Comparing Hamburger Menu With Swipe Gestures for Usage in a Page View

Lukas Galke
Kiel University, Germany
lga@informatik.uni-kiel.de

Steffen Goos
stg@informatik.uni-kiel.de

Felix Paur
stu95611@mail.uni-kiel.de

Florian Mai
stu96542@mail.uni-kiel.de

ABSTRACT

For several years, the swipe gesture has been a very-well established way to navigate among adjacent pages on websites as well as in native apps on a mobile device. The users appreciate its ease and its comfort of use. However, it can only be applied in scenarios where the content is provided in a sequential way, e.g. in a picture gallery, in an e-book reader, or when filling in a form. In this paper, we examine the swipe gesture's suitability for the case when the target page is not adjacent. This use case appears for example when an exam is taken through a mobile device. Although typically the user progresses through the exam's questions sequentially, it is a common case that they want to revise their answers later due to a change of mind or a misunderstanding of the question, for example. We conjecture that navigating back using swipe gestures is unsatisfactory and time-costly as the distance in pages increases. In order to research this question, we conduct a study in which we compare the use of swipe gestures to the use of the also well-established hamburger menu. We design tasks that capture the use cases described above. We alter the distance in pages that has to be covered in order to complete the task. For each task, we measure effectiveness, efficiency, and satisfaction. We introduce a definition of effectiveness that is based on whether the user made unnecessary interactions. We assess the user satisfaction through a questionnaire. For efficiency, we take the speed of navigation into account, as well. The findings indicate that the burger menu is better for distances of eight pages and more, but the swipe gestures are much more efficient when the distance is about two pages. The user satisfaction and effectiveness do not significantly favor either of the navigation methods for small distances. However, the findings can be of practical use for UI designers in the future.

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

Keywords

mobile application, swipe, hamburger menu, learning application

1. INTRODUCTION

Smartphones have become very established for the masses in recent years and the popularity of tablets is still rising. As users become more and more familiar with such devices, it seems natural that businesses and institutions from an educational context would like to use these for the testing of their examinees through a testing application (app). However, as for any kind of application, a practical barrier is the usability. For several years, the swipe gesture has been a

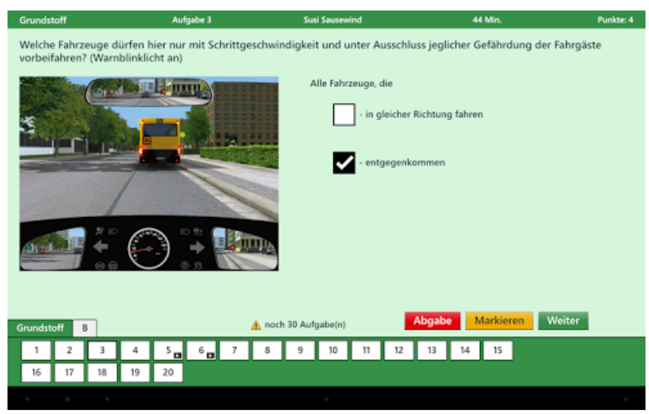


Figure 1: Mobile application 'Fahren Lernen Max'

very well-established way to navigate among adjacent pages in a page view on websites as well as in native applications on a mobile device. The users appreciate its ease and comfort of use. However, it can only be applied in scenarios where the content is provided in a sequential way, e.g. in a picture gallery, in an e-book reader, or when filling in a form. In this paper, we examine the swipe gesture's suitability for navigating through a set of questions in the scenario of an exam. Although typically the user progresses through the questions sequentially, it is a common case that they want to revise their answers later due to a change of mind or a misunderstanding of the question, for example. To this end, they often need to browse several pages back or forth. A case that applies very well to this scenario is the theory test that

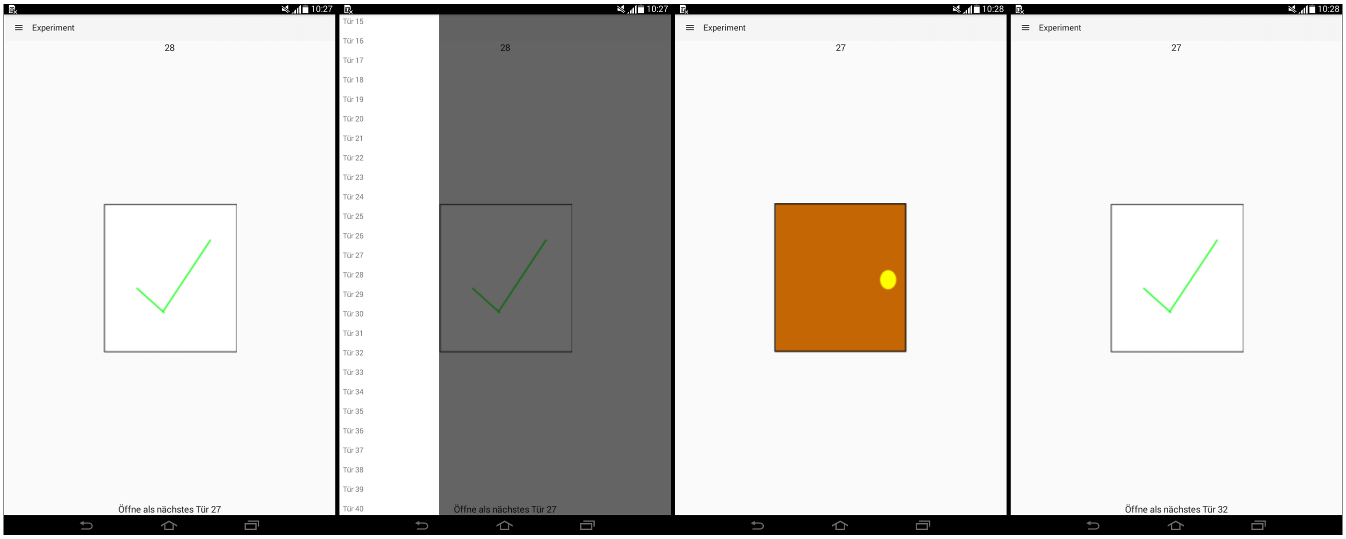
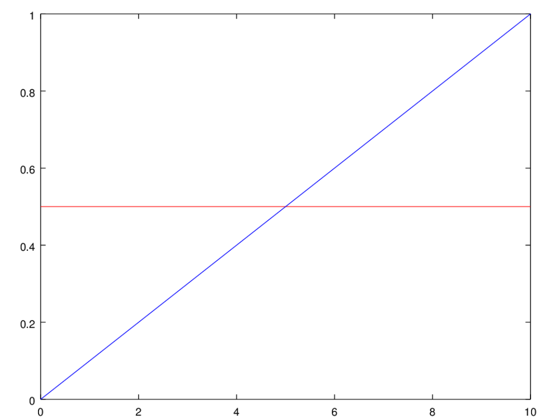


Figure 2: Hamburger navigation sequence

has to be passed in order to obtain a driving license in Germany. It consists of a relatively large number of questions and it's common practice to prepare for this examination by taking several exemplary tests. For example, consider Figure 1 which displays a screenshot of the mobile application 'Fahren Lernen Max'. It shows the typical environment that is also used for the actual exam. The questions are ordered sequentially, but the examinees may answer them in any order. During examinations, it is common that some questions are skipped and returned to at some later point. In the example of this application, navigation between questions can be done in any direction via menu at the bottom of the screen or only forward via the 'Weiter'-button.

We conjecture that using only swipe gestures would be unsatisfactory in this application, since it may be not uncommon to skip one of the first questions and return to it only after the next 20 questions have been solved. Hence, swiping back all the way would be time-costly which can put some extra pressure on the examinee. From the perspective of the application designer, a better design choice in this case would be the use of the also well-established hamburger menu. This kind of navigation is mostly represented by three horizontally lines (the hamburger) on the left or right upper corner of the screen. By touching the hamburger button, a menu appears at the side of the screen. This solution, however, might be inconvenient when the examinee only needs to browse two pages back, for example, and here swipe gestures may come in handier. In order to research this question, we conduct a study in which we compare the use of swipe gestures to the use of the hamburger menu. Our goal is to find a generalizable statement about what navigation methods make sense in which scenario. More particularly, we analyze how the methods perform when the distance in pages is altered. However, it is hard to artificially construct an environment similar to exams and ensure at the same time that the examinees navigate through the app in such a way that the data are comparable. Therefore, our evaluation approach is detached from this specific context. Instead, we design a game that requires the participant to navigate through a page view in a predefined way. We analyze

Figure 3: Expected behavior of dependent variable (y-axis) depending on the distance (x-axis) when using a menu (green line) and when using swipe gestures (blue line)



the common three dependent variables efficiency, satisfaction, and effectiveness, while varying the distance d. We use a definition of effectiveness that measures whether participants made unnecessary interactions. For the efficiency, we additionally consider the required time, and the satisfaction is assessed through a questionnaire. It seems reasonable to conjecture that for very small distances swipe gestures are superior to the hamburger menu. Because the number of swipes increases proportionally with the distance, the time required can be expected to increase proportionally as well. However, when the distance increases, the hamburger menu may have an advantage because one can directly select the desired page. Corresponding to this conjecture we examine the distances $d = 2, 5, 8, 11, 14$ and expect the relation between distance and the independent variables as indicated in Figure 3. The figure presumes an intercept point between the lines to exist. Corresponding to our conjecture that the

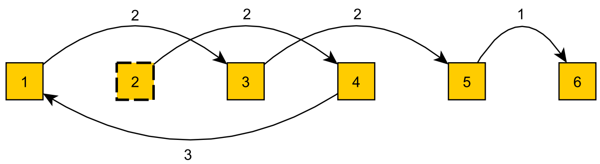


Figure 4: Example of a distance-two-jump

swipe gestures are superior for very small distances and the hamburger menu's superiority becomes apparent for large distances, we choose $d = 2, 5, 8, 11, 14$ as values for d. Formally, for each d, we formulate the H_0 hypotheses that there is no difference between swipe gesture navigation and hamburger menu navigation when the browsing distance is d. Our main results are the following:

- Burger menu navigation is better with respect to all applied metrics when the distance is large (8 and higher)
- Swipe navigation is significantly much more efficient than burger menu for small distances (2)
- When the distance is small, neither of the methods performs significantly better than the respective other in terms of satisfaction and effectiveness

The results of our research give important insight into the usability of different navigation methods, potentially leading to guidelines for the design of future apps developed for the purpose of exams.

2. APPARATUS

In order to simulate the scenarios of jumping from one question to another, we developed a series of treasure hunt like tasks. Here, the goal is to find the treasure which is hidden behind a door. A door is represented by a page. The application was developed for Android devices. For our data collection, we have used two Samsung Galaxy Note 10.1 tablets running Android 4.4.2.

First of all an initial page containing the identification number of the participant is shown so the task can be linked to the demographic questionnaire at the very end. Starting at the first page, an instruction displayed on the bottom of the screen guides the participant to the next one by exactly telling them the number of the next door. Now the participant has the possibility to navigate to the next door by swiping left or right, or using the hamburger menu (Figure 2) on the upper left corner. The navigation method depends on the group in which they are located. When the participant has opened the respective door, the next instruction is displayed. The distance d to cover from one door to another is always about the same ($d \pm 1$, *see Section 3.3*) within the same task. However, the distances differ for different tasks. Consider the example in Figure 4 which has distance of two. The arrows indicate the door to open in the next step and the arrow captions show the jump distance. Hence, the sequence of doors to open in order to reach the treasure is 2, 4, 1, 3, 5, 6. The tasks cannot be failed to be completed because there is no way to open the doors in the wrong order. The participant has to complete a training set and five task sets. Every set consists of five jumping instructions. After each set the participant has to answer two questions (see

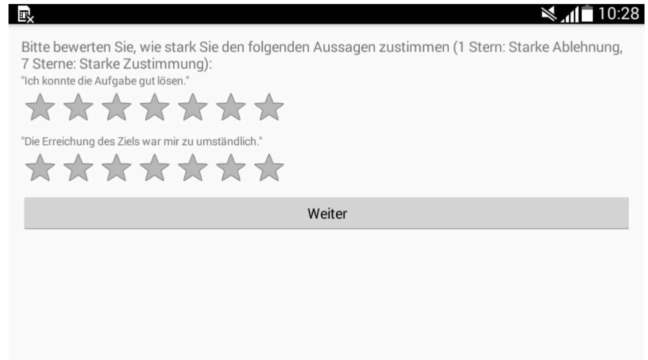


Figure 5: Between questionnaire

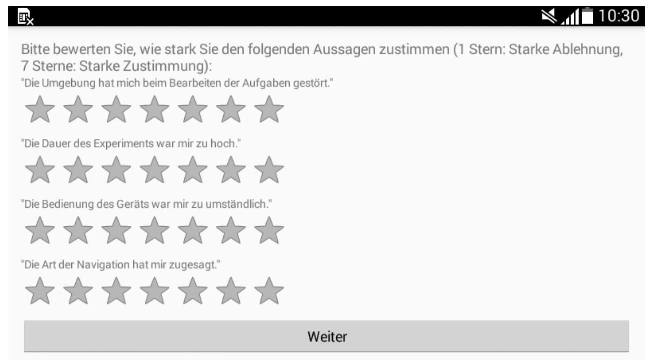


Figure 6: Final questionnaire

Section 3.4) and at the very end there will be a questionnaire of four questions.

3. EVALUATION DESIGN

3.1 Procedure

For our study, we decided for a between-group design for two reasons. First, when solving the tasks with the second type, the test persons might rate their satisfaction more drastically, as they are biased from having solved the tasks with the respective other navigation method. Secondly, in order to avoid an extreme learning effect, the tasks for swipe gesture navigation and menu navigation would have to be distinct if a within-group design were used. This should be avoided to ensure comparability. The participants are assigned to each group randomly (at uniform) by employing the random number generator of the respective android device.

The entire flow for one participant is as follows: First, the participant is explained the purpose of the study (evaluation of navigation methods on mobile (touch)-devices) and the type of task to solve (navigation through app). After completing an informed consent form, a unique identifier is generated and handed over to the participant alongside contact information. Since often students stay in the dining hall in groups, the participant is separated from the group in order to avoid distraction. Before the actual experiment starts, the participant completes a training task during which the goal, the different layout items, and the navigation are explained. Before finishing the questionnaire for the training task, the participant is informed that the experiment will start and is asked to focus on the tasks. When they have finished, the last step for them is to provide demographic data age, gender, profession (including field of study), and whether they own a smartphone with an Android operating system, a different operating system, or no smartphone.

3.2 Participants

Since our target scenario is an application used during examinations, a typical user of such an application is a university student. Hence, the participants were acquired in a field study conducted in an academic context, namely the dining hall of Christian-Albrechts-University in Kiel. The participants are representative for all students as the statistics show: In total, there are $N = 50$ participants. All of them are students and the most common courses are: economic computer science (12 %), agricultural sciences (10 %), mathematics (8 %), psychology (8 %) and economics (6 %). For the entire distribution consider table 6. The participants split up into 42 % male and 58 % female. The average age is 24.1 (SD:2.9), with values between 20 and 35. All but one of the participants own a smartphone, 75.51 % of which run Android. This indicates that we may assume good knowledge of how to handle the device used in our experiment. The participants signed an informed consent form before the start of the experiment. After finishing the experiment, they were rewarded with sweets or fruits.

3.3 Tasks

During task creation, we faced several design issues: In order to avoid learning effects and for the results to generalize better, the sequence should require both forward- and backward-jumps. This raises the problem of collisions, that is, a door is visited twice, which should also be avoided due to a learning effect. Hence, it is not possible to always have the exact same distance from one door to the next in the sequence without collisions, but instead they have to differ slightly. However, it is always possible to create a non-colliding sequence of length five by using distances d three times and distances $d - 1$ and $d + 1$ once each. Hence, on average we have a distance of d. Furthermore, the distances being limited to two neighbored values, it is reasonable to assume negligible impact on the experiments outcomes in terms of user-satisfaction and time.

To further reduce any bias of learning effects the concrete permutation of tasks is generated by subsequently drawing (uniformly at random, without returning) from the remaining tasks.

Every task consist of a sequence of five doors that had to be opened in the right order. In the beginning the participant would be told which door to open first. After each door there would be a message which door to open next.

For the concrete tasks, please see Table 1. It shows the task's identifier, average distance, and the sequence in which the doors have to be opened as j1 through j5. Every task starts at door 25.

3.4 Measurements

We aim to measure the common three dependent variables efficiency, satisfaction, and effectiveness of each of the two navigation methods.

Effectiveness.

Table 1: The tasks to be conducted by a participant

task ID	d	j1	j2	j3	j4	j5
1	2	27	26	24	21	23
2	5	21	26	31	37	32
3	8	34	42	35	27	19
4	11	14	3	13	1	12
5	14	39	24	10	23	9

During the experiment, the application tracks how often the page is changed by the user, that is by either swiping or selecting a different page in the menu. Given some task that requires 5 jumps, where the distances to cover in the jumps are given as $d_i, 1 \leq i \leq 5$, let $n_{interactions_i}$ be the number of page changes by the user for jump i. The optimal number n_{opt_i} of page changes required equals 1 for the burger menu navigation and d_1 for the swipe gesture navigation. As a measure of effectiveness, we measure in how many of the jumps unnecessary page changes were made, that is, the effectiveness for a task equals $\frac{|\{i \in \{1,\ldots,5\} | n_{interactions_i} = n_{opt_i}\}|}{5}$.

Efficiency.

For the efficiency, we calculate the effectiveness of a task divided by the time required to complete the task in seconds. We start measuring the time at the point when the first instruction is displayed and stop when the last door is opened.

Satisfaction.

In order to assess the user satisfaction depending on the distance, after each task, the user is asked to estimate their agreement on two statements concerning the task. The statements are the following.

1. 'I was able to solve the task well.' (in German: 'Ich konnte die Aufgabe gut lösen.')
2. 'Achieving the goal was too intricate.' (in German: 'Die Erreichung des Ziels war mir zu umständlich.')

The rating scale follows the common 7-Likert-scale from strongly disagree (1) to strongly agree (7). Sauro et al. [4] pointed out that the reliability of a questionnaire increases with increasing range of scale. They also pointed out that the effect would be very small after having 11 possibilities to choose. Since our mobile device has a small screen the 7 seemed to be a good compromise between space and choice possibilities.
The first statement will be used to quantify whether the participants have understood the task and how to solve it. The second statement will be used to quantify the participant's satisfaction with the navigation method. After all tasks were completed, the participant also rates these statements that aim to assess the overall experience throughout the experiment.

1. 'The environment during working on the tasks was disturbing.' (German: 'Die Umgebung hat mich beim Bearbeiten der Aufgaben gestört.')
2. 'The duration of the experiment was too high.' (German: 'Die Dauer des Experiments war mir zu hoch.')

3. ‘Handling the device was intricate.’ (German: ‘Die Bedienung des Geräts war mir zu umständlich.’)

4. ‘The type of navigation was appealing to me.’ (German: ‘Die Art der Navigation hat mir zugesagt.’)

Question (1) and (2) serve the purposes to check whether the quality the experiment’s results is impaired due to the lack of a laboratory environment or due to participants’ fatigue, respectively. Question (3), as well as Question (4), directly assesses the overall satisfaction with the navigation method. Since obtaining meaningful results on user satisfaction might be difficult without revealing too much about the experiment design to the participants, we formulate the question for satisfaction both in a negative and in a positive fashion.

4. EVALUATION RESULTS

In this section, we formalize our hypotheses, explain our procedure and notations for testing the hypotheses, conduct a retrospective power analysis and finally, present both descriptive statistics and test statistics for each distance, navigation method, and dependent variable. In the end, we present further statistics related to our experiment.

4.1 Hypotheses

Let d be the distance in pages to be covered when navigating in a sequence of pages. We formulate the following H_0 hypotheses, using the notions of effectiveness, user satisfaction and efficiency presented in Section 3.4:

1. Null hypothesis $H_0^{(efficiency)}$ for efficiency: There is no difference between swipe gestures and hamburger menu in the time required to navigate over a distance of d pages.

2. Null hypothesis $H_0^{(effectiveness)}$ for effectiveness: There is no difference in effectiveness between swipe gestures and hamburger menu when navigating over a distance of d pages.

3. Null hypothesis $H_0^{(satisfaction)}$ for user satisfaction: There is no difference in user satisfaction between swipe gestures and hamburger menu when navigating over a distance of d pages.

In fact, we test these hypotheses for different values of $d = 2, 5, 8, 11, 14$, hence, we examine two independent variables: navigation method and distance.

4.2 Testing Procedure

Our main focus in this research is on comparing the two navigation methods when the distance is equal. The results for all independent variables as introduced in Chapter 3.4 are shown in Table 2. Recall that the satisfaction is quantified through task question (2). For additional results please see the Appendix. In order to test for significant differences in any of the independent variables, we first apply a Shapiro–Wilk test to test for normality and apply Welch’s t-test (marked with *) if the Shapiro-Wilk-test does not reject for either of the groups. In the case of a t-test, we report the following data in this order: t value, significance level p, Cohen’s d as a measure of effect size, and the degrees of freedom df. In the case where the test for normality fails, we apply a Mann-Whitney-U test (marked with $\sim$), in which case z value, significance level p, and the rank-biserial correlation, denoted as r, are provided.

For all our computational analyses, we used the SCIPY.STATS package for Python.

Table 2: Comparing burger and swipe for same d

d	effectiveness	efficiency	satisfaction
2	$z = 289, p = 0.22, r = .06$	$t = -6.70, p < .001, d = -1.91, df = 38.78$	$z = 247.5, p = .08, r = .20$
5	$z = 265, p = 0.14, r = .14$	$z = 263, p = .192, r = .15$	$z = 225.5, p < .03, r = .37$
8	$z = 226, p < 0.03, r = .27$	$t = 2.47, p < .02, d = 0.70, df = 37.32$	$z = 230, p < .05, r = .30$
11	$z = 248, p < 0.05, r = .20$	$t = 6.22, p < .001, d = 1.77, df = 36.91$	$z = 192.5, p < .01, r = .38$
14	$z = 211.5, p < 0.01, r = .31$	$t = 7.97, p < .001, d = 2.27, df = 39.44$	$z = 190.5, p < .01, r = .38$

4.3 Power Analysis

As already suggested in the introduction, preliminary to conducting the study, we expected to find strong differences between swipe navigation and menu navigation, that is, we expected to find significances with an effect size of $d \geq .8$. Furthermore, we assumed normally distributed data with about equal variances, so using Student’s t-tests would be adequate. Hence, in order to achieve statistical power of .8 in a two-sided two sample Student’s t-test, the number of participants should be 25 per group. For this reason we aimed for 50 participants in total. In 9 out of 10 cases, we found the efficiency values per task to be normally distributed and in 4 out of 5 cases of the tasks’ direct comparison, we can report large effect sizes (> 0.5). However, because we didn’t find the variances to always be equal and because it is discouraged in the literature to choose between Student’s and Welch’s t-test per case [3], we decided to constantly use Welch’s t-test.

4.4 Results for Dependent Variables

4.4.1 *Effectiveness*

Overall, the average effectiveness of the burger menu and swipe gestures are .98 (SD:0.063) and .93 (SD:0.106), respectively. For burger menu navigation, Table 3 shows that the error rates were slightly lower for task 1 and task 5 than for the others. For the swipe navigation, the error rate of task 1 is much lower than the error rates of all the other tasks. While most of these have an error rate below 10 %, task 3 stands out negatively by exceeding this value. When comparing both navigation methods to each other, burger menu navigation yields a higher effectiveness than swipe in any case. The differences are significant for tasks 3, 4, and 5, but not for tasks 1 and 2.

Table 3: Mean (standard deviation) of effectiveness

navigation	2	5	8	11	14
burger	.99(0.043)	.96(0.078)	.96(0.079)	.98(0.059)	0.99(0.043)
swipe	.98(0.063)	.94(0.095)	.89(0.139)	.94(0.092)	.92(0.113)

4.4.2 *Efficiency*

On the one hand, the box plot of efficiency by task for burger menu navigation (see Figure 7) reveals heterogeneous variances, no trend, negative skews for tasks 1 and 3, positive skews for tasks 4 and 5, and a single outlier at task 4 below the central value.

On the other hand, the box plot of efficiency by task for swipe navigation (see Figure 8) reveals heterogeneous vari-

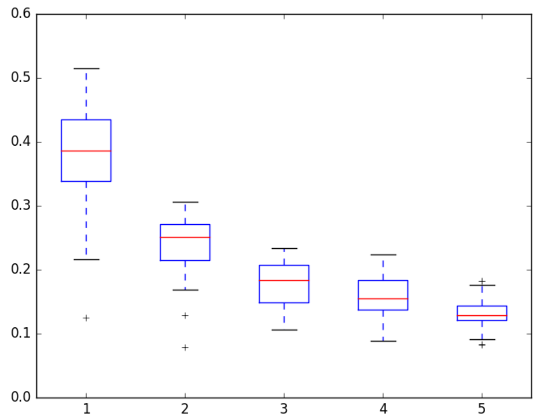


Figure 7: Efficiency by task with swipe navigation

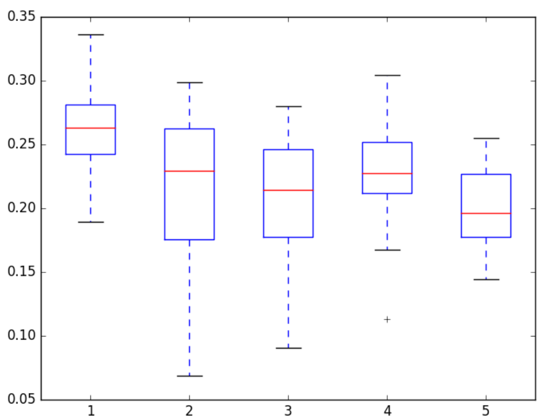


Figure 8: Efficiency by task with burger menu navigation

ances, a descending trend, negative skews for tasks 2 and 3, positive skews for tasks 4 and 5, and no skew for task 1, four outliers at tasks 1, 2, and 5 below the central value, and a single outlier at task 5 above the central value.

Averaged over all tasks, burger menu navigation (0.223 (SD:0.05)) is slightly more efficient than swipe navigation (0.216 (SD:0.1)). Looking at the development of the means depending on the tasks shown in Table 4, it can be seen that swipe is very efficient compared to burger for small distances (task 1 and 2), but also that the efficiency decreases rapidly. The burger's efficiency also decreases (almost) constantly as the distance increases, but at a smaller rate, so that it is more efficient than swipe in tasks 3, 4, and 5. Except for task 2, these reported differences between the navigation methods are significant (see Table 2). The effect sizes are mostly large and increase drastically from task 3 to 5.

4.4.3 *Satisfaction*

The average rating of task question 2, our main measurement for the user's satisfaction with the navigation method, is 1.89 (SD:1.809) for the burger menu group and 2.86 (SD:2.13) for the swipe gesture group, which indicates that the participants perceived the burger menu to be more convenient. This matches the results from final question (4) where the burger menu and swipe gestures yield an average rating of 5.14 (SD:1.67) and 4.5 (SD:1.453), respectively. See Table 5 for the statistics of individual tasks. While the satisfaction with swipe navigation has the tendency to increase as the distance increases, the satisfaction with the burger menu navigation does not show such a tendency. In any case, the burger menu generates higher satisfaction than swipe navigation. The differences are significant for all tasks except task 1, which is shown in Table 2. The effect size is below .4 in any case.

Table 4: Mean (standard deviation) of efficiency

navigation	2	5	8	11	14
burger	0.26(0.036)	0.22(0.062)	0.21(0.05)	0.22(0.045)	0.2(0.034)
swipe	0.37(0.083)	0.024(0.052)	0.18(0.037)	0.16(0.032)	0.13(0.027)

Table 5: Mean (standard deviation) of satisfaction

navigation	2	5	8	11	14
burger	1.96(2.01)	1.77(1.88)	2(1.75)	1.77(1.54)	1.96(1.99)
swipe	2.75(2.37)	2.68(2.21)	2.82(2.16)	2.86(1.88)	3.21(2.11)

4.5 Other Results

We look into the results of the questionnaires that are designed to check the validity of our study's design. Task question (1) was rated with 6.91 on average (SD:0.4) on average in the burger group and 6.84 (SD:0.46) on average in the swipe group. The lowest rankings were 4 and 5, respectively. Final question (1) was rated with on average 1.52 (SD:1.182) for swipe. In total, three participants rated it with four or more, one out of which rated it with 7. The overall mean of final question (2) was 1.92 (SD:1.496), 9 of the participants rated it with 4 or higher. In final question (3) there is a noticeable difference between the rating for the burger menu group (1.22 (SD:0.429)) and the swipe group (1.75 (SD:1.378)). Testing with a Mann Whitney U test, the difference does not show significance, z = 255.5, p=0.10, d = 0.171. The highest rating was 7 and in total three participants rated with 4 or higher.

5. DISCUSSION

5.1 Main Findings

Effectiveness.

The way we defined effectiveness, burger menu navigation is never worse than swipe navigation, for large distances it is even significantly better. However, traditionally, effectiveness measures if users complete a task correctly. In our tasks, there is no right or wrong. Here, effectiveness rather measures how precisely a participant navigates through the set of pages. In an exam scenario as described in the introduction, a user rather cares about how fast he can move through between pages. Clearly, making a navigation error affects the experience for the user differently depending on the navigation method. While a miss-click in the burger menu always requires the user to perform twice as many interactions as in the optimal case to reach the goal, an accidental swiping mistake only requires one additional swipe to correct it. This may affect both the speed and the user satisfaction differently for each navigation method. These considerations make the results difficult to interprete. On top of that, the effect sizes are rather small which allows only vague practical implications from the results.

Efficiency.
The test results indicate clear superiority of swipe gestures in terms of efficiency when the distance is small. On the other hand, there is good evidence that this relation indeed flips as the distance increases, favoring burger menu navigation for distances larger than 8. This supports the hypothesized behavior. However, it remains unclear whether the burger menu's efficiency is independent of the distance. During the experiments we observed that, given the device's display size and the layout of the app, the menu displayed a range of 30 items. Hence, in task 5 where the average distance is 14, a participant is required to scroll within the menu by one item. This raises the question how the burger menu performs when the distances are even larger than the ones investigated in this study.

Satisfaction.
As expected, the participants perceived navigation with burger menu significantly less intricate when the distance grows. On the other hand, the effect size is rather low and a significant superiority of swipe navigation could not be shown for very small distances. Quite the contrary, on average the burger menu navigation was perceived less intricate even for these very small distances. This stands in contrast to oral and written statements made by some participants in the burger group. They were missing swipe gestures in order to simplify navigation further. One reason for this contradiction could be that the evaluation method used in this study is not appropriate to capture the user satisfaction. On the other hand, since for small distances the differences were not significant, there is a reasonable chance that in a repeat of the experiment this contradiction may not appear again.

5.2 Study Design Validation

The results for task question (1) suggest, that the participants did not feel overstrained by the complexity of the tasks. The results of the person who indicated large distraction from the environment did not stand out in particular. However, 18% felt the study was relatively long (rated 4 or more at final question (2)). Hence, there is a chance that these participants became tired or bored at some point during the experiment. The whole process required up to ten minutes. The study could only have been shortened by reducing the number of tasks, but without a sufficient range of distances, finding the key results for navigation in a page view would be very difficult. During answering the final questionnaire, a participant remarked his confusion about the difference in meaning between question (3) and (4). The relatively large difference between the swipe and the burger group may also be understood as indication that participants grasped both questions in a similar way. However, as the difference is not significant, we will assume that the question was understood correctly by most participants and did not lead to too much confusion. Overall, we think that we were able to eliminate some of the major potential threats to the study's validity.

5.3 Summary

Overall, in terms of efficiency, the presumed intercept point before which swipe navigation and after which burger menu navigation is better was indeed found at distance 5. For distances 8 or higher, burger menu navigation outperforms swipe in all applied measures. Before that point (distance 2) the superiority of swipe is backed by the results of satisfaction and effectiveness in so far that burger navigation is not significantly better. It should further be taken into account that these two measurements are more difficult to assess than efficiency which mainly depends on the time required to complete the task. Future work should take a closer look at approaches to capture satisfaction. In order to understand the value of effectiveness better, a systematical investigation of how users perceive a navigation error depending on the navigation method would also be interesting.

It should further be kept in mind that this study was designed for a very specific context that was discussed in the introduction. The navigation behavior that was required when conducting the tasks will not apply to the common case. However, as discussed, we believe that adequate application scenarios will become more frequent in the future. Therefore, as a practical implication of this work, we would like to give UI designers who plan to use a page view in scenarios similar to the discussed ones to consider the following recommendation cautiously:

- if the design of the application allows it, including both burger menu and swipe navigation methods is beneficiary
- if there are more than 5 pages, allowing for burger menu navigation or a similar navigation method is strongly recommended

6. RELATED WORK

During the last years the usage of mobile devices has increased largely. According to a study from the Pew Research Center the number of Americans that owned a tablet in 2015 was about 45%, which is roughly ten times more than in 2010 (4%). One reason for this development could be that touch gestures are far easier to comprehend. Therefore, mobile devices become available for every generation. While in general older adults are slower in completing tasks with desktop and mouse compared to younger adults, the use of touch gestures has reduced that gap. The use of a touch screen resulted in a significant movement time reduction by 35% for older adults. Younger adults still had a movement time reduction of 16% [1].

To further improve the user experience it is necessary to compare the different gestures as well as the menu elements, like we compare the swipe gesture with the hamburger menu given a specific task. We were able to detect differences between the two navigational methods regarding the time needed to fulfill a task and the user satisfaction given a specific task.

Diefenbach et al. [2] compared the usability of manipulative touch gestures to that of an established interaction style based on buttons and menus and indicated a negative influence of gestures when an application is used for the first time. They also stated that new gestures are learned in a shorter time span than expected, which might explain the movement time reduction pointed out by Findlater et al. We can confirm that our participants were as well fast in learning and using the gestures, but since our navigational methods are very simple one has to be careful with a generalization.

7. CONCLUSIONS

We have conducted a study that compares the effectiveness, efficiency, and user satisfaction of a hamburger menu navigation to a navigation with swipe gestures in a page view in applications on mobile devices. The main goal was to investigate how these navigation methods behave when the distance in pages that has to be covered is altered. To this end, we developed an application that requires the user to navigate through the page view in predefined paths. Here, the paths' lengths differ to accommodate for different distances. We introduce a definition of effectiveness that measures whether participants made unnecessary interactions. We measure the efficiency by tracking the time required to complete a task and assess the user satisfaction with a questionnaire.

The findings indicate that the burger menu is better when the distance is large (8 and higher) and that swipe gestures are much more efficient when the distance is small (2). However, the user satisfaction and effectiveness do not significantly favor either of the navigation methods in this case. Investigating the reasons for this will be work of the future.

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] L. Findlater, J. E. Froehlich, K. Fattal, J. O. Wobbrock, and T. Dastyar. Age-related differences in performance with touchscreens compared to traditional mouse input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 343–346. ACM, 2013.

[2] B. Heuwing, I. Köller, V. Schanz, and T. Mandl. Usability of gesture-based mobile applications for first-time use. *Mensch und Computer 2015–Proceedings*, 2015.

[3] D. Lakens. Always use welch's t-test instead of student's t-test. January 2015.

[4] J. Sauro and J. S. Dumas. Comparison of three one-question, post-task usability questionnaires. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1599–1608. ACM, 2009.

APPENDIX

A. SOME APPENDICES

The raw data and test-results as well as descriptive statistics on the tests presented in this paper as well as additional results (which compare different tasks within the same navigation method) can be found in the respective folders of the submission.

Table 6 shows the exact distribution of the participants among the fields of studies.

Table 6: Course distribution

Course	**Count**
Wirtschaftsinformatik	6
Agrarwissenschaften	5
Mathematik	4
Psychologie	4
Volkswirtschaftslehre	3
Ernährungs- und Verbraucherökonomie	2
Finanzmathematik	2
Wirtschaftsingenieur	2
Agribusiness	1
Betriebswirtschaftslehre	1
Biochemie	1
Biologie	1
Informatik	1
Mathematik / Chemie	1
Mathematik / Deutsch / Psychologie	1
Mathematik / Geologie	1
Mathematik / Geschichte	1
Mathematik / Informatik	1
Mathematik / Philosophie	1
Mathematik / Physik	1
Mathematik / Spanisch	1
Mathematik / Sport	1
Medizin	1
Musikwissenschaft / Philosophie	1
Physik	1
Politikwissenschaft / Ur- und Frühgeschichte	1
Rechtswissenschaften	1
Soziologie / Pädagogik	1
Wirtschaftswissenschaften Profil: Handelslehrer	1
N/A (1-Fa-Ma)	1