Human-computer interaction in radiology

Jorritsma, Wiard

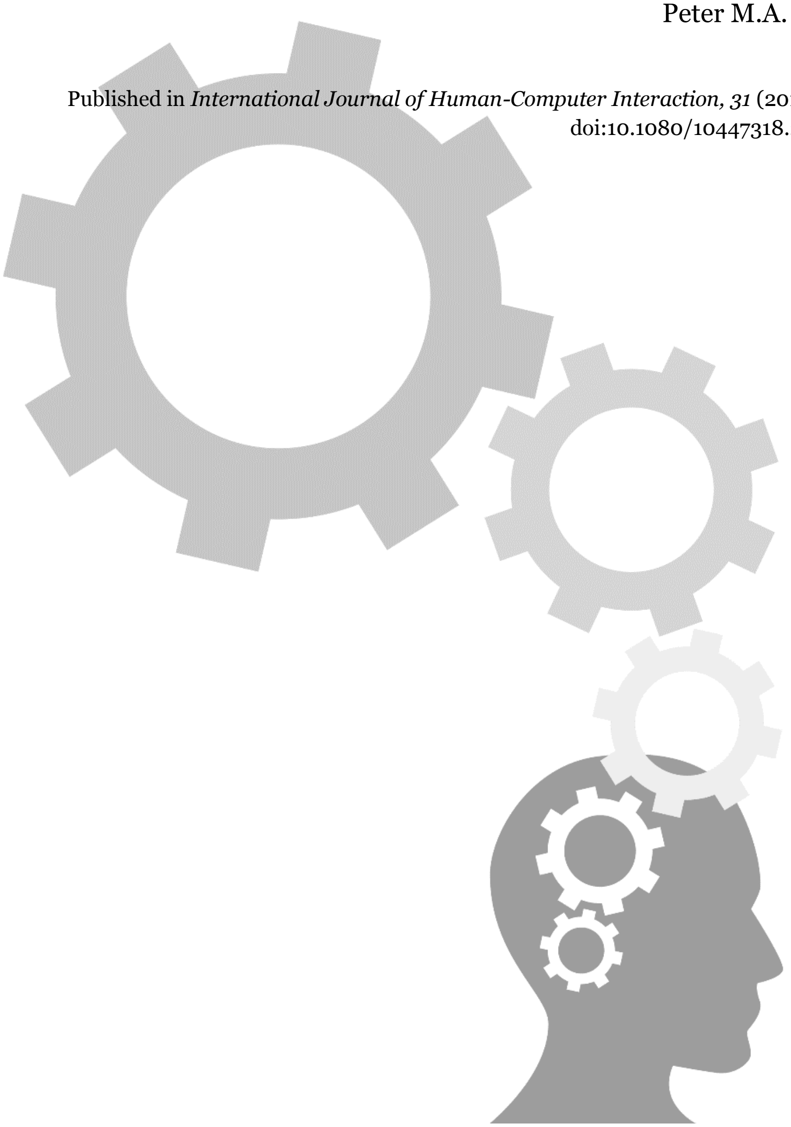
Publication date:
2016

5

Comparing four touch-based interaction techniques for an image-based audience response system

Wiard Jorritsma
Jonatan T. Prins
Peter M.A. van Ooijen

Published in *International Journal of Human-Computer Interaction, 31* (2015), 440–450.
doi:10.1080/10447318.2015.1039437



Abstract

This study aimed to determine the most appropriate touch-based interaction technique for I2Vote, an image-based audience response system for radiology education, in which users need to accurately mark a target on a medical image. Four plausible techniques were identified: land-on, take-off, zoom-pointing and shift. The techniques were implemented in such a way that they could be used on any modern device. An empirical study was performed in which users marked a target on an image using all four techniques on either a smartphone or a tablet. The techniques were compared in terms of accuracy, efficiency, ease-of-use, intuitiveness, and compatibility with the different devices. The results showed that shift was the most accurate technique, but it was hampered by its high complexity and low intuitiveness. Land-on was the fastest technique, but also the least accurate. Take-off and zoom-pointing provided the best trade-off between accuracy, efficiency, ease-of-use, and intuitiveness. We therefore conclude that both take-off and zoom-pointing are viable interaction techniques for I2Vote.

Interaction techniques are the primitive building blocks from which a user interface is crafted.

– Foley et al., Computer Graphics: Principles and Practice

Introduction

Audience response systems (ARSs) aim to create interactivity between a presenter and his or her audience. They allow presenters to incorporate a set of questions into their presentation to which each member of the audience can respond using a handheld device. After each question, the presenter can display the response data on the projection screen. An ARS can make lectures more engaging and fun for the attendees, and potentially improves the information transfer between the presenter and the audience. Various positive effects of using an ARS have been demonstrated, including increased lecture attendance [1–4], increased attention [5,6] and engagement [4,7] during lectures, and improved learning performance [2,8,9].

Although most ARSs only facilitate simple multiple-choice questions, it is also possible to support more advanced types of questions. This is done by the I2Vote system [10], which supports image-based questions for radiology education. With I2Vote, the teacher presents students with a radiological image and asks them to identify a certain region in the image (e.g. a lung nodule in a chest X-ray). Students can view the image on a PDA and place a mark in the region of their choice. When the response period is closed, the marks of all students are displayed on the image on the projection screen. All correct marks (within a predefined region set by the teacher) are displayed in green and incorrect marks in red (Fig. 1).

Supporting these types of questions means that the user interaction with the handheld device becomes much more complex. Instead of simply selecting an

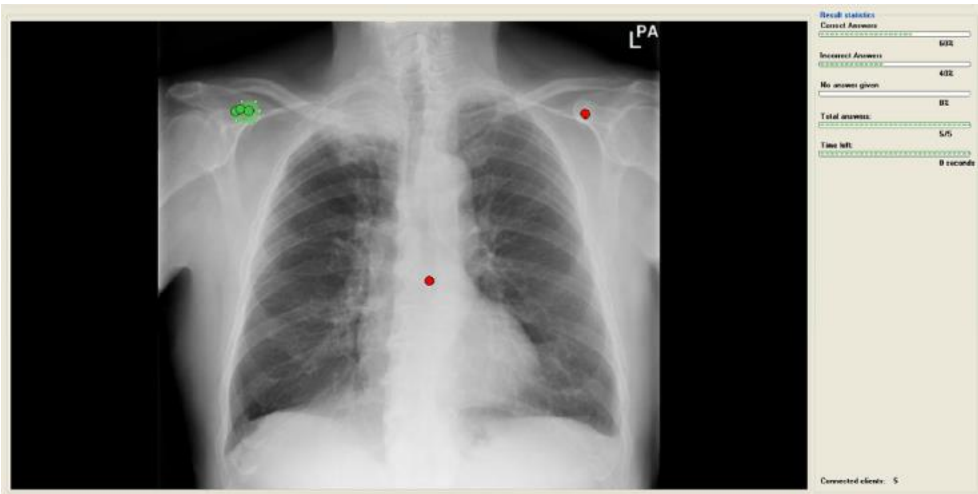


Figure 1. Example of the displayed I2Vote response data for the task of marking the right-hand side processes coracoideus. Circles represent the marks made by the students. Green circles are correct and red circles are incorrect marks. The pane to the right of the image shows the percentage of correct answers, incorrect answers, and no answer given, the total number of respondents, and the time left to submit a response.

answer to a multiple choice question, users now have to accurately mark a region on an image. It is very important to design an easy-to-use and intuitive user interface for this task, to ensure that users do not feel intimidated to use the ARS and can easily and enjoyably interact with it.

Following the BYOD (bring your own device) trend, several ARSs now allow users to use their own devices instead of devices provided by the presenter (e.g. eClicker [11], iResponse [12] and Socrative [13]). This means that the costs for dedicated ARS handheld devices are eliminated, users can interact with the ARS using a device that is familiar to them, and data from a presentation can be stored on the device so that it can be reviewed later. However, supporting BYOD poses the additional challenge of designing a user interface that works well on a wide variety of devices.

Due to technical limitations of the PDA it was originally developed on and to facilitate BYOD, the I2Vote system is currently being modified to support modern smartphones and tablets as its audience input devices. This also means that the user interaction with the devices changes from stylus-based to finger-based touch interaction. In this study, we aimed to determine which touch-based interaction technique would be the most appropriate for the new I2Vote system. To do this, we compared several touch-based interaction techniques for the task of marking regions on an image. Techniques were compared in terms of accuracy, efficiency, ease-of-use, intuitiveness, and compatibility with different devices.

Touch-based interaction techniques

The standard touch-based interaction technique is the *land-on* technique. With this technique, users select a target by directly touching it. The land-on technique is extremely simple and intuitive, but it has two problems when fingers are being used as input: (1) fingers have a large contact area with the screen, resulting in an ambiguous selection point, and (2) fingers may occlude the selection target [14].

To eliminate these problems, the *take-off* technique [15] draws a cursor above the finger while it touches the screen. Users can drag the cursor to the target and select it by removing their finger from the screen. Potter et al. [15] showed that users made fewer errors and were more satisfied with the take-off technique than with the land-on technique. However, it took them longer to select targets.

Another technique is *zoom-pointing*. This technique allows users to zoom in on the target (by using any zoom technique), thereby making it easier to touch. By enlarging the contact area of the target, this technique reduces, or even eliminates, the negative effects of the two problems discussed above. However, it requires users to perform additional zooming actions and when zoomed in, the context of the target is lost. Albinsson and Zhai [16] implemented this technique using bounding box zoom, which zooms in on a rectangular area that the user draws around the target. They found that zoom-pointing outperformed take-off in terms of speed, error rate and subjective preference.

The *shift* technique [14] detects targets within the finger's contact area with the screen. If the area underneath the finger contains a target that is smaller than a predefined occlusion threshold, a circle containing a copy of the occluded area and a cursor at the selection point of the finger is drawn in a non-occluded location. The user can now move the cursor over the target and release the screen to select it. For large targets, the circle does not appear and users can select the target with one touch. Shift is similar to the take-off technique, but it has the advantage that users can aim directly for the target and can select large targets directly, without needing the cursor.

Vogel and Baudisch [14] compared shift to the take-off and land-on techniques. They found that both shift and take-off had a lower error rate than land-on for small targets, but there was no difference between shift and take-off. For small targets, target selection was equally fast with shift and take-off, but much faster with land-on. For large targets, shift was faster than take-off and comparable to land-on. Users preferred shift to take-off and take-off to land-on.

Vogel and Baudisch also implemented a high-precision version of shift, in which the area in the circle was magnified and the circle followed the movement of the finger. The control display ratio was lowered, allowing the circle, and the cursor within it, to be moved more precisely. Like zoom-pointing, this version of shift increases the target's contact area, making it easier to touch. In contrast with zoom-pointing, shift preserves the context of the target and does not require users to perform any zooming actions. This version of shift is theoretically more accurate than the original version, but it has not been empirically validated.

There are many more interesting and possibly accurate interaction techniques (e.g. back-of-device interaction [17,18], dual finger selection [19] and "gravity" selection (where the cursor is "pulled" towards the target) [20–22]), but the four techniques discussed above are the most plausible for an image-based ARS. Other techniques are (currently) not suitable for such a system for various reasons (e.g. back-of-device interaction is not supported by most devices, dual finger selection is awkward on small screens, and gravity selection only works when the system knows the user's intended target, which is impossible in an image-based ARS).

Research objectives

We have identified four plausible interaction techniques for an image-based ARS (i.e. for accurately marking regions on an image): land-on, take-off, zoom-pointing and shift. Although there are studies that have compared several of these techniques, all four techniques have never been compared directly. Furthermore, the techniques have not been evaluated on modern devices. Touch screen quality and screen size of these devices might differentially affect the accuracy of the techniques. This is especially important if the technique is used in a BYOD setting, since different users use different devices and the technique should work well on any of them.

Also, take-off and shift interfere with the native gestures of modern devices, since the touch-and-drag action used by these techniques is identical to the native scrolling action. This means that the implementation of these techniques needs to change in order for them to be compatible with today's devices. Zoom-pointing on the other hand, could benefit from the native "pinch to zoom" gesture, which can be used effectively by most smartphone or tablet users, and could therefore increase the efficiency of this technique.

In this study, we implemented the land-on, take-off, zoom-pointing and shift techniques in such a way that they could be used on any modern touch-based device without interfering with the device's native gestures. In order to determine which technique would be the most appropriate for I2Vote, we performed an empirical study in which users marked a target on an image using all four techniques on either a smartphone or a tablet. The techniques were evaluated in terms of accuracy, efficiency, ease-of-use and intuitiveness. We also investigated how the different devices affected the performance of the techniques.

Implementation

We implemented the four interaction techniques using KineticJS: an HTML5 Canvas JavaScript framework. This allows them to be used inside a browser on any device. With all techniques, users can place a mark on an image. When they are satisfied with the placement of the mark, they press a submit button located below the image. Fig. 2 shows an example of the software running on a smartphone.

Land-on

With the land-on technique, users mark the target by tapping on it directly (Fig. 3).

Take-off

In the original take-off technique, a cursor appears above the user's finger while it touches the screen. The user then drags the cursor to the target and releases the screen to select it. Because this implementation interferes with the native scroll gesture of modern devices, we implemented the technique in a different way while staying true to its main idea: creating a distance between the finger and the cursor to avoid finger occlusion and the problem of the finger's ambiguous selection point.

Figure 2. Example of our software running on a smartphone. The user interface consists of an X-ray and a submit button. This example shows a target (black dot) and a mark (red dot) placed by the user.

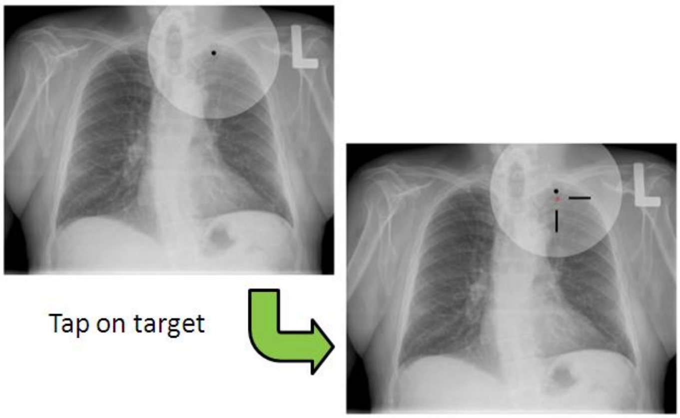


Figure 3. The land-on technique. Users mark the target (black dot) by tapping on it. With each tap, a mark (red dot) appears. When users are satisfied with the placement of the mark, they click on a submit button.

In our implementation (Fig. 4), a crosshair appears when users tap the screen. The crosshair consists of a large gray circle with a mark (red dot) in the center. The mark is placed at the location at which the user tapped the screen. Users can "pick up" the crosshair by touching anywhere within the gray circle and drag it in order to place the mark on the target.

Note that in the original take-off implementation, the cursor (or mark) is placed above the user's finger, immediately creating the distance between the mark and

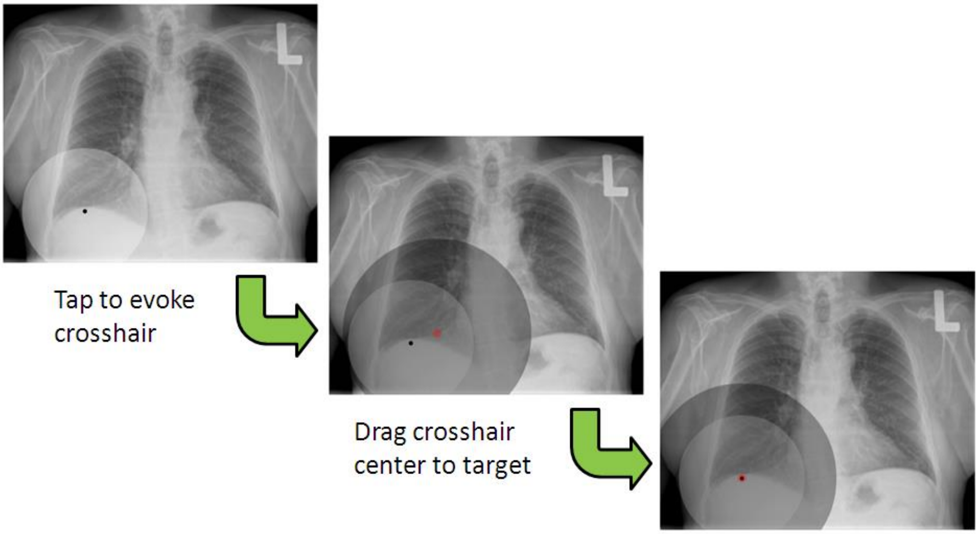


Figure 4. The take-off technique. Users tap the screen to evoke a crosshair (gray circle with a mark (red dot) in the center). They can then drag the crosshair in order to place the mark on top of the target (black dot). When users are satisfied with the placement of the mark, they click on a submit button.

the cursor. In our implementation, the mark is initially placed underneath the user's finger. The distance between the finger and the mark is created when the user drags the crosshair using the gray circle. This has two advantages: (1) users can aim directly for the target and adjust their initial mark – if necessary – by dragging the crosshair. This means that large targets can be selected with one tap, without the need to drag the crosshair. (2) When moving the mark, users have control over where the mark is placed relative to their finger. In the traditional take-off, the mark is always placed above the finger. In our implementation, the location of the mark depends on where the user picks up the crosshair. This can be advantageous when the target is located near the borders of the image. If the target is close to the lower border for example, it might be easier to drag the crosshair by its top, thereby having the mark below the finger.

Zoom-pointing

With zoom-pointing, users zoom in on the target and then mark it by tapping on it (i.e. zoom followed by land-on) (Fig. 5). Zooming is done with the native pinch to zoom gesture. Note that this technique often requires users to zoom out or scroll down in order to press the submit button, which is located below the image.

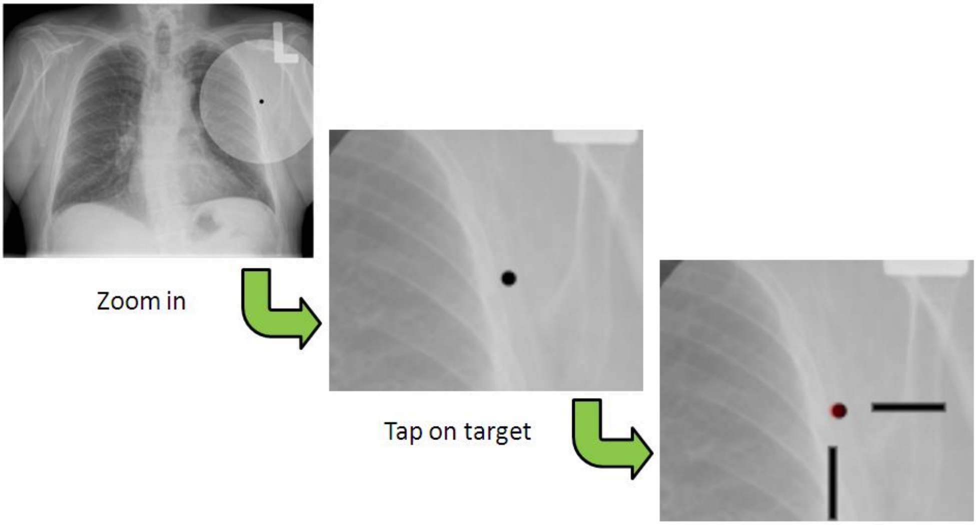


Figure 5. The zoom-pointing technique. Users first zoom in on the target (black dot) by using the pinch to zoom gesture and then mark the target by tapping on it. With each tap, a mark (red dot) appears. When users are satisfied with the placement of the mark, they click on a submit button.

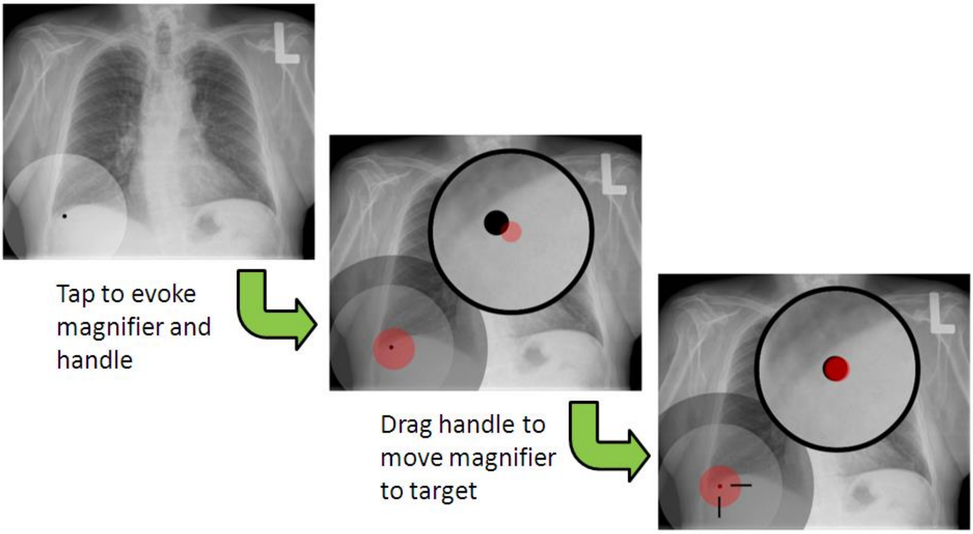


Figure 6. The shift technique. Users tap the screen to evoke a magnifier (circle with a black border) and a handle (gray circle with a red circle in the center at the location of the tap). The magnifier shows a 6x magnification of the red circle and has a mark (red dot) in the center. Users can move the magnifier, and thereby the mark, over the image by dragging the handle. Releasing the handle places the mark. When users are satisfied with the placement of the mark, they click on a submit button.

Shift

Our shift implementation is based on the high-precision version of shift. Like the original take-off technique, high-precision shift interferes with the native scroll gesture of modern devices. Therefore, our implementation differs slightly from the original. In our implementation (Fig. 6), a tap on the screen evokes a magnifier (circle with a black border) and a handle (gray circle with a red circle in the center at the location of the tap). The magnifier shows a 6x magnification of the red circle and has a mark (red dot) in the center. Users can drag the handle, thereby moving the magnifier over the image. This allows them to place the mark within the magnifier on the target. When the handle is released, the mark is also placed in the unmagnified image. Another tap on the screen (in an area that does not contain the handle) removes the magnifier and the handle and leaves the mark.

The location of the magnifier is determined as follows: if the center of the handle is in the lower half of the screen, the magnifier is placed above the handle (and vice versa) and if the center of the handle is in the left half of the screen, the magnifier is placed to the right of the handle (and vice versa).

As in high-precision shift, we lowered the control display ratio for moving the magnifier. We used a ratio of 1:4 instead of the default 1:1. This means that the magnifier moves at 25% of the speed of the finger, allowing for more precise placement of the mark.

Unlike the original shift, which only displays the magnifier for small targets, our implementation always displays the magnifier. Distinguishing between small and large targets is impossible for our task, since any area within the image is a potential target. However, it is still possible for users to mark large targets in one tap by simply not moving the magnifier.

Empirical evaluation

Participants

24 participants participated in this experiment (age range: 20–45, mean age: 26.7). 23 participants owned either a smartphone or a tablet, so they were familiar with finger-based touch interaction. The one participant who did not own such a device had no problems performing the experiment.

Apparatus

The smartphone was a Huawei Y300 (4 inch screen with 233 ppi, running Android 4.1.1). The tablet was an Apple iPad (9.7 inch screen with 132 ppi, running iOS 5.1.1). The experiment was developed using JavaScript and HTML and ran inside a browser (Google Chrome on the smartphone and Apple Safari on the tablet).

Materials

Participants were presented with a chest X-ray containing a target and a submit button (Fig. 2). The target was a black dot of 0.6 mm on the smartphone and 0.9 mm on the tablet. It was placed at a random location in between the X-ray's top/bottom and left/right margins. The top and bottom margins were 10% of the X-ray's height, the left and right margins were 10% of the X-ray's width. Participants had to mark the target using one of the four interaction techniques (land-on, take-off, zoom-pointing or shift).

For each technique, participants filled out a questionnaire consisting of three statements: "This technique was easy to use", "This technique was easy to learn" and "I could accurately mark the target using this technique". Responses were made on a five-point Likert scale with the following levels: "strongly disagree", "disagree", "neutral", "agree", and "strongly agree".

Design and procedure

We used a mixed factorial design with device (smartphone or tablet) as between-subject factor and interaction technique as within-subject factor. Participants were randomly assigned to one of the devices, while the number of participants on each device was balanced (i.e. 12 participants performed the experiment on the smartphone and 12 on the tablet). The order of the techniques was counterbalanced using a balanced Latin square design [23].

For each technique, participants read instructions on how to use the technique and then performed eight trials of marking the target. A trial consisted of marking the target and, when satisfied with the placement of the mark, pressing the submit button. Participants were informed that the first two trials were considered training trials. These trials were not included in the data analysis. They were instructed to mark the target as accurately and quickly as possible. All participants used the device in the upright (portrait) orientation.

After the last trial, participants filled out the questionnaire and then continued with the next technique. At the end of the experiment, participants were asked if they had any additional comments regarding the techniques.

Data analysis

Accuracy was measured as the mean deviation of the mark from the target. For each trial, the coordinates of the mark and the target were stored and the distance between them was calculated. To allow for comparison between the smartphone and the tablet, the distances were converted from pixels to mm according to the devices' ppi specifications.

Efficiency was measured as the mean time and the mean number of attempts needed to mark the target. For each trial, the timer started when all graphical elements were loaded and stopped when participants pressed the submit button. For take-off and shift, an attempt was defined as a drag and release of the crosshair/handle. If participants considered their initial tap accurate enough, and therefore did not drag the crosshair/handle, this was considered as one attempt. For land-on and zoom-pointing, an attempt was simply a tap on the screen.

Ease-of-use, intuitiveness and subjective accuracy were measured as the responses to the statements "This technique was easy to use", "This technique was easy to learn" and "I could accurately mark the target using this technique", respectively. To allow for quantitative analysis of the response data, they were converted to a numerical scale ranging from 0 (strongly disagree) to 4 (strongly agree). Participants' comments during and after the experiment, and observations made during the experiment were also used in the assessment of these constructs.

For each of the quantitative measures, the median and the 95% confidence interval (CI) for the median were calculated for each technique on each device. The median was used instead of the mean because the data were not normally distributed. Note that unlike the CI for the mean, the CI for the median is not typically symmetric.

Wilcoxon signed rank tests for paired samples were used to analyze differences between the techniques. Wilcoxon signed rank tests for independent samples were used to analyze differences between the devices. These non-parametric tests were used because the data were not normally distributed. P-values smaller than 0.05 were considered statistically significant. To account for multiple comparisons, P-values were adjusted according to the Holm-Bonferroni method.

Results

Accuracy

Fig. 7 shows the accuracy (i.e. median deviation from the target) for each technique on both devices. Shift was the most accurate technique, with a lower deviation than the other techniques on both devices (shift vs. land-on: $W(12) = 0$, $p = .003$ (smartphone and tablet); shift vs. take-off: $W(12) = 0$, $p = .003$ (smartphone and tablet); shift vs. zoom-pointing: $W(12) = 5$, $p = .01$ (smartphone), $W(12) = 0$, $p = .003$ (tablet)).

Take-off and zoom-pointing were both more accurate than land-on (take-off vs. land-on: $W(12) = 1$, $p = .003$ (smartphone), $W(12) = 0$, $p = .003$ (tablet); zoom-pointing vs. land-on: $W(12) = 0$, $p = .003$ (smartphone and tablet)). There were no

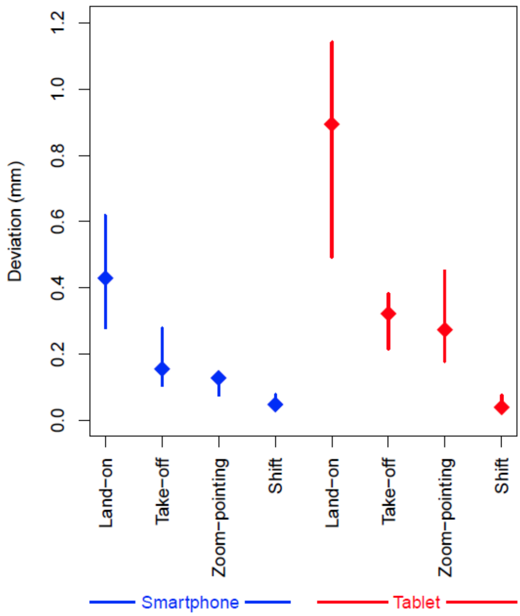


Figure 7. Median deviation from the target for each technique on both devices. Bars indicate the 95% CI.

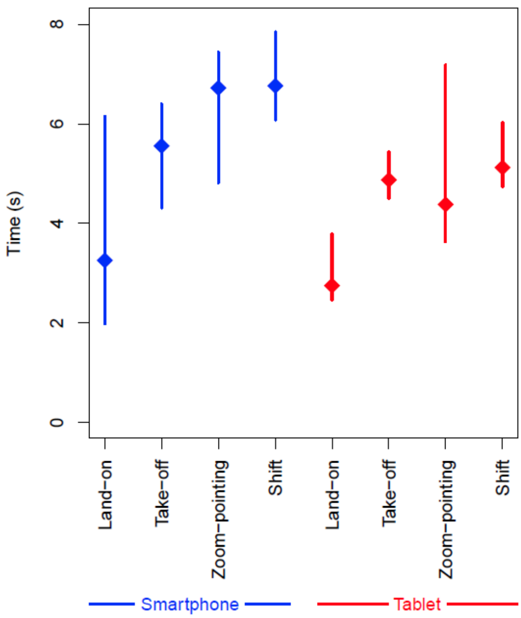


Figure 8. Median time needed to mark the target for each technique on both devices. Bars indicate the 95% CI.

statistically significant differences between take-off and zoom-pointing ($W(12)$ = 60, p = .11 (smartphone); $W(12)$ = 38, p = .97 (tablet)).

Participants using the smartphone were more accurate than participants using the tablet when using the land-on ($W(12)$ = 23, p = .012), take-off ($W(12)$ = 27, p = .016) and zoom-pointing techniques ($W(12)$ = 13, p = .001). There was no statistically significant difference between the devices when using shift ($W(12)$ = 96, p = .18).

Efficiency

Fig. 8 shows the median time needed to mark the target for each technique on both devices. Participants performed faster with land-on than with the other techniques on both devices (land-on vs. take-off: $W(12)$ = 0, p = .003 (tablet); land-on vs. zoom-pointing: $W(12)$ = 0, p = .003 (smartphone), $W(12)$ = 1, p = .004 (tablet); land-on vs. shift: $W(12)$ = 1, p = .005 (smartphone), $W(12)$ = 0, p = .003 (tablet)) However, the difference between land-on and take-off on the smartphone was not statistically significant ($W(12)$ = 10, p = .06).

Take-off was faster than shift on the smartphone ($W(12)$ = 7, p = .037). There were no other statistically significant differences between the techniques (take-off vs. zoom-pointing: $W(12)$ = 22, p = .41 (smartphone), $W(12)$ = 37, p = .91 (tablet); take-off vs. shift: $W(12)$ = 18, p = .33 (tablet); zoom-pointing vs. shift: $W(12)$ = 26, p = .41 (smartphone), $W(12)$ = 26, p = .68 (tablet)).

Participants using the tablet were faster with shift than participants using the smartphone ($W(12)$ = 28, p = .048). There were no statistically significant differences between the devices when using land-on ($W(12)$ = 62.5, p = .60), take-off ($W(12)$ = 46, p = .42), or zoom-pointing ($W(12)$ = 50, p = .44).

Fig. 9 shows the median number of attempts needed to mark the target for each technique on both devices. On the smartphone, the median number of attempts for take-off was lower than for the other techniques, but this difference was not statistically significant (take-off vs. land-on: $W(8)$ = 1, p = .12; take-off vs. zoom-pointing: $W(10)$ = 8, p = .26; take-off vs. shift: $W(8)$ = 11.5, p = .56). There were also no other statistically significant differences between the techniques (land-on vs. zoom-pointing: $W(10)$ = 41, p = .56; land-on vs. shift: $W(9)$ = 39, p = .26; zoom-pointing vs. shift: $W(10)$ = 40, p = .56).

On the tablet, the median number of attempts was lower for land-on, take-off, and shift than for zoom-pointing, but this difference was not statistically significant (take-off vs. zoom-pointing: $W(7)$ = 1, p = .21; land-on vs. zoom-pointing: $W(8)$ = 10.5, p = .98; shift vs. zoom-pointing: $W(6)$ = 1, p = .29). There were also no other statistically significant differences between the techniques (land-on vs. take-off: $W(7)$ = 24.5, p = .36; land-on vs. shift: $W(6)$ = 15, p = .98; take-off vs. shift: $W(4)$ = 6, p = .98).

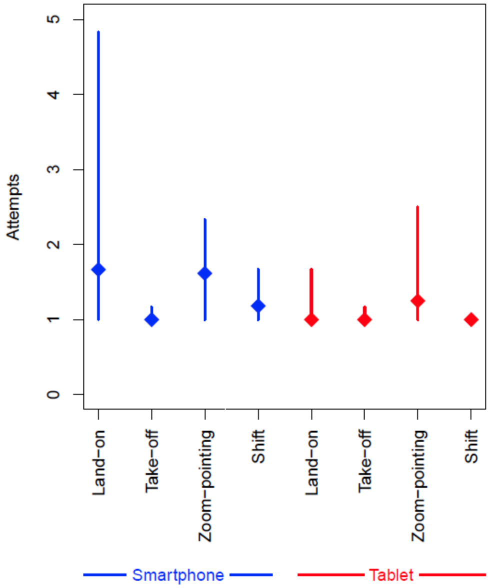


Figure 9. Median number of attempts needed to mark the target for each technique on both devices. Bars indicate the 95% CI.

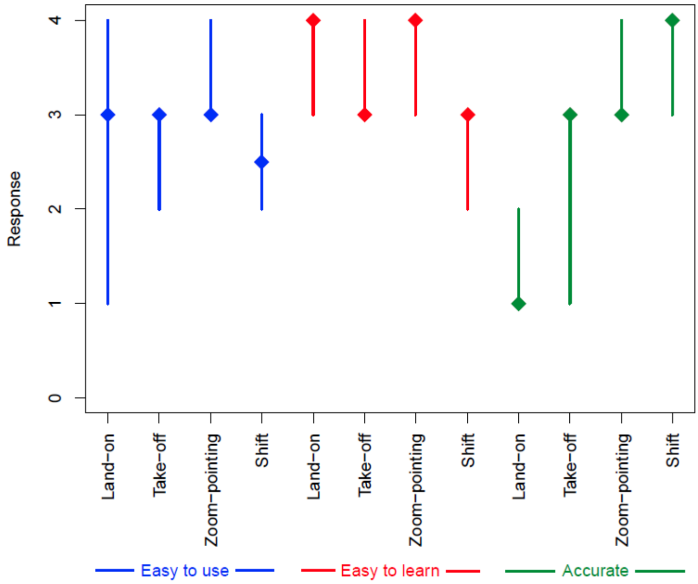


Figure 10. Median response to the questionnaire statements "This technique was easy to use" (blue), "This technique was easy to learn" (red) and "I could accurately mark the target using this technique" (green) for each technique. Responses were made on a five-point Likert scale with the following levels: strongly disagree (0), disagree (1), neutral (2), agree (3) and strongly agree (4). Bars indicate the 95% CI.

The median number of attempts was lower on the smartphone than on the tablet for land-on, zoom-pointing and shift, but these differences were not statistically significant (land-on: $W(9) = 44$, $p = .29$; take-off: $W(7) = 64$, $p = 1$; zoom-pointing: $W(11) = 65$, $p = .70$; shift $W(7) = 39$, $p = .1$).

Ease-of-use

Because the questionnaire responses for the smartphone and tablet were similar, we aggregated the questionnaire data over these devices. There were no statistically significant differences between the devices for any of the techniques on any of the questionnaire statements ($p > .52$).

Fig. 10 (blue data points) shows participants' responses to the statement "This technique was easy to use" for all techniques. The median response for all techniques was positive. The median response for shift was the least positive, but this difference between the other techniques was not statistically significant (shift vs. land-on: $W(21) = 108$, $p = .80$; shift vs. take-off: $W(16) = 53.5$, $p = .92$; shift vs. zoom-pointing: $W(20) = 63$, $p = .33$).

Participants made the following comments regarding the ease-of-use of the techniques. Land-on: "This gives me RSI" (one participant). Take-off: "The mark sometimes moves when I lift my finger from the screen" (three participants); "Dragging the crosshair was a bit jerky" (two participants); "Maybe the crosshair should be magnified" (four participants). Zoom-pointing: "It's annoying to zoom out in order to press the submit button" (two participants); "Zoom is easiest, but my nails are a bit too long" (one participant); "How can I adjust the mark" (two participants who had used take-off and shift before zoom-pointing). Shift: "My thumb sometimes occluded the magnifier" (three participants); "The magnifier sometimes jumped to another location" (one participant).

Intuitiveness

Fig. 10 (red data points) shows participants' responses to the statement "This technique was easy to learn" for all techniques. The median response for take-off and shift was positive and the median response for land-on and zoom-pointing was strongly positive. The difference between zoom-pointing and shift was statistically significant ($W(13) = 85.5$, $p = .016$), the other differences were not (land-on vs. take-off: $W(13) = 69.5$, $p = .091$; land-on vs. shift: $W(17) = 124$, $p = .063$; zoom-pointing vs. take-off: $W(13) = 69$, $p = .18$).

Participants made the following comments regarding the intuitiveness of the techniques. Zoom-pointing: "Anyone knows how to use pinch-zoom" (one participant). Shift: "Too complex; there are many things on the screen" (three participants); "In the beginning, I found it unclear when the target had been marked" (one participant); "The difference between the target and the mark is unclear" (one participant).

The following observations were made that are relevant for the intuitiveness of the techniques. Take-off: two participants expected the crosshair to appear while they touched the screen; one participant tried to drag the crosshair by touching the screen at a location that did not contain the crosshair; one participant expressed frustration when he dragged the crosshair by its center. When he realized he could drag it by touching anywhere within the crosshair, he said that the technique was actually quite good. Zoom-pointing: two participants asked for extra instructions (in addition to the written instructions already provided). One participant tried to zoom by holding his finger on the screen. Shift: seven participants asked for extra instructions; four participants sometimes dragged the magnifier instead of the handle. Three participants' first tap, which evokes the magnifier and the handle, was always quite far from the target. This means that the handle had to be dragged over a relatively large distance, which was slow and awkward because of the handle's low control display ratio.

Subjective accuracy

Fig. 10 (green data points) shows participants' responses to the statement "I could accurately mark the target using this technique" for all techniques. The median response for shift was strongly positive, the median response for take-off and zoom-pointing was positive, and the median response for land-on was negative. The differences between land-on and the other techniques were statistically significant (land-on vs. take-off: $W(19) = 18$, $p = .006$; land-on vs. zoom-pointing: $W(22) = 0$, $p = .0002$; land-on vs. shift: $W(23) = 0$, $p = .0001$). The difference between shift and take-off was statistically significant ($W(18) = 155$, $p = .007$), the difference between shift and zoom-pointing was not ($W(13) = 70.5$, $p = .09$). There was also no statistically significant difference between take-off and zoom-pointing ($W(18) = 40$, $p = .09$).

Participants made the following comments regarding the accuracy of the techniques. Zoom-pointing: "The mark with zoom seems less accurate because you are zoomed in. Therefore you keep trying to get the mark exactly right" (one participant). Shift: "Because shift is so accurate, you spend more time getting the mark exactly right" (one participant).

Discussion

In this study, we implemented four touch-based interaction techniques, land-on, take-off, zoom-pointing and shift, in such a way that they could be used on any modern touch-based device without interfering with the device's native gestures.

We aimed to evaluate which of the techniques would be the most appropriate for I2Vote (i.e. for accurately marking regions on a medical image). To do this, we performed an empirical study in which users marked a target on an image using all four techniques on either a smartphone or a tablet. The techniques were evaluated in terms of accuracy, efficiency, ease-of-use and intuitiveness. We also investigated how the different devices affected the performance of the techniques.

Shift was the most accurate technique. Take-off and zoom-pointing were less accurate than shift, but more accurate than land-on. There was no significant difference between take-off and zoom-pointing. Take-off, zoom-pointing and shift were all quite accurate, with a median deviation from the target of less than 0.2 mm on the smartphone and 0.4 mm on the tablet. Land-on was much less accurate, with a median deviation of 0.43 mm on the smartphone and 0.89 mm on the tablet. Land-on's poor accuracy is in line with previous studies [14,15], confirming that it is very difficult to accurately touch a small target with this technique.

Our results indicate that creating a distance between the user's finger and the mark (as done by take-off) and enlarging the target (as done by zoom-pointing) both have a positive effect on accuracy, and that the size of this effect is roughly the same for both approaches. A combination of these approaches (as done by shift) improves accuracy even further.

Participants marked the target faster with land-on than with the other techniques, but this increase in speed came at the cost of reduced accuracy. There were no statistically significant differences between the other techniques, with the exception that take-off was faster than shift on the smartphone.

On the smartphone, the median number of attempts needed to mark the target was lower for take-off than for the other techniques. On the tablet, the median number of attempts was lower for land-on, take-off, and shift than for zoom-pointing. However, none of these differences were statistically significant.

The comparison between the techniques in terms of the number of attempts needed to mark the target should be interpreted carefully, because this was measured differently for land-on and zoom-pointing than for take-off and shift. An attempt for land-on and zoom-pointing was simply a tap on the screen, while it was defined as a drag and release action of the crosshair/handle for take-off and shift. This makes this measure a lot less meaningful for the latter techniques, since a very long drag action only counts as one attempt. However, repeated tapping might be more frustrating for the user than one continuous dragging action; dragging provides a sense of movement towards a goal, while tapping feels more like trial and error.

Participants rated the ease-of-use of all techniques positively. The median rating for shift was lower than for the other techniques, but this difference was not statistically significant. Participants' comments revealed several problems with the techniques. Three participants (all in the smartphone group) indicated that when using take-off, the mark sometimes moved when they lifted their finger from the screen. Further investigation revealed that when the finger is lifted at an angle, the smartphone registers a movement and therefore moves the mark. This problem did not occur on two more high-end smartphones (a HTC One S and a LG Google Nexus 5), indicating that the touch screen of the smartphone used in this study did not function optimally for the take-off technique. A possible solution for this problem is to block all movements for a few ms at the start of detecting a finger movement. In this way, the mark will not move when the device detects a very small finger movement due to users lifting their finger. However, this might reduce the technique's responsiveness. The effects of the suboptimal touch screen were much smaller (and therefore not noticed by participants) for shift, because of its increased control display ratio.

Two participants indicated that it was annoying to zoom out with zoom-pointing in order to press the submit button. This is an inherent problem of zoom-pointing. Zooming in on one part of the interface (the target in the image) reduces the accessibility of other interface elements (the submit button). This problem could be avoided by always displaying the submit button on the screen, irrespective of the zoom factor. This solution would be appropriate for the simple interface used in this experiment, but it becomes infeasible as the number of interface elements grows. Furthermore, most participants did not consider zooming out to be a problem.

A problem with shift was that the user's finger sometimes occluded the magnifier. This problem did not occur in the original shift [14], since this implementation always places the magnifier above the user's finger. However, placing the magnifier above the finger is impossible when the target is too close to the top edge of the screen, which occurs especially frequently on devices with a small screen. A related problem, which was only mentioned by one participant, was that the magnifier sometimes jumped to another location. This occurred when the target was near the center of the screen, i.e. the threshold for determining whether the magnifier should be drawn on the left or right half of the screen. When the handle is dragged across this threshold, the magnifier jumps from one half of the screen to the other. A possible solution for both of these problems is to ask users whether they are right- or left-handed and use this information to determine the optimal location for the magnifier.

The median intuitiveness rating was positive for take-off and shift, and strongly positive for land-on and zoom-pointing. This indicates that techniques in which the target has to be touched directly are perceived as more intuitive than techniques in which a mark can be dragged over the target.

Several participants complained about shift's complexity. They found that there were too many things on the screen, that it was unclear when the target had been marked, or that the difference between the mark and the target was unclear. A high number of participants asked for extra instructions for shift. Participants sometimes tried to drag the magnifier instead of the handle and several participants used shift in an inefficient way (with the first tap being far from the target). Although the median intuitiveness rating for shift was positive, these comments and observations indicate that shift is not a very intuitive technique.

Based on our results, we conclude that take-off and zoom-pointing provide the best trade-off between accuracy, efficiency, ease-of-use, and intuitiveness. The differences between these techniques were small. The number of attempts was lower for take-off, but the subjective ratings were slightly more favorable for zoom-pointing. Shift was the most accurate technique, but it was hampered by its high complexity and low intuitiveness. Land-on was the fastest technique, but it was also the least accurate. We believe that both take-off and zoom-pointing are viable interaction techniques for the I2Vote system.

Participants using the smartphone were more accurate than participants using the tablet with the land-on, take-off and zoom-pointing techniques. We believe that this was caused by the larger target size on the tablet. When the target is larger, there is more overlap between the mark and the target, making the mark seem more accurate. Marks with an equal distance from their center to the target's center therefore falsely appear to be more accurate on the tablet than on the smartphone. Participants using the tablet were therefore more easily satisfied with the placement of their marks than participants using the smartphone, causing them to spend less time and attempts on improving the accuracy of their marks. This is supported by the fact that the median time and number of attempts needed to mark the target was lower for participants using the tablet than for participants using the smartphone.

The lack of difference in accuracy between the smartphone and the tablet when using shift can be explained by the fact that shift allows participants to drag a magnified mark over a magnified target, thereby allowing them to precisely align the border of the mark with the border of the target.

The subjective accuracy ratings corresponded well to the objective accuracy data (the actual deviation from the target). Several studies [24–26] found a discrepancy between subjective and objective measures of efficiency, but such a discrepancy did not occur for accuracy in our study. Bailey [24] explained this discrepancy by a tendency of users to integrate their preferences into their judgments of the efficiency of a system. The fact that our subjective accuracy data corresponded well to the objective data, but not to the other subjective data, indicates that such a bias does not occur when users rate accuracy.

Our results are consistent with those of Potter et al. [15], who found that take-off was more accurate, but slower than land-on, and Vogel and Baudisch [14], who found that for small targets, both take-off and shift were more accurate, but slower than land-on. Vogel and Baudisch did not find a difference in accuracy between the original implementation of shift and take-off. Our study showed that the high-precision version of shift these authors proposed, with a magnification of the target and a lowered control display ratio, does results in a higher accuracy for shift compared to take-off.

Our results are inconsistent with those of Albinsson and Zhai [16], who found that zoom-pointing was both faster and more accurate than take-off. The difference in speed results might be due to the different zooming techniques that were used (bounding box zoom in Albinsson and Zhai's study, pinch to zoom in our study) and the fact that participants had to zoom out in order to press the submit button in our study. The difference in accuracy results might be due to the different devices that were used. The devices used in our study likely have higher quality touch screens than the device used in Albinsson and Zhai's study, which is already eleven years old. Since take-off is more dependent on touch screen quality than zoom-pointing, because of the drag and release action it requires, accuracy increases more for take-off than for zoom-pointing when touch screen quality is increased. Also, the difference in accuracy between take-off and zoom-pointing in Albinsson and Zhai's study was exacerbated by the fact that they used a binary measure of accuracy (i.e. target is hit or missed). A small difference in deviation from the target could make the difference between a hit and a miss. To translate our results to their measure of accuracy: the median deviation from the target's center was well below the radius of the target (0.3 mm on the smartphone and 0.45 mm on the tablet) for both take-off and zoom-pointing. Therefore, the median hit rate for both techniques was 100%.

A limitation of this study, and of any touch-based interaction study (e.g. [27–29], is that the results depend on the specific devices that are used. Results may not be generalizable to devices with difference specifications (e.g. different screen, processor speed, or native gestures). We believe that the devices used in our study are a good representation of the smartphones and tablets that most people currently own, and that our results can therefore be generalized to all smartphones and tablets that might be used for an ARS in a BYOD environment. The only relevant device that was not covered by our study is the laptop, but since marking a region on an image using a laptop is a trivial task, no special interaction technique is required on this device.

The results of this study are not limited to the I2Vote system, but can be generalized to any touch-based system that allows users to accurately mark targets on images, without a priori knowledge of the user's intended target. For example, the techniques could be used in a touch-based medical image viewer to allow users to perform accurate diameter measurements, or in an ARS for military training to allow users to mark enemy targets on a satellite image.

Conclusion

Of the four interaction techniques evaluated in this study, take-off and zoom-pointing provided the best trade-off between accuracy, efficiency, ease-of-use, and intuitiveness. The differences between these techniques were small. The number of attempts was lower for take-off, but the subjective ratings were slightly more favorable for zoom-pointing. Shift was the most accurate technique, but it was hampered by its high complexity and low intuitiveness. Land-on was the fastest technique, but it was also the least accurate. Based on these results, we conclude that both take-off and zoom-pointing are viable interaction techniques for the I2Vote system.

References

[1] L. Greer, P.J. Heaney, Real-time analysis of student comprehension: an assessment of electronic student response technology in an introductory earth science course, J. Geosci. Educ. 52 (2004) 345–351.

[2] R.W. Preszler, A. Dawe, C.B. Shuster, M. Shuster, Assessment of the effects of student response systems on student learning and attitudes over a broad range of biology courses, CBE Life Sci. Educ. 6 (2007) 29–41.

[3] J. Cain, E.P. Black, J. Rohr, An audience response system strategy to improve student motivation, attention, and feedback., Am. J. Pharm. Educ. 73 (2009) 1–7.

[4] C. Fortner-Wood, L. Armistead, A. Marchand, F.B. Morris, The effects of student response systems on student learning and attitudes in undergraduate psychology courses, Teach. Psychol. 40 (2012) 26–30.

[5] R. Latessa, D. Mouw, Use of an audience response system to augment interactive learning, Fam. Med. 37 (2005) 12–14.

[6] R.G. Miller, B.H. Ashar, K.J. Getz, Evaluation of an audience response system for the continuing education of health professionals, J. Contin. Educ. Health Prof. 23 (2003) 109–115.

[7] K. Siau, H. Sheng, F.F. Nah, Use of a classroom response system to enhance classroom interactivity, IEEE Trans. Educ. 49 (2006) 398–403.

[8] A. Pradhan, D. Sparano, C. V Ananth, The influence of an audience response system on knowledge retention: An application to resident education, Am. J. Obstet. Gynecol. 193 (2005) 1827–1830.

[9] E.I. Rubio, M.J. Bassignani, M.A. White, W.E. Brant, Effect of an audience response system on resident learning and retention of lecture material, Am. J. Roentgenol. 190 (2008) W319–W322.

[10] P.M.A. van Ooijen, A. Broekema, M. Oudkerk, Design and implementation of I2Vote - an interactive image-based voting system using windows mobile devices, Int. J. Med. Inform. 80 (2011) 562–569.

[11] eClicker, http://eclicker.com. Accessed April 3, 2014.

[12] iResponse, http://iresponseapp.com. Accessed April 3, 2014.

[13] Socrative, http://socrative.com. Accessed April 3, 2014.

[14] D. Vogel, P. Baudisch, Shift: a technique for operating pen-based interfaces using touch, in: Proc. SIGCHI Conf. Hum. Factors Comput. Syst., ACM Press, San Jose, CA, 2007: pp. 657–666.

[15] R.L. Potter, L.J. Weldon, B. Shneiderman, Improving the accuracy of touch screens: an experimental evaluation of three strategies, in: Proceeding SIGCHI Conf. Hum. Factors Comput. Syst., ACM Press, Washington, D.C., 1988: pp. 27–32.

[16] P. Albinsson, S. Zhai, High precision touch screen interaction, in: Proc. SIGCHI Conf. Hum. Factors Comput. Syst., ACM Press, Ft. Lauderdale, FL, 2003: pp. 105–112.

[17] P. Baudisch, G. Chu, Back-of-device interaction allows creating very small touch devices, in: Proc. SIGCHI Conf. Hum. Factors Comput. Syst., ACM Press, Boston, MA, 2009: pp. 1923–1932.

[18] J.O. Wobbrock, B.A. Myers, H.H. Aung, The performance of hand postures in front- and back-of-device interaction for mobile computing, Int. J. Hum. Comput. Stud. 66 (2008) 857–875.

[19] H. Benko, A.D. Wilson, P. Baudisch, Precise selection techniques for multi-touch screens, in: Proc. SIGCHI Conf. Hum. Factors Comput. Syst., ACM Press, Montreal, Canada, 2006: pp. 1263–1272.

[20] A. Sears, B. Shneiderman, High precision touchscreens: design strategies and comparisons with a mouse, Int. J. Man. Mach. Stud. 34 (1991) 593–613.

[21] T. Grossman, R. Balakrishnan, The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area, in: Proc. SIGCHI Conf. Hum. Factors Comput. Syst., ACM Press, Portland, OR, 2005: pp. 281–290.

[22] M. Baldauf, P. Fröhlich, Snap Target: investigating an sssistance technique for mobile magic lens interaction with large displays, Int. J. Hum. Comput. Interact. 30 (2014) 446–458.

[23] J. V. Bradley, Complete counterbalancing of immediate sequential effects in a Latin square design, J. Am. Stat. Assoc. 53 (1958) 525–528.

[24] R.W. Bailey, Performance vs. preference, in: Proc. Hum. Factors Ergon. Soc. 37th Annu. Meet., Human Factors and Ergonomics Society, Seattle, WA, 1993: pp. 282–286.

[25] G. V. Kissel, The effect of computer experience on subjective and objective software usability measures, in: Proceeding SIGCHI Conf. Hum. Factors Comput. Syst., ACM Press, Denver, CO, 1995: pp. 284–285.

[26] W. Jorritsma, F. Cnossen, P.M.A. van Ooijen, Merits of usability testing for PACS selection, Int. J. Med. Inform. 83 (2014) 27–36.

[27] H. Hwangbo, S.H. Yoon, B.S. Jin, Y.S. Han, Y.G. Ji, A study of pointing performance of elderly users on smartphones, Int. J. Hum. Comput. Interact. 29 (2013) 604–618.

[28] M. Romano, L. Paolino, G. Tortora, G. Vitiello, The Tap and Slide keyboard: a new interaction method for mobile device text entry, Int. J. Hum. Comput. Interact. 30 (2014) 935–945.

[29] H. Tu, X. Ren, F. Tian, F. Wang, Evaluation of flick and ring scrolling on touch-based smartphones, Int. J. Hum. Comput. Interact. 30 (2014) 643–653.