224
Task 3								
ASQ high[a]	22	6.1 (1.0)	6.2 (.70)	5.6	7.3	10	-.393	.345
ASQ low[b]	25	6.3 (.90)	6.3 (.81)	6.5	7.7	5	-.495	.311

Note. *p* = one tailed; [a]results for the high location typicality condition; [b]results for the low location typicality condition; [1]the displayed values are not log-transformed; statistical tests are based on the log-transformed data