Learning on Multi-Touch Devices: The influence of the distance between information in pop-ups and the hands of the users

Birgit Brucker (b.brucker@iwm-tuebingen.de)
Lara Scatturin (scalara@gmx.de)
Peter Gerjets (p.gerjets@iwm-tuebingen.de)

Leibniz-Institut für Wissensmedien, Schleichstrasse 6, 72076 Tuebingen, Germany

Abstract

Prior research indicated that information processing is influenced by the proximity of the hands to information: visuospatial processing is fostered near the hands, whereas textual processing might not be affected or even inhibited near the hands. This study investigated how the proximity of the hands to digital information in pop-ups influences learning outcomes on multi-touch devices. Depending on the distance between the information in the pop-ups and the hands of the users there were three conditions: (1) all pop-ups opened near the hands, (2) all pop-ups opened far from the hands, and (3) pop-ups with visuospatial information opened near the hands, whereas pop-ups with textual information opened far from the hands (mixed condition). Results showed better learning outcomes when visuospatial pop-ups are presented near the hands, whereas there was no difference in learning outcomes between near and far presented textual pop-ups. Results and implications for multi-touch designs are discussed.

Keywords: Learning; Information in Pop-Ups; Near-Hand Attention; Hand Proximity; Multi-Touch Devices; Design Implementations

Learning with Multi-Touch Devices

Our hands are the perfect interface of our body to get in contact with the world we live in. They allow us to grasp objects or defend us from potentially harmful things. In our daily life not only real objects, but also digital objects on multi-touch devices become more and more present due to the rapid technological development. One general question that arises, is, how such interaction devices should be designed and implemented to support users during acquiring knowledge with digital objects, such as digital pictures and texts (cf. North et al., 2009). Particularly, the conditions under which interactive multi-touch displays are able to facilitate learning are important. This study investigated how the distance between the hands and fingers of the users to additional digital information in pop-ups influences learning on multi-touch devices. Additionally, this was addressed for different types of information in the pop-ups, namely visuospatial and textual information.

Near Hand Attention

A few studies investigate the directness of manipulation in terms of hand proximity on multi-touch-tables. For example, Schmidt, Block, and Gellersen (2009) compared direct input on a multi-touch table display with indirect input where the input is made on the table but the surface on which the action was visible was a separate vertical display. Schmidt et al. (2009) showed that the direct condition led to better results than the indirect version. Moreover, Brucker et al. (2014) and Brucker, Ehrmann, & Gerjets (2016) showed that direct interaction with visuospatial elements (i.e., pictures of art pieces) was beneficial for learning compared to indirect interaction. This leads to the assumption that hand proximity is particularly beneficial for learning about visuospatial information.

Indeed, prior research on information processing near the hands on computers without multi-touch interaction indicated that the processing of visuospatial information is positively influenced when the hands are near the to-be-processed stimuli. For example, Reed, Grubb, and Steele (2006) showed that visuospatial processing is enhanced near the hands, because objects that are located near the hands receive higher visual attention than objects that are distant to the hands. There is a large amount of studies pointing in the same direction that visuospatial processing is fostered near the hands (e.g., Abrams et al., 2008; Cosman & Vecera, 2010; Tseng & Bridgeman, 2011; Vishton et al., 2007). Tseng and Bridgeman (2011) found evidence that the proximity of the hands lead to deeper and more detailed processing of visual information. Cosman and Vecera (2010) showed facilitated figure-ground-distinctions near the hands. Vishton et al. (2012) showed higher visual precision near the hands (lower Ebbinghaus-illusion).

However, Davoli et al. (2010) showed that not only visuospatial processing is fostered, but that also semantic processing might be impaired near the hands. In their first experiment, participants judged sentences classified by the experimenters as meaningless (e.g., “Tim typed his suitcase to the car” instead of a sentence classified as meaningful, such as “Tim carried his suitcase to the car”) more often as meaningful in the near-hand conditions than when the sentences were presented far from the hands (cf. Figure 1). In their second experiment, Davoli et al. (2010) found a reduced stroop-interference in the near-hand condition on a classical stroop-task (naming the color of a word instead of reading it: for example the word “RED” could be written in red [congruent condition] or in green [incongruent condition]).
condition). Thus, participants could better suppress reading the word when their hands were near the stimuli. Davoli et al. (2010) interpreted the results of both studies as impaired semantic processing. According to Graziano and Gross (1994) stimuli that appear outside the border of 20 cm around the hands do not activate bimodal neurons sensitive to both touch and sight and thus, these stimuli can be considered as presented far from the hands. Furthermore, Adam et al. (2012) demonstrated that the proximity of the hands to the stimuli plays also a role while the hands are moving. This is particularly of interest, because during interaction gestures on multi-touch devices the hands have to be moved almost all the time.

**Design Implementations of Information in Pop-Ups**

During interacting with digital information on multi-touch devices it is common that touching on the information activates a certain functionality. We address one specific functionality of information depiction on multi-touch devices – namely pop-ups displaying additional information – by addressing how the information processing of the additional information is influenced by the proximity of the hands to the pop-ups. If pop-ups are used on multi-touch devices the question arises where these pop-ups should open on the display in relation to the position of the hands and fingers that activate it. We will shortly introduce three possible distances: near, far, or mixed.

*Near distances* between the pop-ups and the fingers of the users have the advantage that users easily find the additional information and do not have to run with their eyes over the display to search for it. However, near distances might also lead to coverings because information that is beforehand near to the finger of the user has to be superimposed by the new additional information in the pop-ups. Regarding *far distances* between the pop-ups and the fingers of the users exactly the opposite advantages and disadvantages occur: Coverings can be prevented, but learners might have to use cognitive resources to find the additional information with their eyes (even though the information might be connected to the position of the finger with a line or an arrow or something similar). This might cause a type of split attention effect (cf. Ayres & Sweller, 2005) between the users previous focus, where her/his hands and fingers are, and the new additional information in the pop-ups. The third way of presenting additional information in pop-ups – the *mixed distances* – depends upon the information that is entailed in the pop-up. As abovementioned there is a large amount of studies showing that visuospatial information processing is fostered near the hands. Thus it might be advantageous to depict pop-ups entailing visuospatial information near the hands (even though coverings might occur). For textual information the result pattern is not that clear: there is one study indicating that it might be better to process textual information further away from the hands (cf. Davoli et al., 2010), but this is not enough evidence to make a strong assumption about the best position for pop-ups entailing textual information.

**Hypotheses**

We assumed that information in pop-ups entailing visuospatial information should be learned better when the visuospatial pop-ups are presented near to the hands of the users. For information in pop-ups entailing textual information we did not expect such a difference, although it might be even advantageous if these textual pop-ups are presented far from the hands of the users.

**Methods**

**Participants and Design**

Fifty-six university students (average age: 24.39 years, SD = 4.58 years; 43 female) from a German university were randomly assigned to one out of three conditions in a between-subjects-design with the factor “pop-up distance” (near versus far versus mixed). Each student received 12 Euro for participating in the study. Art history majors were excluded from the study.

**Materials and Domain**

**Learning Materials and Multi-Touch Table** As instructional domain, art history was chosen. The learning materials consisted of five paintings (cf. Figure 2) from the Herzog Anton Ulrich-Museum in Braunschweig, Germany.

![Figure 2: Pictures of the five paintings used in this study.](image)

High quality photographs of the paintings (in the following termed *pictures*) were displayed on a multi-touch table:

1) "Selbstbildnis" (1547) - Ludger tom Ring d. J. (1522-1584),
2) "Porträt des Reinhard Reiners und seiner Ehefrau Gese" (1569) - Ludger tom Ring d. J. (1522-1584),
3) "Frühstücksstilleben" (1642) - Willem Claesz. Heda (1594-1682),
4) "Die Hochzeit des Peleus und Thetis" (1602) - Joachim Anthonisz. Wtewael (1566-1638), and
5) "Die Heilige Katharina" (around 1620/24) - Bernardo Strozzi (1581/82-1644).

The size of the display of the multi-touch table was 128x135 cm with a resolution of 1920px x 2160px via 2-x-Full-HD-projection. We implemented the following interaction
possibilities with the pictures on the multi-touch table. For all five pictures additional information was accessible by touching a “i”-symbol on the bottom right corner. By pressing the “i” the picture turned around and a menu appeared. The menu entailed on the top left a small version of the painting in the middle on the top a short introduction text to the painting, and moreover four thematic index cards with teaser sentences (see Figure 3 for an example).

Figure 3: Menu of “Die Hochzeit von Peleus und Thetis” (1602) with the four thematic index cards.

Each index card could be opened via the “i”-symbol and gave additional information about a certain aspect concerning the painting (e.g., the artist, the story of the painting, details and imagery, space and composition, light and color; see Figure 4 for an example). By touching the “x”-symbol the index card could be closed again to get back to the menu, which could then be itself closed again by touching the respective “x”-symbol on it to go back to the picture. This structure with its three layers (painting – menu – index card) was developed in cooperation with the curators from the museum as well as the computer scientists that implemented the information on the multi-touch table in the context of developing an informal visitor-information-system for the museum (Gerjets et al., 2013).

Figure 4: Example of an opened thematic index cards. The arrows mark the pop-ups within the text and the picture.

We decided to stick to three layers at the maximum, however for some aspects there was more additional information that would not have fit the limited space of the respective index card. For this additional information we decided not to open another layer to avoid the user of getting lost (e.g., Conklin, 1987), but instead used pop-ups. These pop-ups appeared by touching highlighted words or parts of the pictures. The distance in which the pop-ups opened when participants touched them was subject to experimental manipulation (see next subsection).

Hand Proximity of Pop-Up Distance According to the factor “pop-up distance” we compared experimental conditions in which all pop-ups opened near the touching hand/finger of the users (see Figure 5) with conditions in which all pop-ups opened far from the touching hand/finger of the users (at least 25 cm; this distance was chosen to definitely exceed the peripersonal space of 20 cm around the hands, cf. Graziano & Gross, 1994; see Figure 6) with conditions in which pop-ups containing visuospatial information opened near, whereas pop-ups containing textual information opened far (at least 25 cm) away from the touching hand/finger of the users (mixed, see Figure 7).

Figure 5: Near pop-ups condition.

Figure 6: Far pop-ups condition.

Figure 7: Mixed pop-ups condition.

Measures
The measures comprised a questionnaire on demographics and on participants’ familiarity with the domain, a visuospatial ability test, and learning outcome measures.
Demographic Data and Familiarity with the Domain

The demographic questionnaire assessed age, gender, body size, need of glasses or contact lenses, major, and study progress. Moreover, this questionnaire assessed participants’ familiarity with the domain to determine participants’ familiarity with the content domain of this study (i.e., art) and to ensure that all students were novices with respect to this domain. Questions comprised details of the participants’ school education in art (e.g., number of courses taken, grades) and their familiarity with and interest in art, for instance, indicated by their visits in museums or galleries within the last year. Participants received points for answers that indicated at least some familiarity with the domain. Depending on the question they could receive only positive points (e.g., 0 to +4 points), whereas for some questions they could also receive negative points (e.g., -2 to +2 points). Depending on these calculations, a participant could receive points within the range of -28 to +40 points.

Learners’ Visuospatial Abilities

Visuospatial abilities of the participants were assessed with a short version of the paper folding test (PFT, Ekstrom, French, Harman, & Dermen, 1976). This short version consists of ten multiple-choice items. Participants have to choose the correct answer out of five options for each item. The stimuli are depictions of stepwise folded sheets of paper that were perforated in their folded state. The answer options depict the holes of various unfolded sheets of paper with the holes being either in the correct or incorrect positions. A maximum of three minutes is assigned to work on the items. For each correct answer participants received one point (max. 10 points).

Learning outcome measures

Learning outcomes were measured by means of 60 multiple-choice items about the contents entailed in the pop-ups. For each of the 60 implemented pop-ups, there was one multiple-choice item. Most of the items (88 %) had four answer possibilities of which always only one was correct. The remaining items had two answer possibilities. Depending upon the content of the pop-up (visuospatial versus textual), the respective item asked for visuospatial or textual information. Visuospatial items asked for certain details (e.g., depicted objects or parts) from the picture, showed different versions due to color filters or mirroring (see Figure 8 for an example). Textual items asked for specific information that was only given in the texts of the corresponding pop-up (see Figure 9 for an example). For each correct answer participants received one point (max. 60 points).

Procedure

Participants were tested individually. Subsequently to reading a short overview on the study, they worked on the demographics, the questionnaire on participants’ familiarity with the domain, and the PFT. Afterwards, participants were instructed to stand at a fixed position in front of the multi-touch table to control the distance from the table. Then, they started with a practicing task on the multi-touch table to get – depending on the experimental condition – used to manipulating the digital objects and the way they could interact with the depicted information (about four minutes). Subsequently, participants started with the learning phase in which they could – again depending on the experimental condition – freely explore the five pictures of the paintings with the corresponding menus, index cards, and pop-ups (maximal 45 Minutes to explore the five paintings; participants took on average 35.84 Minutes [SD = 7.47]). During the learning phase participants were allowed to zoom and move the digital objects and freely switch between the paintings, their menus, index cards and pop-ups. They were instructed to focus on the information in the pop-ups to ensure that they open preferably all of them to extract the relevant information. Subsequently to the learning phase, the participants answered the 60 multiple-choice items. One session lasted approximately 80 Minutes.
regard to learners’ prerequisites in terms of familiarity with the domain, age and gender.

However, for participants’ visuospatial abilities there was a significant difference between the three experimental conditions ($F(2, 53) = 3.76, MSE = 5.81, p = .03, \eta^2_p = .12$). Bonferroni-adjusted post-hoc comparisons showed that only participants in the mixed condition had higher visuospatial abilities than participants in the far condition ($p = .04$). Thus, we calculated an analysis of covariance (ANCOVA) with the between-subjects factor “pop-up distance” and learning outcomes as dependent variable, in which we included visuospatial abilities as a covariate and moreover the interaction term pop-up distance * visuospatial abilities to test whether there was an interaction between pop-up distance and visuospatial abilities. Because this interaction term did not reach statistical significance, we report for reasons of simplicity the ANCOVA with visuospatial abilities included as covariate, but without incorporating the interaction term (pop-up distance * visuospatial abilities).

Table 1: Means and standard deviations (in parentheses) of learner prerequisites and learning outcomes (% correct) as a function of hand proximity.

<table>
<thead>
<tr>
<th></th>
<th>pop-ups near (n = 18)</th>
<th>pop-ups far (n = 18)</th>
<th>pop-ups mixed (n = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain Familiarity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-28 to +40)</td>
<td>6.83 (9.04)</td>
<td>2.11 (11.65)</td>
<td>-0.15 (11.09)</td>
</tr>
<tr>
<td>Visuospatial abilities (PFT 1-10)</td>
<td>6.61 (2.45)</td>
<td>4.89 (3.01)</td>
<td>6.90 (1.65)</td>
</tr>
<tr>
<td>Age (in years)</td>
<td>24.06 (5.86)</td>
<td>25.00 (4.06)</td>
<td>24.15 (3.84)</td>
</tr>
<tr>
<td>Female Participants</td>
<td>15 (10.67)</td>
<td>12 (8.29)</td>
<td>16 (9.79)</td>
</tr>
<tr>
<td>Learning Outcomes in % correct</td>
<td>58.53 (10.67)</td>
<td>52.98 (8.29)</td>
<td>59.82 (9.79)</td>
</tr>
</tbody>
</table>

**Learning Outcomes**

An ANCOVA with the between-subjects factor pop-up distance and visuospatial abilities as covariate revealed a significant main effect of pop-up distance for learning outcomes ($F(2, 52) = 3.17, MSE = 92.77, p = .05, \eta^2_p = .11$, achieved power = 0.62), whereas there was no effect of visuospatial abilities on learning ($F(1, 52) = 1.08, MSE = 92.77, p = .30, \eta^2_p = .02$, ns). To disentangle the main effect of pop-up distance for the three groups, we calculated contrast analyses in which we compared on the one hand the two conditions with near pictures (pop-ups near and pop-ups mixed) to the condition with far pictures (pop-ups far) and on the other hand the two conditions with far texts (pop-ups far and pop-ups mixed) to the condition with near texts (pop-ups near). These contrast analyses revealed that near pictures lead to better learning outcomes than far pictures ($F(1, 52) = 6.10, MSE = 92.77, p = .02, \eta^2_p = .11$), whereas there were no differences in learning outcomes for near and far texts ($F < 1$, ns). This result pattern is in line with our hypothesis (for means and standard deviations see Table 1).

**Discussion**

This study addressed how the position of pop-ups in relation to the hands and fingers of the users on multi-touch devices influences learning about the information entailed in the pop-ups. Results showed that learning outcomes are better if pop-ups that contain pictures are presented near the hands, whereas there was no difference for learning outcomes between near-presented and far-presented pop-ups that contain texts. This result pattern is in line with prior findings that visuospatial information is better processed near the hands (e.g., Reed et al., 2006). Moreover it is in line with our hypothesis that visuospatial contents should be presented near the hands of users on multi-touch displays.

Both alternative explanations of split-attention (cf. Ayres & Sweller, 2005) or possible coverings of important information were not solely valid. A split-attention effect would have favored the near condition because in the far hand conditions the attention of the learners would have been split between the origin of the pop-up where the finger activates it and the location where the pop-up really opens (at least 25 cm distance). A prevent-coverings explanation would have favored the far hand condition, in which the pop-ups open at least 25 cm away from the users fingers, because this implementation prevents coverings of the relevant contents that might have been caused by the opening pop-ups or the fingers, hands or arms of the users.

For textual contents it seems to be indifferent of how near or far these information is presented to the hands of the users. Maybe we were not able to find any differences for near and far texts because we measured recognition of specific details from semantically correct sentences. In contrast to this measure, the only paper that found evidence for differences regarding textual information so far, addressed semantic processing and contrasted meaningful with meaningless sentences (Davoli et al., 2010). However, we did not investigate this more basal type of semantics, but rather a more complex type of semantics. Weidler and Abrams (2014) showed enhanced cognitive control near the hands. Admittedly, they did not address textual processing directly in their studies, but their results indicate that tasks that highly focus on cognitive control should be enhanced near the hands. This result pattern might also explain parts of the Davoli et al. (2010) study, namely the reduced Stroop-inference-effect that can be also explained by higher cognitive control near the hands and not by worse semantic processing. Hence, particularly the complexity of the textual materials might have influenced the information processing. Thus, one might have also assumed that textual information should also be better processed near the hands. However, our results showed neither better nor worse performance for pop-ups containing textual information near the hands. We cannot preclude that both processes – the worse processing of semantic stimuli (cf. Davoli et al., 2010) and the enhanced processing due to higher cognitive control (cf. Weidler & Abrams, 2014) – might have influenced learning about textual information in this study. Further research is needed to disentangle these concurring explanations.
Another important difference from our study compared to prior research in this field is the direct interaction with the materials. Participants directly manipulated the pop-ups during learning, instead of only holding their hands next to the stimuli as in many prior studies. Thus, in the present study the participants were involved by their active manipulation of the given materials (e.g., freely choosing with which object they want to interact, moving and zooming of objects). Under the evolutionary assumption visuospatial stimuli are potential candidates for manipulation, because grasping of desirable objects or withdrawing the hands in case of dangerous or harmful objects was important in the evolution of human beings. Thus visuospatial stimuli are much more likely to be interacted with than textual stimuli for which no such evolutionary assumption exists. This might also give some hints for the result pattern that we found differences for pop-ups with pictures, but not for pop-ups with texts. Future studies should investigate the importance of the direct manipulation for our result patterns by comparing interactive with non-interactive conditions.

Moreover, in this study the hands and the fingers of the participants, with which they opened the pop-ups, was visible for both the near pop-ups, as well as the far pop-ups, whereas in prior research the hands of the participants were often not visible in the far hand conditions as the hands are for example positioned in the lab of the participants. The visibility of the hands and fingers might have influenced the result pattern even though the far pop-ups opened at least 25 cm away from the finger of the participants.

Further research is needed to replicate our findings. In future studies the exact distance of the pop-ups in the different conditions should be assessed. Moreover, the pop-up attendance should be gathered as a manipulation check whether all participants really accessed all relevant information by opening all pop-ups. Additionally, eye-tracking data would deliver more insights in the question which information the participants really processed during learning. Furthermore, assessing verbal and visual memory skills in addition to visuospatial abilities might contribute to the understanding of the different result patterns for visuospatial and textual contents in the pop-ups.

In sum, the results from this study yield direct implications for designing multi-touch environments: Let pop-ups containing visuospatial information open near the hands, but let pop-ups containing textual information open further away to prevent coverings if the size of the display and the number of users allow for such a far presentation.

References


