Visual Search on E-commerce Category Navigation: A Multi-language Study

Bin Liu, Lijuan Cao, Yumei Tang, and Ming Zhong

Alibaba.com & AliExpress.com, P.R. China
rock.liub@alibaba-inc.com

Abstract. E-commerce websites need effective category navigation especially when they need to be accessible to global users with multi-language backgrounds. An eye tracking laboratory study with English, Chinese, and Russian participants was conducted to investigate search behavior that influenced by different category level and layout design (vertical versus horizontal). The results showed that users employed different search strategies at different levels of categories, and they showed significant differences in search patterns on different layouts. The results also indicated the impact of language on search behavior. These findings provided useful considerations for visual research in category as well as implications for the category design of multi-language e-commerce websites.

Keywords: E-commerce, Category, Multi-language, Eye movements, Visual search.

1 Introduction

The global B2C market has been growing rapidly over the past 10 years. According to a report by Ystats.com, a research company for international market, global business to customer (B2C) e-commerce grows at 20% annually [28]. Emerging markets like China and Russia are leading the growth. In 2012, B2C e-Commerce in Russia has grown by over a quarter. Meanwhile in the third quarter 2013, the Chinese B2C e-commerce market reached 26.66 billion U.S. dollars [27].

However, People using e-commerce websites are frequently faced with the problem of finding products on a category due to inefficient category navigation systems. The major challenge is how to design screen layouts that people can search efficiently. Therefore it is important to understand how users search in the category navigation in order to improve the efficiency and effectiveness of e-commerce websites.

There are plenty of researches on visual search from the perspectives of cognitive psychology and eye movement, and areas cover widely such as search array, perceptual span, eye movement control, and eye movement models. However, visual search behaviors vary a lot depending on the search environment or search array [4,17]. Many of the previous studies were conducted in simulation experiment, or focused on evaluation of designs. Although several studies had shed lights on the visual search on

P.L.P. Rau (Ed.): CCD/HCII 2014, LNCS 8528, pp. 470–481, 2014.

menu [1,14,18], limited researches involve category, a special kind of search array. And there are few researches on multi-language category design. Therefore, visual search is worth further investigation on category in e-commerce website with multi-languages.

This study investigated user's information search behavior when navigating by category items on e-commerce websites within the context of three languages: English, Chinese, and Russian. We started out using eye tracking data to generate search pattern, to gain insight into users' visual search strategy.

2 Related Work

There has been an enormous amount of researches done on visual search, which is one of the major cognitive processing tasks (i.e. reading, visual search, and scene view) that has gained considerable attention over the past decades [4]. Researchers started to use eye tracking technology increasingly in recent years, since the eye movements have been considered as a very important way in understanding cognitive processing in visual search.

The most important factor about visual search is significantly influenced by the characteristics of the search array such as density. Complicated and cluttered array gives rise to more fixations and longer durations [4]. The structure and configuration of the search array can influence the target search efficiency. Other detailed factors such as color, shape, information type (target, distractor) also have effect on fixation position and duration [25]. Apart from array, several studies have tried to investigate the factors associated in eye movement control in visual search [8]. Various visual search models were proposed [7]. The models such as simple search model [15] and guided search model [5] discuss visual search patterns from cognitive and perceptual aspects. Hornof found that people use both random and systematic search strategies [11]. However, these researches were conducted in different visual environment, and there is no universal visual search pattern proposed so far.

2.1 Menu and Category Search

As a type of menu navigation, the category search study was originated from the research on menu, which is regarded as one of the most important visual components and has received increasing attention from researchers. The study of visual search on computer-based menu selection started from forty years ago [3]. Previous theories of menu search focused on the cognitive strategies. Unlike the eye movement pattern in reading narrative text, Aaltonen (1998) found that users read menus in sequential sweeps [1].

Plenty of researches focused on the menus style efficiency. The comparison between vertical menu and dynamic menu was conducted to investigate internal pattern [24]. The efficiency and accuracy study of vertical menus found that users needed less eye fixations, and can complete a search task faster in a vertical menu than in a horizontal menu [2,4]. The comparison of select time in array of digit numbers also

indicates that users were quicker in vertical menus than in horizontal menus [26]. Other study about menu details like grouping navigation and group labels showed grouping allows users to access a large number of items with less mouse movement over a few levels [19], and group labels help visual processing [10]. Another eye tracking study found group labels affect users in determining menu reading scan paths [11].

There were other related researches about the manipulate factors which would likely have effect in designing a more reliable visual search experiment [9]. Design elements, including menu item content, the menu items order, the number of items, the items size and spacing, and menu item grouping can influence the menu efficiency. Users adopt diverse mechanisms in interacting with the online environment, and screen usage experience depends largely on the specific task (experiential, goal-directed) intended [12,13]. In addition, the practice effect was also found in target search menu experiment, the more users practice, the quicker they complete the search [15].

In recent years, the research interest on menu has shifted from traditional computer interface to web environment. And the eye tracking research has gradually become popular in investigating how users search on web pages [12]. However, among all the above, limited research involved category menu.

2.2 Linguistic and Semantic Influences on Visual Search

Over the past decades, people are increasingly aware of the various linguistic factors that influence eye fixations. Eye movement patterns like fixation duration and saccade length are directly linked to the difficulties of the text being read. The fixation duration on that word is also determined by the frequency, predictability, meanings, familiarity and semantic relations of the fixated word [17].

Furthermore, there are natural linguistic differences in people's search behavior online. The direction of the language, for example, is shown to influence how people scan and search information [16]. According to Dong and Salvendy (1999), for Chinese menu, Chinese users perform better on vertical layout than horizontal layout. However, when it comes to English menu, the same group of users responded faster when presented with a horizontal layout [6]. Mainland Chinese appear to flexibly adapt their search pattern to the layout of the screen. They used predominantly horizontal searches for row layouts and used vertical searches in column layouts [16]. But that is not applicable in all languages.

3 Hypothesis

Based on previous studies, we expected to find that users' category search behaviors are influenced by their language backgrounds and we held the following hypotheses: H1, Fixation durations are different at different category areas (areas refer to different category levels, target and non-target area for the search task). H2, The visual search

strategies employed are influenced by category design. H2a, Users will show different saccade length at different levels of category. H2b, Users will show different saccade direction on vertical and horizontal layout category.

4 Method

In order to study the effects of language and navigation design on user visual search, a 3×2 design was used with the factors category (vertical, horizontal) and language (Chinese, Russian, English). The goal-directed search tasks were designed to uncover eye movement. The dependent variables of this study were eye movement control operationalized by fixation duration and saccade in different area of category menu during task execution, and other dependent variables were visual search strategy measured by calculating the saccade location in particular areas.

4.1 Participants

A total of 31 participants, including 10 native Chinese speakers (5 female, 5 male, M=21.4 years, SD=1.7, range: 19-24), 10 native English speakers (4 female, 6 male, M=20.2 years, range: 19-21), and 11 native Russian speakers (3 female, 8 male, M=20.7 years, range: 19-22) were recruited from Zhejiang university. They had normal or corrected-to-normal vision, and previously had at least one year experience of using e-commerce websites in their own countries. Participants were compensated for their participation.

4.2 Materials

The stimuli for this experiment were the HTML category prototypes to imitate the category menu in real e-commerce environment. A total of 379 category items were presented in the prototype. They were organized by eight top-level categories. Each of them was furthermore divided into four groups, and for each group there were eight to twelve secondary-level category items. Three language versions, that is, English, Chinese and Russian prototypes were created. All items were originally in English with reference to several top e-commerce websites (AliExpress.com, eBay.com, and Amazon.com). The Russian and Chinese versions were translated from the English version without changing its structure or content. The wordings were based on the items used in local e-commerce websites (OZON.ru and Taobao.com), and also double checked by native speakers. The average width of category items of Chinese were ranged from 12 to 96 pixels (M=39.29, SD=14.29), English were ranged from 14 to 171 pixels (M=76.94, SD=35.16), and Russian were ranged from 22 to 191 pixels (M=94.51, SD=40.55).

For each language, we created two designs (horizontal layout and vertical layout) (see Fig. 1). To eliminate the order and memory effects, a program was written to create HTML prototype webpages with randomized item position within the category hierarchy. 16 (tasks per layout) ×2 (layout) ×3 (language) factorial prototypes were

generated so that each experimental task was conducted on an exclusive prototypes. The web style of prototypes was based on AliExpress website with a consideration of the menu design guidelines proposed by Mayhew (1992) [21]. Both design layouts shared the same elements style, such as font, line-height, and the width of secondary level category menu. The category items are evenly spaced out. The space between each secondary item in the horizontal layout is 20 pixels.

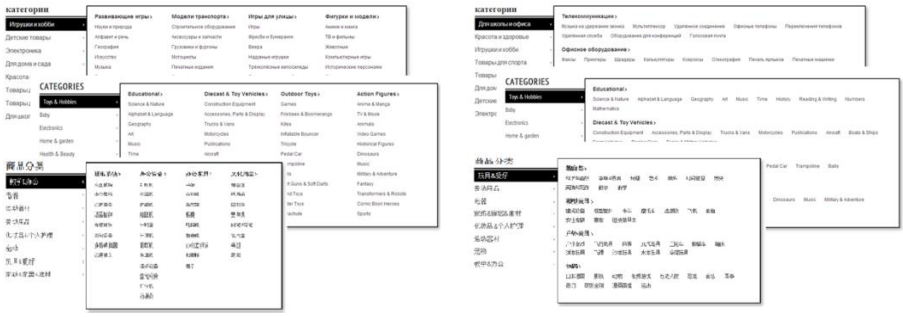


Fig. 1. Samples of horizontal and vertical category designs

4.3 Procedure

After a short introduction about the experiment procedure, participants were briefed with a list of items (such as, basketball, digital camera) they were going to search to prevent misunderstandings caused by semantic reasons. The instructions and prototypes were also in the subjects' native language. They were then asked to perform 32 search tasks with a short break in the middle, during which they need to find and click the best semantically matched target category on the prototype.

4.4 Data Collection and Reduction

Eye-tracking data were recorded by a Tobii T60 eye-tracker with screen resolution of 1280×1024, and were processed with Tobii Studio software. Raw data were exported and processed using MATLAB programs to extract user performance and behavior variables.

In our analysis, the category design was divided into the following areas (see Fig. 2): top-level (A), group titles of secondary-level category (B), non-target secondary-level category (C), target secondary-level category (D), and search target item (E).

All data were aggregated and averaged over all tasks on both the horizontal prototype and the vertical prototype conducted by an individual user. The data we collected focused on the fixation duration, saccade length, and saccade location. A fixation was recorded when a user fixated an area with a radius of 35 pixels for at least 100 ms.

On the top of that, the last fixation before click and the last fixation before the courser moving to the secondary level category were excluded. The reason is that these fixations were accompanied by decision process for navigation and mouse movement and thus were significantly longer than the other fixations during visual search. Saccade length was calculated based on the numbers of items user jumped (how many items between two fixations). For top-level and secondary-level in vertical layout, the saccades in Y coordinate are divided by the item line height. For the secondary level in vertical layout, the saccades in X coordinate are divided by the average category item width. And the saccade directions were measured based on the percentage of saccade from secondary level to the same group, to the other groups, and to the group titles respectively.

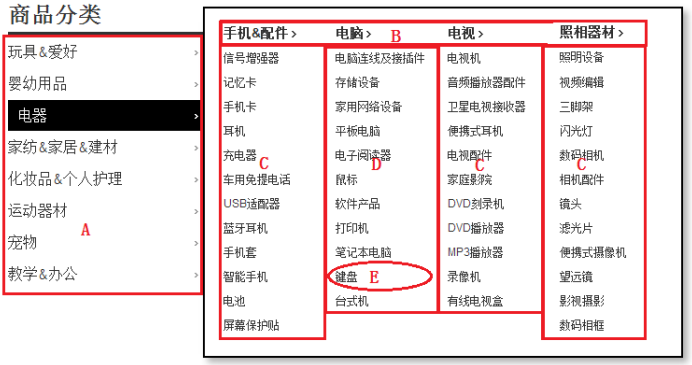


Fig. 2. Sample of area definitions in category design

5 Results

5.1 Fixation Durations

An analysis of repeated measures was conducted to evaluate average duration of fixations with category levels (top level, secondary level) as within-subjects factors and languages as between-subjects factors. This revealed a significant main effect of category level, $F(1, 28) = 24.925$, $p<.001$, indicating that the fixation duration on top level were significantly longer than that on the secondary level. There was also a significant main effect of language, $F(2, 28) = 7.491$, $p<.01$, indicating that fixation duration were significantly different in language condition. And the levels × language interaction was also significant, $F(2, 28) = 6.711$, $p<.01$, indicating that English and Russian users take longer when processing the top level category, while Chinese users didn't show any significant difference. The means and standard deviations for each of the groups are reported in Table 1.

Table 1. Average duration (ms) of fixations on two level categories

	Chinese	English	Russian
Top level category	397(14)	396(14)	340(10)
Secondary level category	384(16)	328(7)	321(10)

An analysis of repeated measures was conducted to evaluate average duration of fixations with searching area (target, non-target) on category, as within-subjects factors, and languages as between-subjects factors. There was a significant main effect of searching area, $F(1, 28) = 34.874$, $p<.001$, and a significant main effect of language $F(2, 28) = 5.474$, $p<.05$, and also a significant interaction $F(2, 28) = 5.223$, $p<0.05$. This indicated that Chinese and English users have longer fixation duration on target category area than in non-target area, while Russian users didn't show any significant difference. The means and standard deviations for each of the groups are reported in Table 2.

Table 2. Average duration (ms) of fixations on target and non-target area

	Chinese	English	Russian
Target category area	399(13)	346(13)	328(13)
Non-target category area	345(12)	310(12)	318(11)

5.2 Saccade Length

A Kruskal-Wallis test was conducted to evaluate differences among the three groups (Chinese, Russian, English) on average saccade length (number of items) on the top level category ($\chi2(2)=19.374$, $p<.001$), on vertical secondary level category ($\chi2(2)=5.737$, $p>.05$), and on horizontal secondary level category ($\chi2(2)=22.229$, $p<.001$) (See Table 3). Follow-up Mann-Whitney tests indicated that Russian users' saccade lengths are shorter than Chinese and English users on the top level category, and the saccade lengths of Chinese users are longer than that of English users and longer than Russian users on the horizontal secondary level category. A Wilcoxon signed-rank test showed that there was a statistically significant different saccade length between top level and secondary level for all languages.

Table 3. Saccade length (measured by the number of menu items)

	Chinese	English	Russian
Top level	1.75 (0.03)	1.90 (0.10)	1.27 (0.07)
Secondary level (Vertical)	2.76 (0.09)	2.77 (0.13)	2.41 (0.12)
Secondary level (Horizontal)	1.62 (0.06)	1.11 (0.05)	0.95 (0.03)

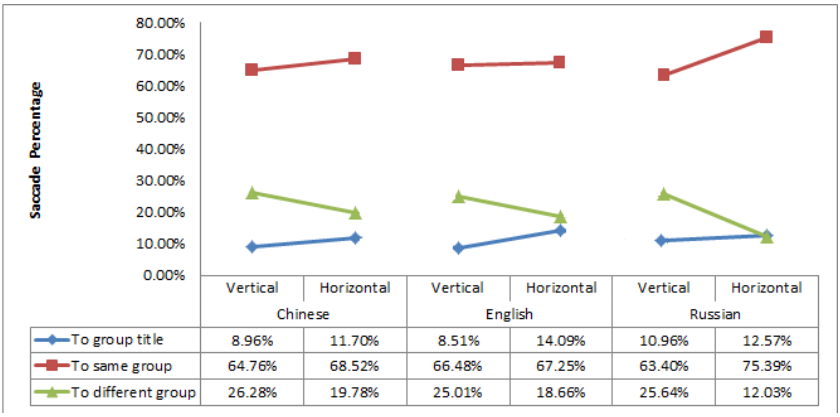
5.3 Saccade Direction

The saccades from secondary level were divided into the following types: to the same category group (A), to the other groups (B), and to the group titles (C) (see Fig. 3). And Fig. 4 shows the percentage of saccade direction of each group respectively.

An analysis of repeated measures was conducted to evaluate average saccade directions (from group to group titles) with layouts (vertical, horizontal) as within-subjects factors, and language (Chinese, English, and Russian) as between-subjects factors. There was a significant main effect of saccade directions, $F(1, 28) = 4.567$, $p<.05$, indicating that there are more saccades from secondary category group to group title in horizontal layout than in vertical layout. However, there was no significant main effect of language ($F (2, 28) = 0.442$, $p=.647$) and interaction between the two factors ($F (2, 28) = 0.297$, $p=.745$).

An analysis of repeated measures was conducted to evaluate average saccade directions (from group to other groups) with layouts (vertical, horizontal) as within-subjects factors, and language (Chinese, English, Russian) as between-subjects factors. There was a significant main effect of saccade directions, $F (1, 28) = 17.120$, $p<.001$, indicating that there are more saccade from secondary category group to other groups in vertical layout than in horizontal layout. However, there was no significant main effect of language ($F (2, 28) = 0.913$, $p=.413$) and interaction between the two factors ($F (2, 28) = 2.125$, $p=.138$).

Fig. 3. Samples of the saccade direction on the secondary level category



	Chinese		English		Russian	
	Vertical	Horizontal	Vertical	Horizontal	Vertical	Horizontal
To group title	8.96%	11.70%	8.51%	14.09%	10.96%	12.57%
To same group	64.76%	68.52%	66.48%	67.25%	63.40%	75.39%
To different group	26.28%	19.78%	25.01%	18.66%	25.64%	12.03%

Fig. 4. Percentage of saccade direction on secondary level category

6 Discussion

Comparisons on eye fixations between different levels in this study supports that the fixation durations are different at different category areas (H1). Previous research argued that fixation duration is longer on complicated search array [23]. It is contrary to see in this case that fixation durations on the top level items that are less complicated than the secondary category items are longer. Since cognitive processes associated fixated words are important to the fixation duration [4], it is improper to judge fixation duration only based on visual complexity in search. Moreover, there is no measurement of reading difficulty among all category items, which can lead to different fixation duration. From a semantic point of view, the words in the top category are the superordinate terms which are more general and higher in the semantic hierarchy while the secondary category words are the subordinate terms which are more concrete and are lower in the semantic hierarchy. As words higher in the hierarchy have wider scope than those they dominate, it takes longer to comprehend and process, which, in this case, result in longer fixation stay in the top category.

In addition, comparisons between fixation durations in target and non-target areas provided further support to (H1). In the targeted area, all words belong to the same semantic field and carry similar semantic meanings when comparing with words in other groups. The search task requires users to find and select the best semantically matched (either identical match or close approximation) item. Users need to distinguish target from distractors and determine whether it is the desired one or not. Previous studies have pointed out that during searching tasks, users distinguish targets from distractors based on primary features such as shape, size, and orientation [22]. This experiment provided furthermore evidence that the semantic relationship in the array is another fundamental feature for user to distinguish.

H2 states that visual search strategies employed are influenced by category design. This experiment figured out where to look next in category from two main saccade features: direction and extent. The data analysis used number of items as a unit and used saccade location to examine interpretable eye movement patterns. Hornof and Kieras (1997) reported that people process more than one menu item at a time and the average saccade length in vertical direction was 2.21 menu items [11]. Our results demonstrate that the saccade lengths are not fixed. The reason of shorter saccade lengths in the top level category was due to the semantic reason rather than the search array complicity. The higher in semantic hierarchy, the more cognitive workload required and the shorter saccade lengths will be. Furthermore, saccade length varied a lot across different languages. The language influences are especially noticeable on horizontal layout. The reading direction of the Chinese, English, and Russian are congruent, the widths of the words are different. The results indicate that the shorter language word width the more category items covered in one saccade, which may related to the visual search efficiency. Hypothesis (H2a) was confirmed by the findings above.

From the quantitative fixation location data, it is obviously to see that most visual search follow the direction of the category layout within the same group. Most of previous researches focused on the efficiency comparisons of vertical and horizontal

layout, however, our study emphasized on the search path. Previous research found that group labels play an important role in facilitating visual processing [13,19], because grouping will introduce another level of search. When titles are presented, an initial search conducted can only examine the group headings to determine whether reading more group items [10]. Our experiment result illustrated how group titles used in searching tasks is largely influenced by the layout. User during visual search rely more on group titles on horizontal layout, while on vertical layout, with better aliment cues, they tend to do direct searches between groups. Thus, we conclude that our data is in accordance with the hypothesis (H2b) that design layout affect users' visual search direction on category navigations.

A limitation of our work was the limited number of participants in our study cannot provide persuasive support for the conclusion associated with the language factor. Another limitation was uncontrolled web elements related to language which can also affect fixations, saccades and regressions such as typographical variables [20]. As far as our prototypes concerned, the participants were novices and their search pattern might be different with experienced users. More research will be needed to validate those differences in the future.

7 Conclusion

This study provides new theoretical and practical insights into the elementary visual search pattern in e-commerce category navigations. The results showed that the semantic scope of the fixated word and semantic relationship in the array have an effect on fixation durations and saccade lengths in visual search. On category design with horizontal layout, the search saccade lengths are highly associated with the language factor which indicated that menu layout is language-sensitive. Different language patterns are observed. Significant differences in search patterns (saccade direction) on different layouts are also found. Insights provided in this study have implications for both fundamental visual search study and improvements in the practical e-commerce category design.

Acknowledgements. We would like to show our sincere gratitude to our colleague Zhao Chen for her valuable reviews and suggestions and the Alibaba International User Research team for their dedicated support and contribution to this paper.

References

1. Aaltonen, A., Hyrskykari, A., Räihä, K.-J.: 101 Spots, or How Do Users Read Menus? In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 1998, vol. 42, pp. 132–139 (1998)
2. Backs, R.W., Walrath, L.C., Hancock, G.A.: Comparison of horizontal and vertical menu formats. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 31, pp. 715–717. SAGE Publications (1987)

3. Card, S.K.: Visual search of computer command menus. Attention and Performance X: Control of Language Processes, pp. 97–108 (1984)
4. Castelhano, M.S., Rayner, K.: Eye movements during reading, visual search, and scene perception: An overview. In: Cognitive and Cultural Influences on Eye Movements, vol. 2175, p. 46 (2009)
5. Cave, K.R., Wolfe, J.M.: Modeling the role of parallel processing in visual search. Cognitive Psychology 22(2), 225–271 (1990)
6. Dong, J., Salvendy, G.: Designing menus for the Chinese population: horizontal or vertical? Behaviour & Information Technology 18(6), 467–471 (1999)
7. Eckstein, M.P.: Visual search: A retrospective. Journal of Vision (2011)
8. Greene, H.H.: The control of fixation duration in visual search. Perception 35(3), 303–315 (2006)
9. Hornof, A.J.: Computational models of the perceptual, cognitive, and motor processes involved in the visual search of pull-down menus and computer screens. Dissertation Abstracts International: Section B: The Sciences and Engineering. ProQuest Information & Learning, US (2000)
10. Hornof, A.J.: Visual search and mouse-pointing in labeled versus unlabeled two-dimensional visual hierarchies. ACM Transactions on Computer-Human Interaction (2001)
11. Hornof, A.J., Kieras, D.E.: Cognitive modeling reveals menu search in both random and systematic. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 1997, pp. 107–114 (1997)
12. Jacob, R.J.K., Karn, K.S.: Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. Work 2(3), 573–605 (2003)
13. Leuthold, S., Schmutz, P., Bargas-Avila, J.A., Tuch, A.N., Opwis, K.: Vertical versus dynamic menus on the world wide web: Eye tracking study measuring the influence of menu design and task complexity on user performance and subjective preference. Computers in Human Behavior 27(1), 459–472 (2011)
14. MacGregor, J.N., Lee, E.: Menu search: Random or systematic? International Journal of ManMachine Studies 26(5), 627–631 (1987)
15. Neisser, U.: Decision-time without reaction-time: Experiments in visual scanning. The American Journal of Psychology 76(3), 376–385 (1963)
16. Plocher, T.A., Choong, Y.-Y.: Cross-cultural Design for IT Products and Services, vol. 36. CRC Press (2012)
17. Rayner, K.: The Thirty Fifth Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. Quarterly Journal of Experimental Psychology 62(8), 1457–1506 (2009)
18. Shen, M., Tao, R., Liu, Y.: A study on visual search and selection of pull-down menu. Hang Tian Yi Xue Yu Yi Xue Gong Cheng Space Medicine Medical Engineering 16(1), 72–74 (2003)
19. Shneiderman, B., Plaisant, C.: Designing the User Interface: Strategies for Effective Human-Computer Interaction, 4th edn. Pearson Addison Wesley (2004)
20. Slattery, T.J., Rayner, K.: The Influence of Text Legibility on Eye Movements During Reading. Applied Cognitive Psychology 24, 1129–1148 (2010)
21. Smith, S.L., Mosier, J.N.: Guidelines for designing user interface software, Citeseer (1986)
22. Treisman, A.: Features and objects in visual processing. Scientific American 255(5), 114–125 (1986)
23. Vlaskamp, B.N.S., Hooge, I.T.C.: Crowding degrades saccadic search performance. Vision Research 46(3), 417–425 (2006)

24. Ware, C.: Visual thinking: For design. Morgan Kaufmann (2010)
25. Williams, C.C., Henderson, J.M., Zacks, R.T.: Incidental visual memory for targets and distractors in visual search. Perception & Psychophysics 67(5), 816–827 (2005)
26. Woodward, R.M.: Proximity and direction of arrangement in numeric displays. Human Factors: The Journal of the Human Factors and Ergonomics Society 14(4), 337–343 (1972)
27. Xinhuanet. China's B2C e-commerce market booms in Q3 (2014), http://news.xinhuanet.com/english/china/2013-10/29/C_132841948.htm (retrieved February 4, 2014)
28. Ystats. Global B2C E-Commerce Sales & Shares Report (2014), http://ystats.com/en/reports/preview.php?reportId=1040 (retrieved February 4, 2014)