

Quality study of user activity using mobile device

Tap, double tap, flick gestures

Artur ARCIUCH, Antoni M. DONIGIEWICZ

Institute of Teleinformatics and Automation, Faculty of Cybernetics,
Military University of Technology,
ul. Gen. W. Urbanowicza 2, 00-908 Warszawa
{artur.arciuch}, antoni.donigiewicz@wat.edu.pl

ABSTRACT: The article presents the results of research on the quality of gestures performed by users using a mobile device. As a mobile device, the Nokia Lumia 800 smartphone was used. The results of the research concern the basic gestures of tap, double tap and flick, and include the execution time and the precision of the gesture. The results take into account the division of users into age groups and groups using and not using a smartphone every day. A comparison of the designated characteristics between groups is presented. Research on other gestures will be presented in the next paper.

KEYWORDS: tap, double tap and flick gestures, gestures entered (input) with a finger, mobile device, gesture execution time, gesture precision

1. Introduction

Among the quality studies conducted in the field of user activity, particular attention has recently been paid to quality studies of smartphone (and/or tablet) user activity. The reason for this is the common use of such devices, and the multitude of tasks performed with them. More and more frequently, this refers both to professional and personal tasks.

Knowing the characteristics of the quality of the gestures made by smartphone or table users constitutes the basis for the assessment of the performance of tasks comprising several consecutive, similar or different gestures. The most common interesting gesture quality indicators include gesture time, and the precision of indicating objects on a screen. In the context of design, it is necessary to learn whether the quality of gestures made by users belonging to various age groups differs significantly or if the users' ages are irrelevant in this scope. Another factor worth investigating is whether the quality of the gestures made by people using mobile devices on an everyday basis differs from the quality of gestures made by people who use such devices less frequently. The time during which the user's attention is concentrated on making a given gesture is also a relevant factor. At that time, the user is looking at the screen, so they are not focussing on other external objects. Such external objects may include field obstacles, pedestrian crossings, or roads while driving a car (or riding a bicycle). Irrational behaviour of users holding smartphones can often be observed.

This paper presents the results of studies aimed at quantifying the quality of tap, double tap and flick gestures made by users on a mobile device screen, i.e. a Nokia Lumia 800 smartphone. Taking into account ongoing studies and analyses [11], [13], the basic gesture quality characteristics used in this paper are: gesture duration and precision (distance between screen touching point and touched object centre) and gesture error probability. The gestures were made with fingers, and 60 people in the 16-66 age bracket were involved in the studies.

The gestures included in the studies are typical for smartphone users.

The "tap" gesture involves short, single touch of a screen (tapping its surface) and is generally used to start applications, select options or input data. This gesture ends when the finger or stylus is removed from the screen.

When making a "double tap" gesture, a user quickly touches the screen twice (taps the surface twice). This gesture is typically used to toggle between normal and zoomed-in views, e.g. of websites in a browser. This gesture is used relatively rarely. Similarly to the tap gesture, it ends when the finger or stylus is removed from the screen after the second tap.

The "flick" usually involves swiping (sliding) a finger across a screen in a selected direction. This gesture actually ends when the finger is removed from the screen, but the screen objects remain in motion. This gesture has multiple uses. Most commonly, it is used to move content or scroll lists, and can occur after the "pan" gesture.

The results of studies concerning other typical gestures (e.g. pan, pinch and stretch, etc.) will be presented in an upcoming paper.

2. Selected studies documented in literature

While discussing studies related to mobile devices and documented in the applicable literature, special attention has been paid to study results concerning gestures made on the screens of mobile devices, or directly on smartphones.

The majority of studies presented in the literature have been conducted using a stylus on a touch screen.

Studies concerning the selection and control of objects on a Pocket PC mobile computer screen are presented in publication [3]. These studies were aimed at determining the efficiency of younger and older adults while performing simple tasks using a stylus. The mobile device used for the purposes of these studies was a medium-sized Compaq iPAQ PocketPC 3950 computer operated by an Intel PXA250 processor. This device comes with a 57.6 mm × 76.8 mm screen with a resolution of 240 × 320 pixels. During the studies, the mobile computer was connected to a desktop using a USB cable.

Sixty people participated in the studies. The participants were presented with four types of tasks: tapping, touching, direct (straight) and circular swiping. The “tapping” tasks involved touching a green circle with a stylus, which resulted in the display of a red circle (target) to be touched. The target circle diameter was 16, 24 or 32 pixels. Circles (targets) were displayed within a distance three, four or five times larger in relation to the target circle diameter and at one of the eight angles in relation to the green circle’s location.

All studies concerning the above-mentioned tasks were preceded with training tests. The study concerning each individual comprised two blocks with 72 randomly arranged tasks each. Such factors as task performance time and precision were recorded. Based on the recorded results, it was possible to draw the following conclusions:

For the purpose of the “tap” task, such factors as the target dimensions and the user's age had a significant impact on task completion time and precision. The distance from the target had a significant impact on the completion time, which did not apply to precision. The task sequence had a significant influence on precision, which did not apply to completion time.

For the purpose of the “touch” task, such factors as target dimensions, the distance from the target and the location angle had a significant impact on task completion time, which did not apply to precision. However, such factors as the

age group and task sequence did not have any significant impact on task completion precision and time.

Details regarding straight and circular operation as well as the average completion time and precision for each task for each age group are presented in paper [3].

The assessment of the effectiveness of selecting objects using a stylus on a small touch screen is presented in paper [4]. The users' actions were assessed on the basis of experiments involving three different devices, i.e. a mobile phone, a palmtop and a tablet. The phone's screen diameter was 2.9", the palmtop's was 3.6" and tablet's was 10.2". All devices were placed in an upright position. The smaller devices (the phone and palmtop) were used in two positions, i.e. held in the hand and placed on a desk. Due to its size, the tablet was used only in a flat position, i.e. laid on a desk.

The test consisted of indicating two rectangular targets in a predefined sequence, i.e. the first target, the second target. After the second target was indicated, the next set of targets was displayed. The tests were run in two scenarios, i.e. errors are permissible and errors are impermissible.

The size of the first target (a black square) was constantly 6.0 mm, while the size of the second target (a white square) was randomly selected by the phone from the 2-3 mm range. The size of the second target on the palmtop or tablet screen was selected randomly from a range ensuring the same difficulty factor as for the phone.

The test results made it possible to formulate the following conclusions. The throughput of hand-held devices was significantly lower than for devices laid on the desk. The error rate (percentage) for a hand-held palmtop was much lower than for a palmtop laid on the desk. No significant difference in the error rate for a phone in the same conditions was determined. Details regarding device throughput and error rates for such testing conditions are provided in paper [4].

Studies concerning the three methods of data input (with a finger, stylus and mouse, as the performance reference point) are presented in paper [5]. The tests involved various types of basic user actions aimed at selecting targets by tapping, swiping and radial swiping.

A HP Compaq 2710p (screen diameter 12.1", screen resolution 1280 × 800 pixels) was used for test purposes. When a Logitech MX610 laser mouse was used, the screen was set in an upright position, however, with a stylus, the

screen was in a flat position. Data was input with a touch on an HP TouchSmart 600 PC (23" screen, resolution 1920 × 1080 pixels).

Eighteen people (including 6 women) participated in the tests (age bracket: 21-59). The participants were trained in the tasks involved in the test.

The efficiency of tapping and swiping was studied on the basis of a typical single-direction test [14].

The following test results are interesting. During the "tap" gesture test, the input method (finger, stylus, mouse) had a significant impact on the error rate. The significant impact of the difficulty factor on the error rate was determined. A low precision of finger indication resulted from the lack of feedback on the finger position right before touching the screen, and the finger size in relation to the object.

While analysing the test results for object indication time, it was determined that data input was fastest using a finger, and slowest using a mouse. Thus a significant impact of the input method and difficulty factor on the object indication time was detected.

As far as swiping is concerned, the error rate was lower than for tapping. No significant impact of the input method (finger, stylus, mouse) on the error rate was determined.

Details concerning input test results and test results for radial swiping are presented in paper [5].

Three experiments concerning the selection of objects are presented in paper [6]. Two experiments (one-dimensional and two-dimensional) are more interesting from the point of view of this topic, as the third one concerns typing using a touch keyboard. The paper itself concerns another issue, i.e. using the sum of two independent normal distributions to interpret the distribution of the final points input with a finger, on a screen.

A HTC NEXUS smartphone, running Android, was used in these tests. It came with a capacitive touch screen (size 48 × 80 mm, resolution 480 × 800 pixels). 12 people (including 3 women) participated in the tests (age bracket: 18-45).

The one-dimensional experiment consisted of generating a sound, displaying a grey, 6 mm wide, horizontal bar (the initial bar) and displaying a red, horizontal target bar (of varied width). Successful selection of the initial bar resulted in its colour changing to green along with the simultaneous

generation of an audio signal. Raising a finger after touching the target bar ended each task performed by a given person.

The test results demonstrated that the error rate significantly exceeded 4%, i.e. indicated very low test performance precision. The effective object width, i.e. the actual spread of touch points, was more than twice as large as the nominal target width. See Table 1 for average task completion times and error rates for the A and W value combinations.

Tab. 1. Average task completion times and error rates for one-dimensional experiment [6]

A [mm]	W [mm]	Average time [ms]	Error rate [%]
20	2.4	432	29
30	2.4	483	38
20	4.8	383	14
30	4.8	433	11
20	7.2	367	3
30	7.2	394	6

The two-dimensional experiment was similar. The difference consisted of the fact that circles of diameter W were the targets.

Similarly to the one-dimensional experiment, it was determined that the error rate significantly exceeded 4%. Average task performance times in the two-dimensional experiment were similar to the values determined during the one-dimensional experiment, however, the error rates were higher. See paper [6] for details.

An examination of the impact of the device size on performance, and the impact of the screen size on input quality for various tasks is presented in article [7].

30 people (including 15 men) participated in the tests (age bracket: 19-43). All participants had experience in using notebooks, and the majority of them had experience in using touch pads.

ASUS Eee PC 701SD, 900HA, 1000HG and 1101HA models were used during the tests (screen diameters: 7", 8.9", 10.1" and 11.6", respectively). The screen resolution was set to 800×480 , 1024×600 , 1024×600 and 1366×768 pixels, respectively. A comparable RAS (resistive) touch screen was installed for each model to compare the operation of a touch pad and touch screen with the same model.

The tests comprised five tasks. Task 1 (omnidirectional clicking) and task 2 (omnidirectional selection with “tap” gestures) were easy tasks. Task 3 (consecutive selection and clicking) and task 4 (consecutive selection by tapping) and task 5 (typing in Mandarin using a keyboard) were complex tasks.

The tests within tasks 1 and 2 were performed as per Appendix B of ISO 9241-9 [14]. During task 3 (consecutive selection and clicking), users selected icons by means of a touch pad. An icon (button) 1 was selected. Next, buttons from 2 to 5 were selected consecutively. Similarly, during task 4 (consecutive selection by tapping), the user selected icons (buttons) consecutively, using a touch screen. The total distance (from button 1 to button 5) was maintained at 1000 pixels.

The average operation time measurement results in this test showed that the tasks were performed more slowly on models fitted with smaller screens. Taking into account all tasks, a significant relation between a device size and task performance time was demonstrated.

Significant differences in the performance of simple tasks with a touch pad (task 1: multi-directional clicking) and tasks performed with a touch screen (task 2: multi-directional tapping test) were also demonstrated.

Taking into account the fact that the error rate was below 5%, the error percentage values were not significantly different across the various device models.

The details of the input test results, including average task performance times and error rates for all device models and all tasks, are presented in paper [7].

The tests presented in paper [8] are considered to be relevant studies in the field of the touch-input of data. The tests consisted of selecting objects. The input device was a mouse connected to a laptop and a smartphone touch screen (inputting data with a finger).

Sixteen people (including 6 women) participated in the tests. The laptop tests were performed in a seated position, and the touch screen tests were conducted in a standing position. The tests were performed with an LG Nexus 4 smartphone (touch screen and Android 4.2.2 OS). Screen size 61 × 102 mm, resolution 768 × 1184 pixels.

The study included typical, one-dimensional tests (a one-directional test and indication of objects) and two-dimensional tests (a multi-directional test and indication of objects) as per ISO 9241-9 [14].

Each participant performed tests on each device for each block of conditions and each task type (one- and two-dimensional), taking into account the distances between objects and object sizes.

The test results included throughput, movement time and error rate. It was demonstrated that efficiency is higher by 42-88% with touch operation (direct input), in relation to the efficiency obtained when a mouse was used (indirect input). The throughput for the two-dimensional tasks was significantly lower than for the one-dimensional tasks.

No significant differences between movement time and device placement (smartphone held in hand, smartphone laid on a desk) were determined. However, differences between movement duration in relation to given task type (shorter times during one-dimensional tasks) were detected.

No statistically relevant differences in error rates in relation to task types and device placement were determined. Detailed information concerning these test results are published in paper [8].

Apart from the studies presented in the available literature that have been presented in detail, numerous interesting studies can be mentioned whose detailed descriptions are impossible to present within a single paper. For example, they include studies concerning the time and precision of object indication on a smartphone screen while driving a car (in a simulator) presented in paper [9]. Participants were asked to take their right hand from the steering wheel and indicate and touch a target on a smartphone screen, defined by the person running the experiment. Movement times and touch coordinates were recorded.

Paper [10] presents studies concerning the impact of device feedback after the selection of very small (2-4 mm) targets on touch screens. The feedback comprised the contact point on a touch screen after a user raised their finger. The tests showed that application of device feedback resulted in increasing the object indication time, but also in significantly decreasing the error rate value.

The studies presented in article [11] were aimed at such gestures as swiping, pinching (spreading) and complex gestures such as double touch with swiping and swiping with double tapping, as well as swiping down the screen. Children, adults and seniors participated in the tests. The tests confirmed that children and adults achieve better results than seniors.

Attention should also be paid to studies in which precision and speed of two zooming methods are compared, i.e. the tap-and-drag and traditional pinch-to-zoom method [12]. The results showed that the tap-and drag method was 47% more effective than pinch-to-zoom, and the effectiveness was measured on the basis of the number of gestures necessary to complete a test.

3. Test conditions

3.1. Device used in tests

A mobile device, i.e. a Nokia Lumia 800 smartphone (hereafter referred to as “the smartphone”) was used during the tests. Basic smartphone parameters: MS Windows Phone 7.5 Mango operating system, single-core Qualcomm MSM8255T 1.40 GHz processor. Screen type: capacitive display with the multi-touch function, 3.7” in size. Display technology: AMOLED with ClearBlack, facilitating operation in WVGA (480 × 800 px, 252 ppi ~54.7% screen-to-body ratio) resolution [19], [20]. A capacitive display precludes using a stylus.

3.2. Application facilitating test completion

The application facilitating test completion was of modular design (Fig. 1) [13]:

- *Mobile Application* module run on a phone operated by the Windows Phone system;
- *Analysis Program* module run on an IBM PC with Windows 7 installed;
- *Server Program* module, i.e. a webservice run on an IBM PC.

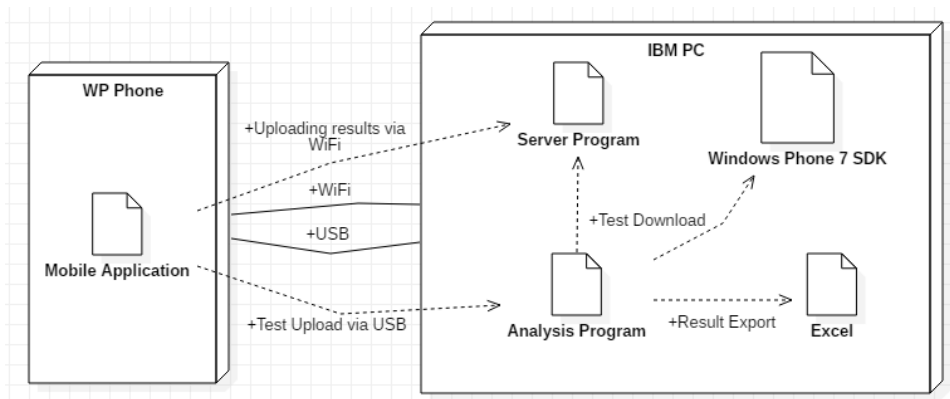


Fig. 1. Architecture of application facilitating test completion

After starting the *Mobile Application* module, a test participant filled in a short survey (Fig. 2a) and made certain gestures, repeating them a preset number of times (Fig. 2b). See Fig. 2c for an example screen showing the status after objects were displayed and before a gesture was made.

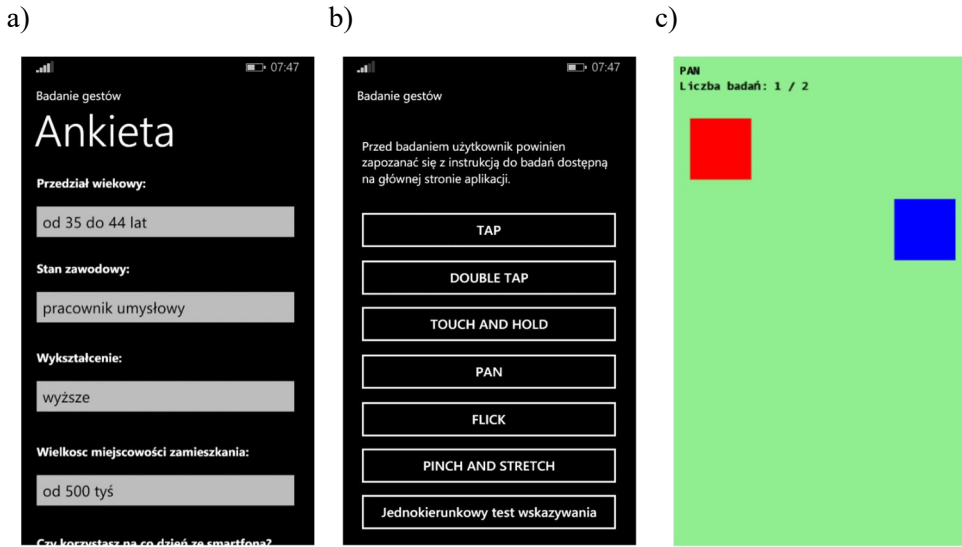


Fig. 2. Mobile Application a) survey, b) measurement types, c) a view during one “pan” gesture

Test results were sent from the *Mobile Application* module to the *Analysis Program* module using the *Server Program* webservice via WiFi or directly from the *Mobile Application* module to the *Analysis Program* module via a USB interface. The *Analysis Program* (Fig. 3) could present total results (all test participants) or single test results (gesture or person).

Selected measurement results could be exported from the *Analysis Program* module to an Excel file. In the studies presented, the results were exported to Excel files for further processing. The basic data included in these Excel files comprised:

- indication of whether a target was hit or missed;
- distance between an object and finger at the moment of touching the screen (for the “flick” gesture, this parameter was unavailable);
- gesture completion time.

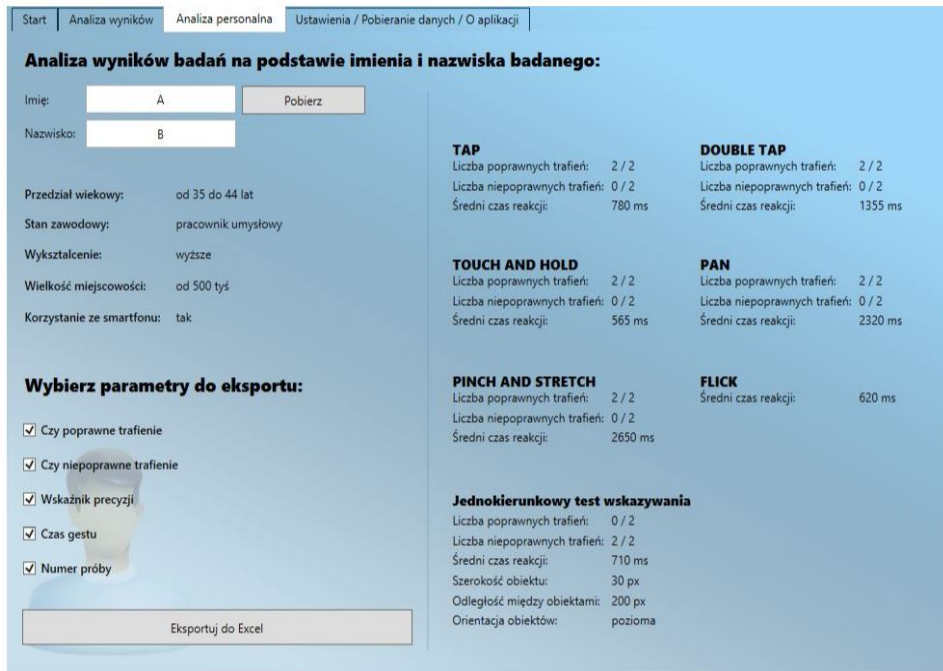


Fig. 3. Example of the display in the *Analysis Program* module

3.3. Number of test participants and test conditions

See Table 2 for the number of test participants. 60 people participated in the tests (age bracket: 16-66, with the majority of young and middle-aged people). Five age groups (1-5) were distinguished. The test participants included 2 women.

Tab. 2. Number of test participants

Age brackets [years of age]	16-24	25-34	35-44	45-54	≥55	Total
Group number	1	2	3	4	5	
Number of participants	14	5	23	13	5	60

The test participants included mostly men, white-collar workers with higher education degrees as well as several students, i.e. secondary school graduates. Some participants used smartphones on an everyday basis, and some did not use them on an everyday basis (Tab. 3). All participants repeated the gestures 30 times.

Tab. 3. Participants using and not using smartphones on an everyday basis

Feature	Uses a smartphone	Does not use a smartphone
Number of participants	48	12

The tests were conducted from 8.00 a.m. to 4 p.m., in one room. A smartphone was held in an upright position, in a left hand (none of the participants were left-handed). The participants did not move (walk). Most of them were sitting. The next object (target) requiring a participant's action was displayed immediately after finishing the previous gesture. No training or familiarising activities were conducted prior to the tests. The participants were informed orally about the following tasks to complete.

The tests presented in this paper involved tapping, double tapping, flicking and panning.

The main parameters determined on the basis of measurements taken were:

- average time of making a gesture by a user;
- average gesture performance precision (average distance between an object and finger).

Statistical tests adequate to the study conditions were used for calculations presented in this paper [15], [16], [18]. Direct calculations and diagrams were elaborated using the MATLAB software for scientific and technical calculations [17].

4. “Tap” test results

During the “tap” test, a red square (100 × 100 pixels) was subjected to manipulations (Fig. 4). Thus the square was slightly larger than a typical finger touching a screen surface. The square was displayed in randomly selected positions, on the smartphone screen.

Values measured for a participant (Fig. 5):

x_{so} , y_{so} - coordinates of the object (square) centre;

x_p , y_p - coordinates of the screen point touched by the participant;

t_a - object (square) displaying time;

t_b - time in which the participant removed their finger from the screen.

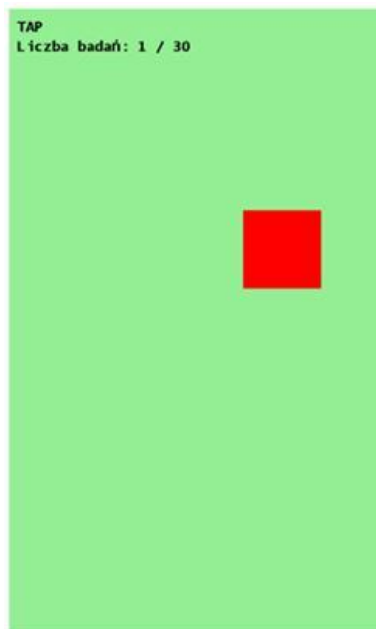


Fig. 4. Nokia Lumia 800 smartphone view while making the “tap” gesture

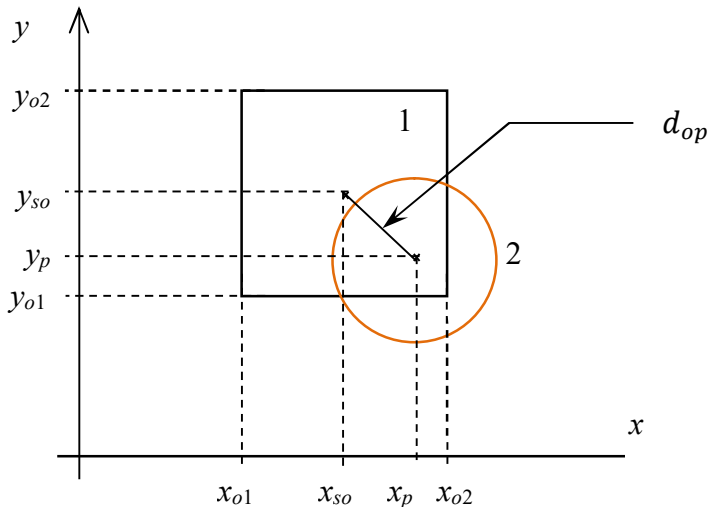


Fig. 5. Example of object 1 (square) and finger 2 locations on the screen, where:
 $(x_{o1}, y_{o1}), (x_{o2}, y_{o2})$ - coordinates of the object 1 (square) centre;

“Tap” gesture correctness condition:

$$(x_{o1} \leq x_p \leq x_{o2}) \wedge (y_{o1} \leq y_p \leq y_{o2}) \quad (9)$$

where: $x_{o1}, x_{o2}, y_{o1}, y_{o2}$ - as in Fig. 5.

The following values were determined during the test.

Distance d_{op} between the object and finger (hereinafter referred to as “precision”):

$$d_{op} = \sqrt{(x_{so} - x_p)^2 + (y_{so} - y_p)^2}. \quad (10)$$

Length of time in which the participant made the gesture:

$$t_1 = t_b - t_a,$$

where: t_a - object displaying time;

t_b - time in which the participant removed their finger from the screen.

Following the measurements, such aspects as average gesture precision, average gesture completion time and the probability of an error while making a “tap” gesture were determined.

Average gesture precision:

$$\bar{d}_{op} = \frac{1}{n} \sum_{i=1}^n d_{op i},$$

where: $d_{op i}$ - gesture precision following the i -th measurement;

n - number of gesture repetitions ($n = 30$).

Average gesture completion time:

$$\bar{t}_1 = \frac{1}{n} \sum_{i=1}^n t_{1 i}, \quad (11)$$

where: $t_{1 i}$ - gesture completion time following the i -th measurement;

n - number of gesture repetitions ($n = 30$).

The probability of an error while making a “tap” gesture was determined by means of the gesture completion correctness condition (compare dependence [9]).

In order to make the necessary comparisons, the average “tap” gesture precision of \bar{d}_{op}^g was determined in age group g , $g \in \{1, 2, 3, 4, 5\}$. The calculation results are presented in Table 4.

See Figure 6 for the average precision, standard deviation of precision and precision confidence ranges for “tap” gesture completion in the given age groups.

Tab. 4. Average precision \bar{d}_{op}^g for “tap” gesture completion in given age group g

Age bracket [years of age]	16-24	25-34	35-44	45-54	≥ 55
Group g	1	2	3	4	5
Average tap gesture precision \bar{d}_{op}^g [pix]	23.15	18.77	18.62	18.18	19.92

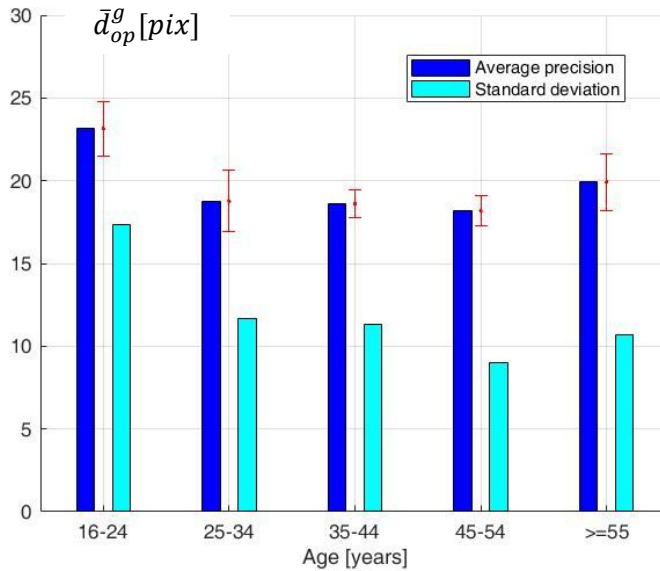


Fig. 6. Average precision \bar{d}_{op}^g , standard deviation of precision and precision confidence ranges for “tap” gesture completion in given age groups

Similarly, the average time for “tap” gesture completion \bar{t}_1^g in a given age group g was determined, $g \in \{1, 2, 3, 4, 5\}$. Apart from the gesture completion time in an age group, gesture completion time standard deviation and confidence ranges for gesture completion time in an age group were determined. See Table 5 for calculation results for gesture completion average time for all gestures and for correctly completed gestures (condition [10] fulfilled).

See Figure 7 for the average time, standard deviation of time and time confidence ranges for “tap” gesture completion in given age groups.

Tab. 5. Average time \bar{t}_1^g for “tap” gesture completion in given age group g

Age bracket [years of age]	16-24	25-34	35-44	45-54	≥ 55
Group g	1	2	3	4	5
Average time \bar{t}_1^g for “tap” gesture completion (all gestures) [ms]	584.2	639.5	751.7	931.0	952.5
Average time \bar{t}_1^g for “tap” gesture completion (correct gestures) [ms]	588.2	641.6	753.1	931.7	952.5

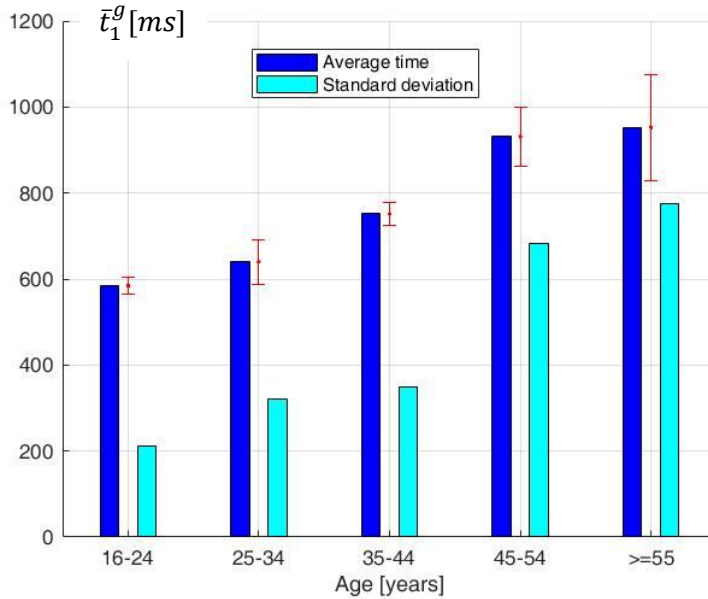


Fig. 7. Average gesture completion time \bar{t}_1^g , standard deviation and confidence ranges for “tap” gesture completion in given age groups (all gestures)

Taking into account the “tap” gesture correctness requirement (condition [9]), the probability Pb_{tap} of gesture completion errors in the given age groups was determined. The results are shown in Figure 8.

Test results among age groups were compared, in particular between the first group and the other groups.

The following hypotheses regarding gesture completion average precision were formulated:

H0: equal average precision for age groups 1 and j ($\bar{d}_{op}^1 = \bar{d}_{op}^j$);

H1: different average precision for age groups 1 and j ($\bar{d}_{op}^1 \neq \bar{d}_{op}^j$).

Similarly, the following hypotheses regarding gesture completion average time were formulated:

H0: equal average time for age groups 1 and j ($\bar{t}_1^1 = \bar{t}_1^j$);

H1: different average time for age groups 1 and j ($\bar{t}_1^1 \neq \bar{t}_1^j$).

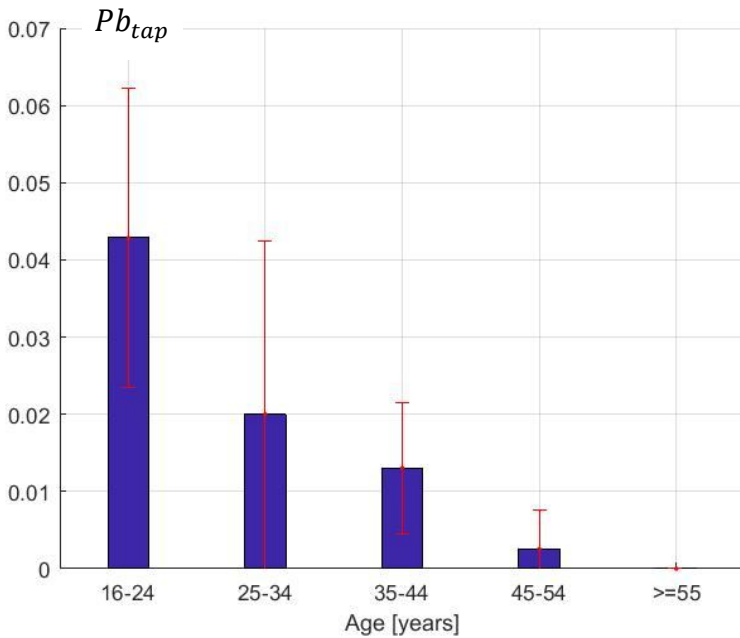


Fig. 8. Probability of errors Pb_{tap} and confidence ranges for “tap” gesture in age groups

Due to the fact that the tests were numerous, a typical test for comparing average values was used. The hypotheses were tested at the significance level of $\alpha = 0.05$, and the hypothesis verification results are given in Table 6.

Tab. 6. Results of comparison between average precision and average time of “tap” gesture completion, between group 1 and the other age groups

Groups subject to comparison	1-2	1-3	1-4	1-5
Decision on average precision comparison result	Reject H0	Reject H0	Reject H0	Reject H0
Decision on average time comparison result	No grounds to reject H0	Reject H0	Reject H0	Reject H0

The following designations were assumed.

\bar{d}_{op}^u – average precision of gesture completion by users using smartphones on an everyday basis;

\bar{d}_{op}^n – average precision of gesture completion by users not using smartphones on an everyday basis;

\bar{t}_1^u – average time of gesture completion by users using smartphones on an everyday basis;

\bar{t}_1^n – average time of gesture completion by users not using smartphones on an everyday basis;

Pb_{tap}^u – probability of errors in gesture completion by users using smartphones on an everyday basis;

Pb_{tap}^n – probability of errors in gesture completion by users not using smartphones on an everyday basis.

On the basis of the results obtained, the average gesture precision, average gesture completion time and probability of a “tap” gesture error were determined within the group of participants using smartphones on an everyday basis and the group of participants not using smartphones on an everyday basis. The results are shown in Fig. 9-11.

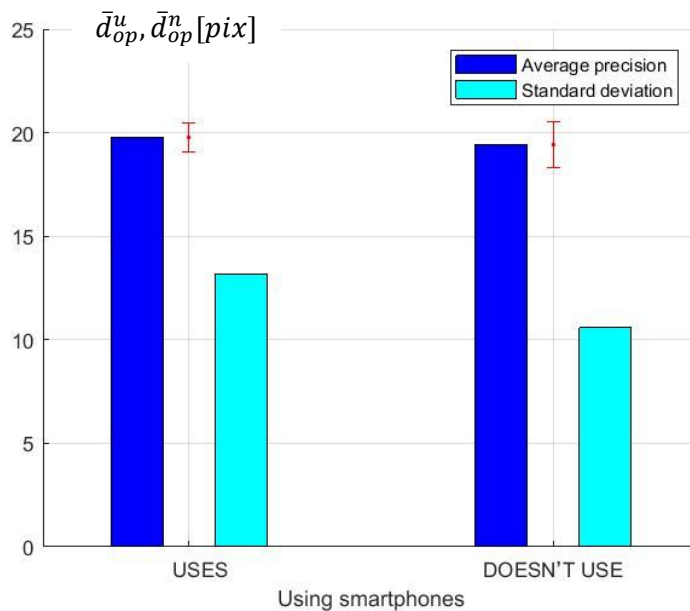


Fig. 9. Average precision, standard deviation of precision and precision confidence ranges for “tap” gesture completion for participants who use and do not use smartphones on an everyday basis

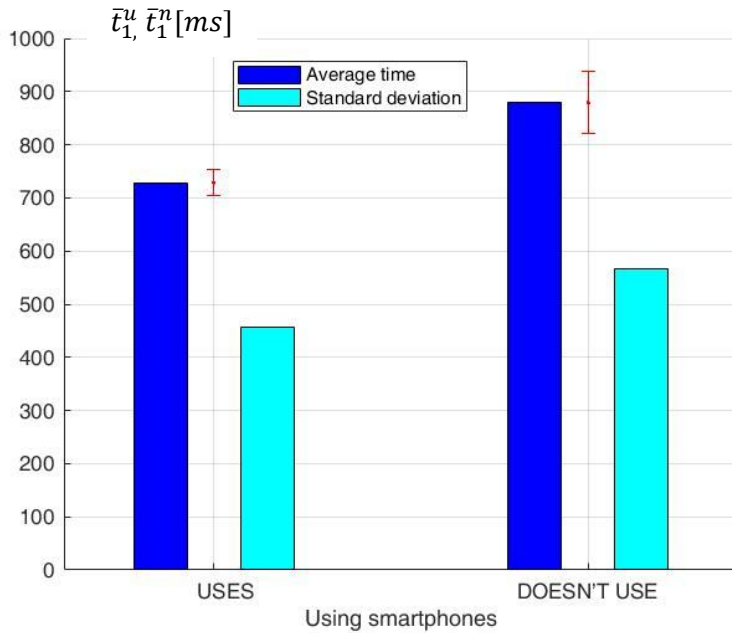


Fig. 10. Average time, standard deviation and precision confidence ranges for “tap” gesture completion time for participants who use and do not use smartphones on an everyday basis

It was also checked whether the fact that participants use smartphones on an everyday basis influenced the quality of their operation. The following hypotheses were formulated:

H0: equal average precision (participants using; participants not using smartphones on an everyday basis) ($\bar{d}_{op}^u = \bar{d}_{op}^n$),

H1: different average precision ($\bar{d}_{op}^u \neq \bar{d}_{op}^n$),

H0: equal average time ($\bar{t}_1^u = \bar{t}_1^n$),

H1: different average time ($\bar{t}_1^u \neq \bar{t}_1^n$),

H0: equal error probability ($Pb_{tap}^u = Pb_{tap}^n$),

H1: different error probability ($Pb_{tap}^u \neq Pb_{tap}^n$).

The hypotheses were tested at the significance level of $\alpha = 0.05$, and verification results are given in Table 7.

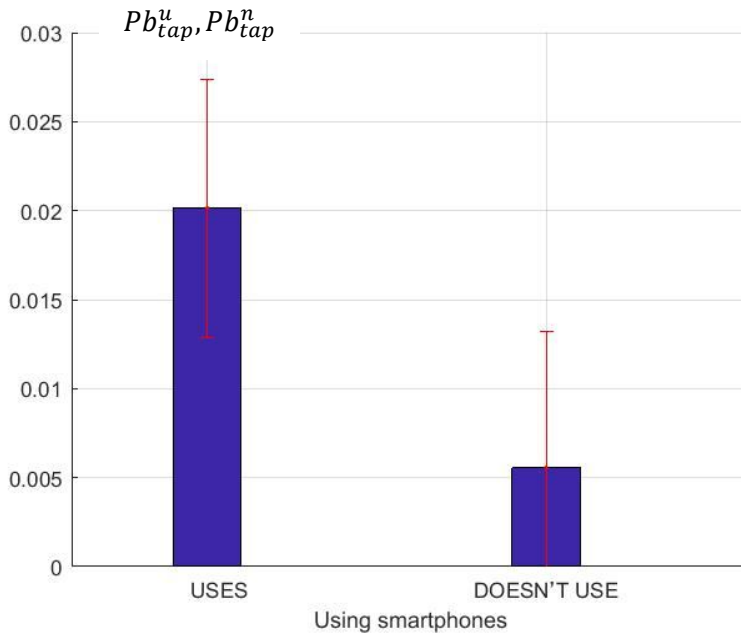


Fig. 11. Error probability and confidence ranges for “tap” gesture completion for participants who use and do not use smartphones on an everyday basis

Tab. 7. Results of comparison between average time, average precision and error probability for “tap” gesture completion for participants who use and do not use smartphones on an everyday basis

Compared parameter	Average time	Average precision	Error probability
Decision on H0	Reject H0	No grounds to reject H0	No grounds to reject H0

Comments and conclusions regarding test results

On the basis of the results obtained, one can observe fairly significant standard deviation values, both for the average “tap” gesture completion precision and average “tap” gesture completion time, in various age groups. Practically, taking into account the average time \bar{t}_1^g of

“tap” gesture completion in an age group, no differences between all gestures and correct gestures (compare Tab. 5) can be observed.

Characteristic values include a decrease in standard deviation of precision (compare Fig. 6), an increase in standard deviation of gesture completion time (compare Fig. 7), and a decrease in gesture completion error probability (compare Fig. 8) in relation to the participants' age. Thus it can be concluded that the older the users are, the more significantly slow and less precise gesture completion is (compare Tab. 6).

In all cases, the hypotheses regarding equality of average precision between the first age group and other groups were rejected. In the majority of cases (apart from the comparison with the second group), hypotheses regarding the equality of “tap” gesture average completion time between the first age group and other groups were rejected.

Taking into account participants using and not using a smartphone on an everyday basis, the hypothesis regarding the equality of the “tap” gesture average completion time between these groups was rejected. For average precision and error probability in “tap” gesture completion, there were no grounds for rejecting the hypotheses regarding the equality of these parameters among participants using and not using smartphones on an everyday basis.

5. “Double tap” test results

In the “double tap” gesture studies, the same object as for the “tap” gesture studies was used (compare Fig. 4). The square was displayed in randomly selected positions, on a smartphone screen.

The values measured for the test participant and the gesture completion correctness requirement were the same as for the “tap” gesture, however, the coordinates of the screen point touched by the participant, and the time of removing the finger from the screen referred to the second tap.

During those tests, values similar to the “tap” test were determined, i.e. the distance d_{op} between the object and finger (second tap on the object – precision) and gesture completion time, and time t_b was the time of removing the participant's finger from the screen for the second time.

Following the measurements, such aspects as average gesture precision, average gesture completion time and probability of an error while making a “double tap” gesture were determined.

In order to make the necessary comparisons, average “double tap” gesture precision \bar{d}_{op}^g was determined in age group g , $g \in \{1, 2, 3, 4, 5\}$. The calculation results are presented in Table 8.

Tab. 8. Average precision \bar{d}_{op}^g for “double tap” gesture completion in given age group g

Age bracket [years of age]	16-24	25-34	35-44	45-54	≥ 55
Group g	1	2	3	4	5
Average “double tap” gesture precision \bar{d}_{op}^g [pix]	17.58	21.11	17.76	18.53	20.61

See Figure 12 for the average precision, standard deviation of precision and precision confidence ranges for “double tap” gesture completion in the given age groups.

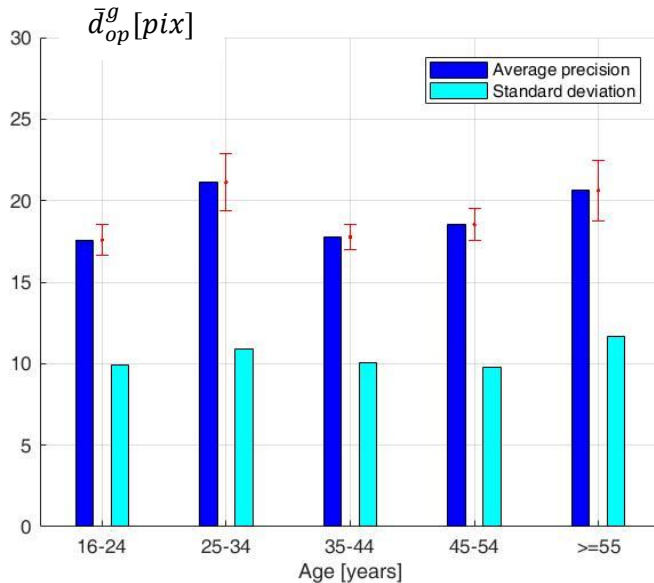


Fig. 12. Average precision \bar{d}_{op}^g , standard deviation of precision and precision confidence ranges for “double tap” gesture completion in the given age groups

Similarly, the average time for “double tap” gesture completion \bar{t}_1^g was determined in a given age group g , $g \in \{1, 2, 3, 4, 5\}$. Apart from the gesture completion time in the age group, gesture completion time standard deviation and confidence ranges for gesture completion time in the age group were

determined. See Table 9 for calculation results for gesture completion average times for all gestures and for correctly completed gestures (condition [10] fulfilled).

Tab. 9. Average time \bar{t}_1^g for “double tap” gesture completion in given age group g

Age brackets [years of age]	16-24	25-34	35-44	45-54	≥ 55
Group g	1	2	3	4	5
Average time \bar{t}_1^g for “double tap” gesture completion (all gestures) [ms]	665.7	695.3	821.0	999.8	1172.4
Average time \bar{t}_1^g for “double tap” gesture completion (correct gestures) [ms]	665.7	696.0	821.2	999.8	1177.2

See Figure 13 for the average time, standard deviation of time and time confidence ranges for “double tap” gesture completion in given age groups.

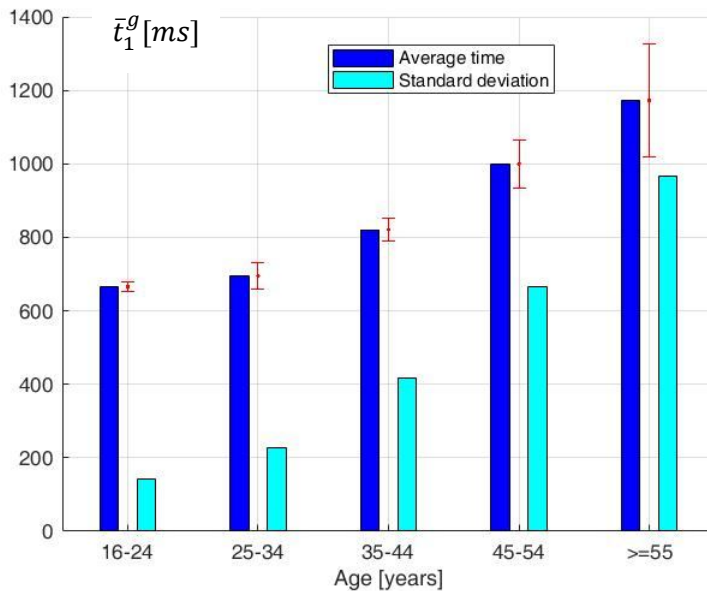


Fig. 13. Average gesture completion time \bar{t}_1^g , standard deviation and confidence ranges for “double tap” gesture completion in the given age groups (all gestures)

Similarly to the “tap” gesture, taking into account the gesture completion correctness requirement (condition [9], when the object was tapped for the

second time) the probability Pb_{Dtap} of the “double tap” gesture completion error, in age groups, was determined. The results are shown in Figure 14.

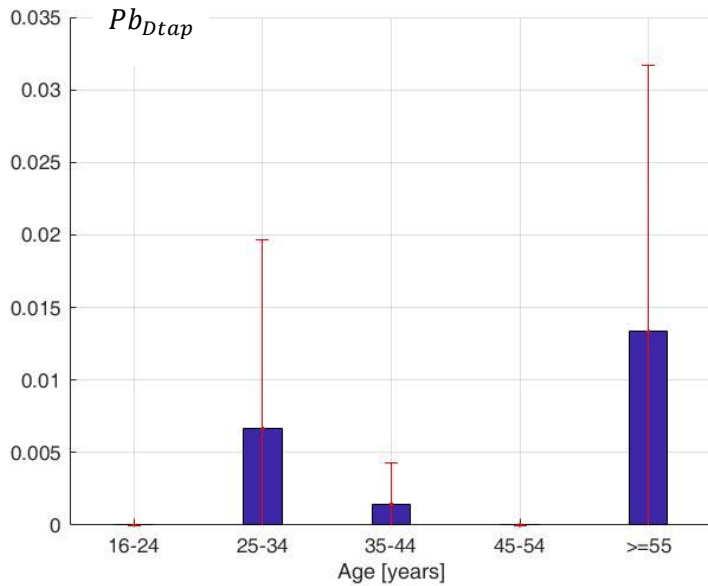


Fig. 14. Probability of errors Pb_{Dtap} and confidence ranges for “double tap” gesture in age groups

The test results among the age groups were compared, in particular between the first group and the other groups.

The following hypotheses regarding “double tap” gesture completion average precision were formulated

H0: equal average precision for age groups 1 and j ($\bar{d}_{op}^1 = \bar{d}_{op}^j$);

H1: different average precision for age groups 1 and j ($\bar{d}_{op}^1 \neq \bar{d}_{op}^j$).

Similarly, the following hypotheses regarding “double tap” gesture completion average time were formulated:

H0: equal average time for age groups 1 and j ($\bar{t}_1^1 = \bar{t}_1^j$);

H1: different average time for age groups 1 and j ($\bar{t}_1^1 \neq \bar{t}_1^j$).

The hypotheses were tested at the significance level of $\alpha = 0.05$, and verification results are given in Table 10.

The following designations were assumed:

\bar{d}_{op}^u – average precision of “double tap” gesture completion by users using smartphones on an everyday basis;

\bar{d}_{op}^n – average precision of gesture completion by users not using smartphones on an everyday basis;

\bar{t}_1^u – average time of “double tap” gesture completion by users using smartphones on an everyday basis;

\bar{t}_1^n – average time of gesture completion by users not using smartphones on an everyday basis;

Pb_{Dtap}^u – probability of errors in “double tap” gesture completion by users using smartphones on an everyday basis;

Pb_{Dtap}^n – probability of errors in gesture completion by users not using smartphones on an everyday basis.

Tab. 10. Results of comparison between average precision and average time of “double tap” gesture completion, between group 1 and the other age groups

Groups subject to comparison	1-2	1-3	1-4	1-5
Decision on average precision comparison	Reject H0	No grounds to reject H0	No grounds to reject H0	Reject H0
Decision on average time comparison	No grounds to reject H0	Reject H0	Reject H0	Reject H0

On the basis of the results obtained, the average gesture precision, average gesture completion time and probability of a “double tap” gesture error were determined within the group of participants using smartphones on an everyday basis and the group of participants not using smartphones on an everyday basis. The results are shown in Fig. 15-17.

It was also checked whether the fact that participants use smartphones on an everyday basis influences the quality of “double tap” gesture completion. The following hypotheses were formulated:

H0: equal average precision (participants using; participants not using smartphones on an everyday basis) ($\bar{d}_{op}^u = \bar{d}_{op}^n$),

H1: different average precision ($\bar{d}_{op}^u \neq \bar{d}_{op}^n$),

H0: equal average time ($\bar{t}_1^u = \bar{t}_1^n$),

H1: different average time ($\bar{t}_1^u \neq \bar{t}_1^n$),

H0: equal error probability ($Pb_{Dtap}^u = Pb_{Dtap}^n$),

H1: different error probability ($Pb_{Dtap}^u \neq Pb_{Dtap}^n$).

The hypotheses were tested at the significance level of $\alpha = 0.05$, and verification results are given in Table 11.

Tab. 11. Results of comparison between average time, average precision and error probability for “double tap” gesture completion for participants who use and do not use smartphones on an everyday basis

Compared parameter	Average time	Average precision	Error probability
Decision on H0	Reject H0	No grounds to reject H0	No grounds to reject H0

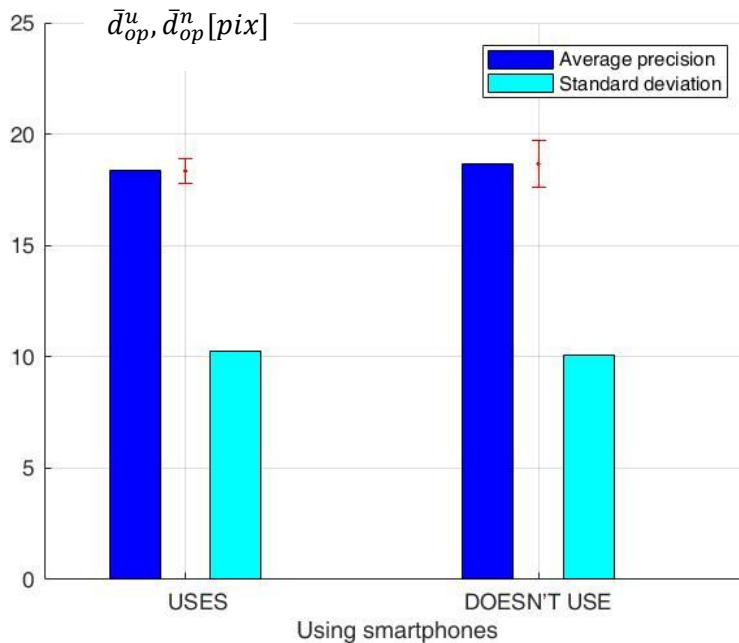


Fig. 15. Average precision, standard deviation of precision and precision confidence ranges for “double tap” gesture completion for participants who use and do not use smartphones on an everyday basis

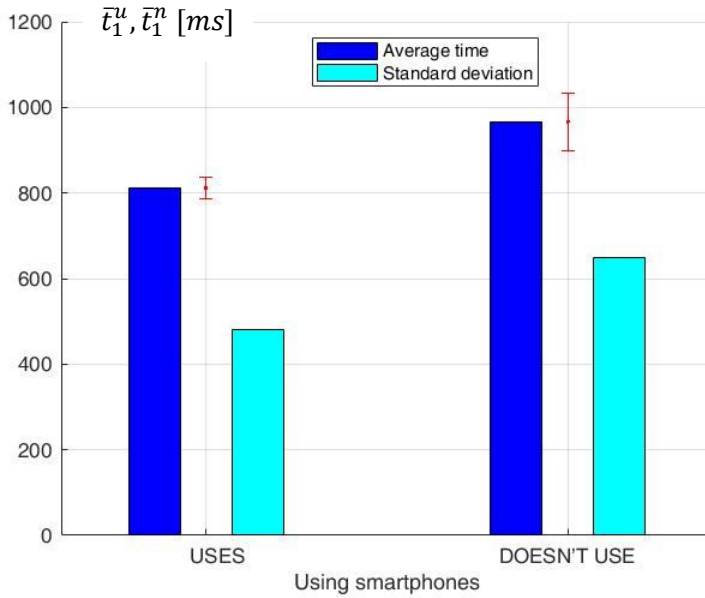


Fig. 16. Average time, standard deviation of time and precision confidence ranges for “double tap” gesture completion time for participants who use and do not use smartphones on an everyday basis

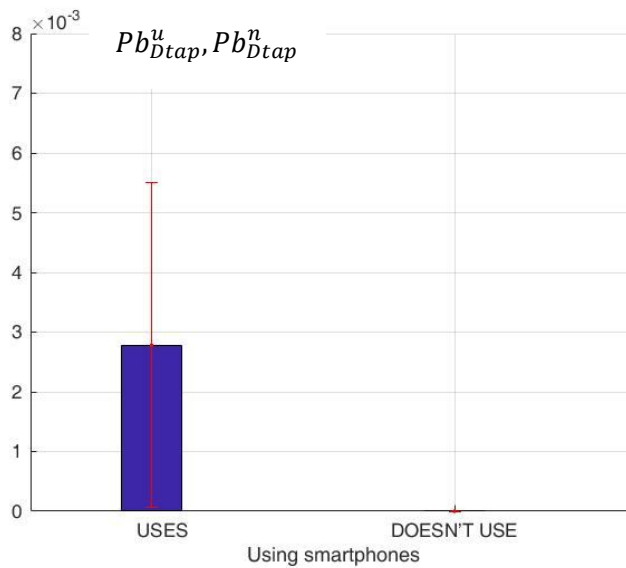


Fig. 17. Error probability and confidence ranges for “double tap” gesture completion error probability for participants who use and do not use smartphones on an everyday basis

Comments and conclusions regarding test results

On the basis of the results obtained, one can observe fairly significant standard deviation values for the average “double tap” gesture completion precision in each age group. A significant increase in the standard deviation of “double tap” gesture completion average time with the increase in the participants’ ages (compare Fig. 13) can be observed.

The result of the comparison of the “double tap” gesture completion average time among age groups is the same as for the “tap” gesture. The hypotheses regarding the equality of the “double tap” completion average time between the first age group and other groups (apart from the comparison with the second group) (compare Tab. 10) must be rejected.

Taking into account participants using and not using smartphones on an everyday basis, the results of the comparison of the average gesture precision, average gesture completion time and probability of an error while making the “double tap” gesture are the same as for the “tap” gesture - the hypothesis regarding the equality of the gesture completion average time was rejected. Nevertheless, for average precision and error probability in “double tap” gesture completion, there were no grounds for rejecting the hypotheses regarding the equality of these parameters among participants using and not using smartphones on an everyday basis.

6. “Flick” gesture test results

During the “flick” gesture tests, the object subjected to manipulations was the same as the one used for the “tap” gesture tests (compare Fig. 4). The square was displayed in randomly selected positions on the smartphone screen.

“Flick” gesture measurements concerned only time, so the values measured for a tested participant were:

t_a - object (square) displaying time;

t_b - time in which the participant removed their finger from the screen.

After the measurements were taken, the average time for “flick” gesture completion was determined (compare formula [11]). Gesture completion time standard deviation and confidence ranges for gesture completion time in an age group were determined. See Table 12 for the calculation results regarding the “flick” gesture completion average time.

Tab. 12. Average time \bar{t}_1^g for “flick” gesture completion in given age group g

Age bracket [years of age]	16-24	25-34	35-44	45-54	≥ 55
Group g	1	2	3	4	5
Average “flick” gesture completion time \bar{t}_1^g [ms]	714.0	706.0	741.8	1017.5	1093.3

See Figure 18 for the average time, standard deviation of time and time confidence ranges for “flick” gesture completion the in given age groups.

Test results (“flick” gesture completion average time) among age groups were compared, in particular between the first group and the other groups.

The following hypotheses regarding “flick” gesture completion average time were formulated:

H0: equal average time for age groups 1 and j ($\bar{t}_1^1 = \bar{t}_1^j$);

H1: different average time for age groups 1 and j ($\bar{t}_1^1 \neq \bar{t}_1^j$).

The hypotheses were tested at the significance level of $\alpha = 0.05$, and verification results are given in Table 13.

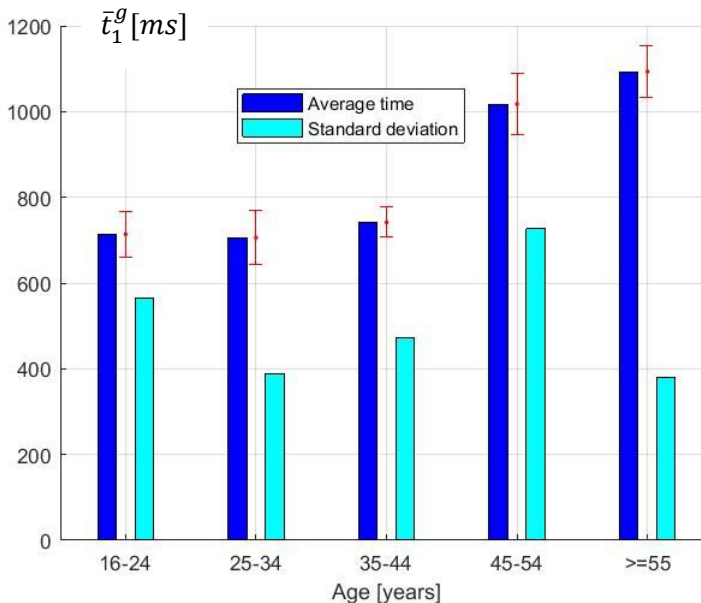


Fig. 18. Average “flick” gesture completion time \bar{t}_1^g , standard deviation and confidence ranges for gesture completion in the given age groups

Tab. 13. Results of comparison between average time of “flick” gesture completion between group 1 and the other age groups

Groups subject to comparison	1-2	1-3	1-4	1-5
Decision on H0	No grounds to reject H0	No grounds to reject H0	Reject H0	Reject H0

The following designations were assumed:

\bar{t}_1^u – average time of “flick” gesture completion by users using smartphones on an everyday basis;

\bar{t}_1^n – average time of gesture completion by users not using smartphones on an everyday basis;

On the basis of the results obtained, the average “flick” gesture completion time within the group of participants using smartphones on an everyday basis and the group of participants not using smartphones on an everyday basis was determined. The results are shown in Figure 19.

It was also checked whether the fact that participants use smartphones on an everyday basis influenced the quality of “flick” gesture completion. The following hypotheses were formulated:

H0: equal average time ($\bar{t}_1^u = \bar{t}_1^n$),

H1: different average time ($\bar{t}_1^u \neq \bar{t}_1^n$).

The hypothesis was tested at the probability level $\alpha = 0.05$, and the verification result indicated that hypothesis H0 must be rejected, to the advantage of an alternative hypothesis.

Comments and conclusions regarding test results

On the basis of the results obtained, one can observe fairly significant standard deviation values for the average “flick” gesture completion time, regardless of the age group (compare Fig. 18).

While comparing the average time for “flick” gesture completion among age groups, it can be concluded that differences between the first and fourth as well as first and fifth age group (for larger age difference among the participants) are significant.

It was also concluded that differences concerning the average time for “flick” gesture completion are significant, taking into account persons using and not using smartphones on an everyday basis.

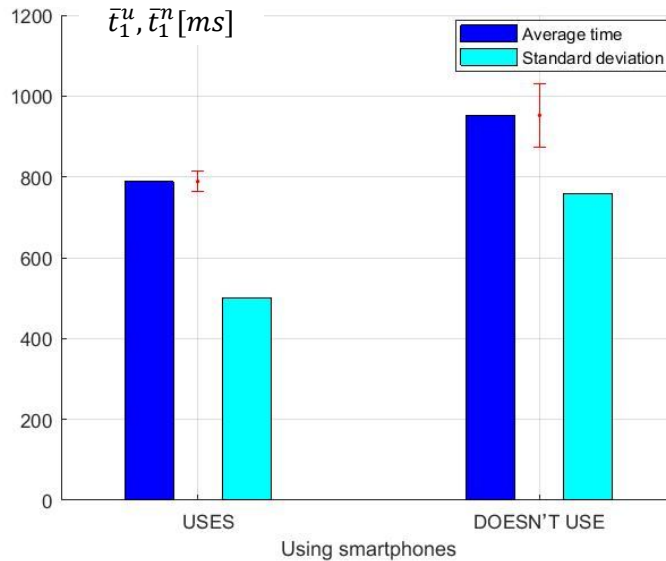


Fig. 19. Average time, standard deviation of time and precision confidence ranges for “flick” gesture completion time for participants who use and do not use smartphones on an everyday basis

7. Summary of conclusions

This paper constitutes the first part of a broader study concerning the results of tests examining the quality of test participants’ operation while using a smartphone (and/or tablet) and indicating objects on the screen using certain gestures. The tests were performed using an application based on recommendations provided in the ISO 9241-9 standard [14]. The test results presented in the paper concern “tap”, “double tap” and “flick” gestures. Gesture completion time and gesture completion precision (distance between the screen point being touched with a finger and the centre of the object being touched) were used as the basic quality characteristics.

While comparing the test results with data available in the literature, it can be observed that, as far as the “tap” gesture is concerned, the age of the test participants had a significant impact on precision (compare Tab. 6 and the conclusions in paper [3]). The partial confirmation regards the gesture completion average time, and one must remember that in paper [3], the “tap” gesture tests were performed differently from the tests performed by us. The

impact of the test participants' ages on the input characteristics subject to comparison is clearly visible. The older the users are, the more significantly slow and less precise gesture completion is (compare Tab. 6). Taking into account test participants using and not using smartphones on an everyday basis, this characteristic has a significant impact only on the average time of "tap" gesture completion, however, it has no significant influence on the average gesture precision or error probability.

As for the "double tap" gesture, the participants' ages exerted no unequivocal impact on the characteristics. A significant impact is only visible for the gesture completion average time; however, it does not concern the comparison between the first and the second age group. Thus the conclusion here is the same as for the "tap" gesture. The results of the comparison of the gesture precision, gesture completion average time and gesture completion error probability for the "double tap" gesture in the case of test participants using and not using smartphones on an everyday basis are the same as for the "tap" gesture - the hypothesis regarding the equality of the gesture completion average time was rejected.

While comparing the average precision of "tap" and "double tap" gesture completion in age groups, slight differences that do not exceed 2.5 pixels, apart from the first age group, can be observed. While comparing the significance of the average precision for "tap" and "double tap" gestures, it was determined that, for the first age group, the hypothesis regarding the equality of average precision of "tap" and "double tap" gestures must be rejected. As for the other groups, there are no grounds for rejecting the hypothesis regarding the equality of average precision of "tap" and "double tap" gestures.

Taking into account the average time of "tap" and "double tap" gesture completion in various age groups, these differences do not exceed 90 ms, apart from the oldest age group, where this difference exceeds 200 ms. While studying the significance of average completion time for (all) "tap" and "double tap" gestures, it was determined that, for the first, third and fifth age group, the hypothesis regarding the equality of average time of "tap" and "double tap" gesture completion must be rejected. As for the other groups, there are no grounds for rejecting the hypothesis regarding the equality of average time of "tap" and "double tap" gesture completion. Exactly the same results of the tests regarding the significance of average time for "tap" and "double tap" gesture completion were obtained, taking into account only correct gestures.

Taking into account the "flick" gesture and its average completion time, significant differences between the first and fourth, and the first and fifth age groups were determined (significant differences at a higher age difference between the tested groups). No significant differences for the other age groups were identified.

The next paper will present the results of tests examining the quality of test participants' operation using a smartphone, taking into account gestures different from the ones presented in this paper.

Literature

- [1] DONIGIEWICZ A.M., *Modelowanie interakcji człowiek-komputer. Problemy oceny jakości i niezawodności*, Wyd. WAT, Warszawa, 2005.
- [2] DONIGIEWICZ A.M., *Badanie czasu wskazywania obiektu przez użytkownika testem wielokierunkowym. Część II – wyniki z lat 2015-2016*. Przegląd Teleinformatyczny, nr 1-2, 2017, s. 13-39.
- [3] HOURCADE J.P., BERKEL T.R., *Simple pen interaction performance of young and older adults using handheld computers*. Interacting with Computers, 20, 2008, pp. 166–183.
- [4] FUJIOKA R., AKIBA T., OKADA H., *Evaluation of Pointing Efficiency on Small Screen Touch User Interfaces*. In: Human Interface and the Management of Information. Information and Interaction. Human Interface, Part II, HCII 2009, LNCS 5618, 2009, pp. 375–384.
- [5] COCKBURN A., AHLSTRÖM D., GUTWIN C., *Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse*. Int. J. Human-Computer Studies, 70, 2012, pp. 218–233.
- [6] BI X., LI Y., ZHAI S., *FFitts Law: Modeling Finger Touch with Fitts' Law*. CHI'13, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2013, pp. 1363-1372.
- [7] LAI C.C., WU C.F., *Display and device size effects on the usability of mini-notebooks (netbooks)/ultraportables as small form-factor Mobile PCs*. Applied Ergonomics, 45, 2014, pp. 1106-1115.
- [8] MACKENZIE I.S., *Fitts' throughput and the remarkable case of touch-based target selection*. Human-Computer Interaction, Part II, HCII 2015, LNCS 9170, 2015, pp. 238–249.
- [9] CONTI A.S., KREMSER F., KRAUSE M., AN D., BENGLER K., *The effect of varying target sizes and spaces between target and non-target elements on goal-directed hand movement times while driving*. Procedia Manufacturing, 3, 2015, pp. 3168-3175.
- [10] YU C., WEN H., XIONG W., BI X., SHI Y., *Investigating effects of post-selection feedback for acquiring ultra-small targets on touchscreen*. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2016, pp. 4699-4710.

- [11] TSAI T.H., TSENG K.C., CHANG Y.S., *Testing the usability of smartphone surface gestures on different sizes of smartphones by different age groups of users*. Computers in Human Behavior, 75, 2017, pp. 103-116.
- [12] FARHAD M., MACKENZIE I.S., *Evaluating tap-and-drag: A single-handed zooming method*. In: Human-Computer Interaction. Interaction Technologies, HCI 2018, LNCS 10903, 2018, pp. 233–246.
- [13] WAWRYNIUK R., *Metodyka oceny jakości działania użytkownika urządzenia mobilnego*. Praca dyplomowa, WAT, Warszawa, 2013.
- [14] *Ergonomic requirements for Office work with visual display terminals (VDTs). Part 9: Requirements for non-keyboard input devices*. ISO 9241-9:2000(E), International Organization for Standardization.
- [15] KOWALSKI L., *Statystyka*. Wyd. Wydział Cybernetyki WAT, BelStudio, Warszawa, 2005.
- [16] CIECIURA M., ZACHARSKI J., *Metody probabilistyczne w ujęciu praktycznym*. Wizja Press&IT, Warszawa, 2007.

Online sources

- [17] <http://www.mathworks.com/>, <http://www.ont.com.pl/>
- [18] <http://statystyka.rezolwenta.eu.org/materialy.html>
- [19] <http://gsmowo.pl/nokia-lumia-800/>
- [20] <https://tech.wp.pl/nokia-lumia-800-6039436541907585c>

Badanie jakości działania użytkownika wykorzystującego urządzenie mobilne

Gesty tap, double tap i flick

STRESZCZENIE: W artykule przedstawiono wyniki badań jakości wykonywania gestów przez użytkowników wykorzystujących urządzenie mobilne. Jako urządzenie mobilne wykorzystano smartfon Nokia Lumia 800. Wyniki badań dotyczą podstawowych gestów tap, double tap i flick i obejmują czas wykonania gestu i precyzję wykonania gestu. Wyniki uwzględniają podział użytkowników na grupy wiekowe oraz grupy używające i nie używające smartfon na co dzień. Przedstawiono porównanie wyznaczonych charakterystyk pomiędzy grupami. Badania dotyczące innych gestów zostaną przedstawione w kolejnym artykule.

SŁOWA KLUCZOWE: gesty tap, double tap i flick, wprowadzanie gestów palcem, urządzenie mobilne, czas wykonania gestu, precyzja wykonania gestu

Paper submitted to the editorial office on: 1.12.2018