

Performance evaluation of gesture-based interaction between different age groups using Fitts' Law

Diana Carvalho
INESC TEC and UTAD
4200-465 Porto, Portugal and
5000-801 Vila Real, Portugal
+351 {222094000, 259350000}
dianac@utad.pt

Luís Magalhães
Centro ALGORITMI, UMinho
4800-058 Guimarães, Portugal
+351 253510180
lmagalhaes@dsi.uminho.pt

Maximino Bessa
INESC TEC and UTAD
4200-465 Porto, Portugal and
5000-801 Vila Real, Portugal
+351 {222094000, 259350000}
maxbessa@utad.pt

Eurico Carrapatoso
INESC TEC and FEUP
4200-465 Porto, Portugal
+351 222094000
emc@fe.up.pt

ABSTRACT

The recent advances made in human-computer interaction have allowed us to manipulate digital contents exploiting recognition-based technologies. However, no work has been reported that evaluates how these interfaces influence the performance of different user groups. With the appearance of multiple sensors and controllers for hand gesture recognition, it becomes important to understand if these groups have similar performance levels concerning gestural interaction, and if some sensors could induce better results than others when dealing with users of different age brackets. In this respect, it could also be important to realize if the device's sensor accuracy in terms of hand / full body recognition influences interaction performance. We compare two gesture-sensing devices (Microsoft Kinect and Leap Motion) using Fitts' law to evaluate target acquisition performances, with relation to users' age differences. In this article, we present the results of an experiment implemented to compare the groups' performance using each of the devices and also realize which one could yield better results. 60 subjects took part in this study and they were asked to select 50 targets on the screen as quickly and accurately as possible using one of the devices. Overall, there was a statistically significant difference in terms of performance between the groups in the selection task. On the other hand, users' performance showed to be rather consistent when comparing both devices side by side in each group of users, which may imply that the device itself does not influence performance but actually the type of group does.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

Interacción '15, September 07 - 09, 2015, Vilanova i la Geltrú, Spain. Copyright is held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-3463-1/15/09...\$15.00
DOI: <http://dx.doi.org/10.1145/2829875.2829920>

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation (e.g., HCI)]: User Interfaces – *evaluation/methodology, input devices and strategies (e.g., mouse, touchscreen)*.

General Terms

Performance, Design, Human Factors.

Keywords

HCI, natural user interfaces, gestural interaction, performance, Fitts' law, Microsoft Kinect, Leap Motion, target acquisition, selection tasks.

1. INTRODUCTION

For many years, the traditional mode of interaction with computers was based on a WIMP interface (Windows, Icons, Menus, Pointing device), which allowed us to interact with the machine via specific pointing devices, usually a computer mouse. With the paradigm shift to Post-WIMP interfaces [5], we have turned towards a user-oriented and task-oriented approach that attempts to simplify the usability of the interface [17], giving preference to the users innate skills [16] and allowing us to take advantage of recognition-based technologies that understand human behaviors, such as gestures or touch. However, there seems to be a constant addition of new modes of interaction without the proper awareness as to which could be the most adequate for different user profiles (e.g., children, elderly users, people with different levels of digital literacy, people with disabilities) and also regarding which types of tasks (selection, insertion, manipulation).

A previous study of ours [3] presented insights on how users interact using different input modalities and which interface holds the best results in terms of usability testing regarding three user groups with different age brackets. Our preliminary findings indicated that the gestural interface presented worse results than the other ones, and we thus acknowledged that the device itself used for the gesture-recognition – the Kinect – could be negatively influencing the user due to the accuracy required by the

task. Different optical sensors, which allow human body acquisition with respected accuracy, have been released and comparable controllers in the same price range include the Kinect and Leap Motion, a sensor with declared sub-millimeter precision that claims to obtain higher levels of accuracy than the Kinect [24]. Therefore we think that it may be important to understand if another device could perform better than the one used in the previous experiment. After all, when it comes to gestural interfaces the precision of the sensor is said to be vital.

In this context, this study aims at throwing some light on the indices of performance of three groups of users with different age-brackets concerning the use of two different gestural sensors. We intended to understand: (1) if the gestural interaction presents significantly better results on specific groups opposed to others with regards to their age; (2) if the devices used and their declared precision influence or not the reaction times and indices of performance in each group.

2. SCIENTIFIC BACKGROUND

2.1 Natural Interfaces

Active and passive input modes of interaction [19] have been implemented in countless fields. However, published studies have not yet provided an understanding of how different user groups perceive distinct tasks and if their performance is directly influenced by the interaction modality.

Little is known about how the different interfaces affect one's performance when it comes to age-related issues. There are no transversal comparisons of different age groups in one same study where more than one natural interface is evaluated. Work has been developed in this area, but not as a systematic approach. Sambrooks and Wilkinson compare gestural, touch and mouse interaction with 15 participants aged between 22 and 34 years old [21]. They reached the conclusion that touch and mouse presented better results, but this interaction performance was not compared between other groups of users and thus it just clarified that the gestural performance was indeed worse than the other interfaces regarding that specific niche of participants. Other studies, as [2, 6, 9], compared the interaction performance when using traditional mouse inputs or touchscreens, but not between other natural recognition-based interfaces as a whole, and they only compare at the most two groups.

Concerning the gestural controllers, tests have been performed in order to understand the accuracy of any given device [20, 24], but they have not been compared amongst different user profiles, such as distinct age groups.

2.2 Fitts' Law: An Overview

Fitts' law is a human performance model that has been widely applied to describe the relationship between movement time, distance and accuracy within specific pointing tasks, as a way to formally evaluate pointing devices [15] and compare their efficiency [23]. The original experimental paradigm Fitts' Law [7] measures the rate of transmission in "bits per second" of a pre-established movement with different indices of difficulty. Also, ISO 9241-9 "Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 9: Requirements for non-keyboard input devices" [10, 11] provides a standard approach for input device evaluation based on Fitts' law. Fitts' performance model can also be applied when comparing and evaluating pointing devices. Task differences, selection techniques, response irregularity and test conditions applied may influence

experimental variations. Understanding these variables increases the validity of across-studies comparisons regarding input techniques [15]. In this sense, researchers use this model to measure multiple movement times and then determine how the different conditions or devices affect the coefficients of the equation, that is, performance.

3. CASE STUDY

The aim of this work is to compare two different gestural input devices in order to understand if different age groups display similar indices of performance between them, and which device could lead to better target acquisition performance rates for different user groups, using Fitts' law.

To accomplish this, we compared the use of Microsoft Kinect and Leap Motion Controller for motion and gesture control with users from three distinct age groups: children, young adults and older-adults. A total of 60 participants who were naïve to the purpose of the experiment took part in this case study, and were grouped by age and device: (1) 20 children from 9 - 12 years old (10 for each device); (2) 20 young adults from 20 - 35 years old (10 for each device); (3) 20 older-adults from 45 - 60 years old (10 for each device).

We intended to emphasize the distinct age groups and work with users that had noticeable differences concerning cognitive performance levels and dexterity and, as such, we tried to select distinct groups. First, we did not consider users younger than 9 years old because: (1) they are not so used to the computer and do not use it on a daily basis, and also they do not have the dexterity capabilities as improved as older children; and (2) these ages are encompassed by the fifth and sixth year of primary school, being these school years the ones that show greater variance in terms of reasoning by the children [12]. As this variance tends to significantly fade towards high school, we selected graduate students, being these users already at a different stage concerning cognitive performance when compared with the group of children. Finally, the group of older-adults consisted of active workers of a secretariat department. This group could present worse dexterity or consider gestures to be more challenging.

Also, we ascertained that all of the participants had the same level of computer proficiency and used the computer on a daily basis. However, when it comes to gestural interaction, this modality was not familiar to the participants, who were conscious of this gesture-based interface mainly due to having seen gaming consoles with support to this modality. In fact, only seven users had already tried gestural interaction once.

The majority of participants were right-handed (right-handed: 58; left-handed: 2), and the setup was calibrated to accommodate each arm as needed: the left-handed participants used the left arm to perform the tests, and the right-handed used the right arm.

3.1 Apparatus

We conducted the experiment in a closed room with artificial light and the tests were performed in a specific setup assembled for the purpose of this research. The system consisted on a 22" screen placed on a desk in front of the user, with a resolution of 1280x800 pixels; a Microsoft Kinect sensor mounted on a tripod behind the screen, about 25 cm above it and facing the user; a Leap Motion sensor placed on top of the desk, between the user and the screen, and facing upwards. The distance between the user and the Kinect sensor was about 75 cm, and was calibrated to see the user's upper body only. The user was seated facing the setup

at all times, and this setup was not altered throughout the different trials. Also, a purpose-built application based on Fitts' law was developed in Python with the support of: the Kivy Framework, TuoKinect and the Leap Motion Python API.

3.2 Experiment Design

The experiment made use of an independent-measures design and the tasks were performed in a controlled environment. Aside from the discrete data collected during the tests, we also gathered qualitative observational analysis on the participants' behavior. At the beginning of each phase of this study, we performed a questionnaire regarding the participants' previous experience with gestural interfaces. Also, at the end of each test we proposed a questionnaire with qualitative Likert Scales [13], in order to understand the users' preferences and their views regarding: ease of use of the device in question, fatigue effect, level of user comfort / frustration, and users' degree of concentration.

The tests were performed on different days and with different users. As such, the participants only tried one of the devices, and not both of them, as we did not intend for the use of one device to influence the movement times recorded on the other due to performance improvement over trials.

In order to understand the users' index of performance when considering target acquisition tasks using gestural interaction with the Kinect sensor versus Leap Motion, we resorted to Fitts' law evaluation paradigm, a frequently used model for measuring movement performance. Fitts quantifies the movement tasks' difficulty, also known as Index of Difficulty (ID), by the metric "bits" and calculates it by using the value of the distance, or amplitude (A), between two specific targets and their width (W) or tolerance area:

$$ID = \log_2(A/W + 1). \quad (1)$$

Although not in its original form, this formulation is a more stable model for the Fitts' Law suggested by Mackenzie [15], since it is resilient to negative values. The Index of Difficulty can be obtained by varying the values of the width of the targets and their distance, thus providing a range of task difficulties. We used six different amplitudes: A = 200, 460, 750, 770, 930, 940 pixels; and five different target widths: W = 16, 26, 50, 120, 200 pixels. These choices attain seven levels of Fitts' Law's index of difficulty, from 1.72 bits to 5.20 bits. The order of appearance of the seven indices of difficulty was randomized. We also included two target directions (left, right) in the study because it could be relevant to understand if the contactless motion was influenced by different directions [4, 8].

In summary, the target acquisition data consisted of: 60 participants divided in three groups; two input devices (Microsoft Kinect; Leap Motion); two target directions (left, right); 25 repetitions per index of difficulty; and seven indices of difficulty (1.72, 2.25, 3.12, 3.14, 4.04, 4.89, and 5.20 bits). Thereby, altogether there were 28 experimental conditions and an overall of 42 000 trials.

The gestural-based cursor illustration was a target badge (Figure 1) grounded on the validated "point and wait" strategy [22] for selection. This cursor was shown on the screen when a pointing gesture was interpreted by the system. The center of the target badge represented the point from which the selection coordinates were referenced by the application. When the system detected the gesture to be motionless, a visual feedback was triggered to make the user aware of the selection progress: a green circle started

growing inside the cursor representation until it reached the whole target and thus the selection was made. Here, the user had to point and wait for 0.8 seconds for the system to recognize the intention.

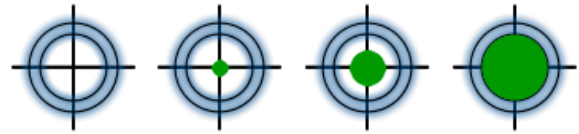


Figure 1. Feedback for the selection strategy.

3.3 Procedure

We analyzed the users' performance considering only movements along the "x" axis and used a reciprocal one-dimensional pointing task based on Fitts' original experiments [7]. This consisted on a horizontal movement between two vertical bars displayed on the screen, representing an initial location and a target (as shown on Figure 2). According to MacKenzie [15], three-dimensional movements may follow the same predictive model as a one-directional task, as a possible substitute for target width when the angle of the approach varies can be the distance through the targets along the approach vector. However, we only intend to understand the one-dimensional movement with homogeneous shapes.



Figure 2. Layout of the targets displayed on the screen

Each test proceeded as follows. At the beginning of the test, the participant would be required to select a "start" button shown on the screen, therefore maintaining a consistent original position. Afterwards, two vertical bars would be revealed: a grey bar representing the initial position, and the green bar being the intended target to be selected. Participants then selected the target bar in each successive trial, which was at all times organized in the opposite direction (from left to right and from right to left). Also, there was a visual feedback to indicate the gesture position was on top of the target object: the bar changed its color to blue. The start time was recorded from each target bar to the next and stopped when the next target was selected. At the end of all of the trials, the bars disappeared and a "successfully completed task" message appeared on the screen.

Participants were given a practice attempt before starting the test. This period consisted in completing a continuous task [15] with 10 consecutive trials, in order for the participant to feel at ease with the interaction device.

4. RESULTS AND DISCUSSION

In this section, we present the results of this case study. We believe that we may be able to throw some light on how or if: (1) the different user groups exhibit significant differences in terms of interaction performance when compared between each other; (2) the accuracy of the devices influences gestural interaction when it comes to target acquisition tasks, concerning users of different age brackets.

4.1 Movement Times and Errors

We present the movement times recorded during each trial of the experimental tests with the Kinect and Leap Motion devices (Figure 3). A boxplot analysis of the data revealed a presence of outliers from standardization failure [18] on some of the movement times regarding the calibration of the equipment, and thus these results were removed in order to prevent distortion of estimates.

In terms of movement times recorded, the device that registered the fastest mean results throughout all the three groups was the Leap Motion: children (1.94 s); young adults (1.76 s); and older-adults (2.24 s). On the other hand, the Kinect sensor registered higher mean movement times: children (2.46 s); young adults (2.03 s); and older-adults (2.64 s).

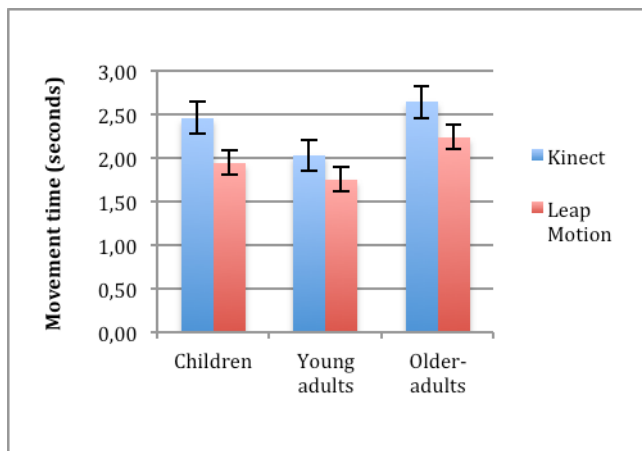


Figure 3. Mean movement times of each group with both devices

We assessed the normality of data with the Shapiro-Wilk Test in order to understand if the data were normally distributed and could thus be considered for statistical analysis. We elected this test over the Kolmogorov-Smirnov due to our narrow sample. As we found that the data were not normally distributed across all of the groups and since we were working with small samples, we performed a non-parametric Mann-Whitney U test to understand if the devices presented significantly different results for each group in terms of movement time. Indeed, every group presented a statistically significant difference between the use of each device. In this regard, the group of children revealed slower movement times using the Kinect sensor, with an average rank of 13.80, while the Leap Motion had an average rank of 7.20, being this difference significant ($U = 17$, $p = .013$). Concerning the group of the young adults, there was also a statistically significant difference between the devices ($U = 23$, $p = .041$), but not as pronounced as the children, being the Kinect the one with the

highest movement times (Kinect, average rank of 13.20; Leap Motion, average rank of 7.80). Also, we determined a statistically significant difference when comparing the older-adults' interaction using both sensors ($U = 21$, $p = .028$), where the Leap Motion remains faster (Kinect, average rank of 13.40; Leap Motion, average rank of 7.60). This may indicate that according to the movement times all the groups display lower results when interacting with the Leap Motion sensor than with the Kinect, which may imply that for gestural interaction concerning selection tasks, the Leap Motion could induce quicker reaction times regardless of the users' age.

Overall, considering a comparison between the three groups, the young adults presented the lowest mean time with both devices, which may indicate that when it comes to gestural interaction the young adults are quicker in completing target acquisition tasks than the other groups. Since our group sample was limited to 10 participants for each group and each device and some data were not normally distributed, a non-parametric independent-samples Kruskal-Wallis 1-way ANOVA was run to determine whether this difference in terms of mean times throughout the groups was indeed significant. The distribution of the times of all the groups using the Leap Motion was considered the same and thus not statistically different ($\chi^2(2) = 5.635$, $p = .060$), presenting a mean rank time of 15 for the group of children, 11.10 for young adults, and 20.40 for older-adults. On the other hand, considering the use of the Kinect, the Kruskal-Wallis test showed that except for the children and older-adults ($p = .875$), there was a statistically significant difference in mean times between the different groups ($\chi^2(2) = 12.742$, $p = .002$), with a mean rank time of 17.30 for the group of children, 7.75 for young adults, and 21.45 for older-adults. Hence, the difference of movement times using of the Kinect was statically significant between the group of children versus young adults ($p = .046$), and the group of young adults versus older-adults ($p = .002$).

In this sense, when using the Leap Motion, the groups did not present statistically different results concerning movement times, which may suggest that their reaction times are not that disparate. However, when using the Kinect sensor the groups show inconsistency with regard to movement times.

Another aspect that we looked at was error rate. We assumed a speed-accuracy tradeoff approach [25] when dealing with errors, as the selection rested on the "point and wait" strategy and thus did not trigger false selections caused by user mistakes. The more precisely the task was performed, the longer it took to be completed, and vice-versa. Here, the selection would only be triggered when the system detected the gesture to be motionless and precisely on top of the target bar. Otherwise, no selection was made.

4.2 Indices of Performance

Figure 4 shows the Index of Performance achieved by the three groups for each interaction device. The chart clearly shows the children as having the worst performance of all the groups, either with the Kinect (IP = 1.64 bits/s) or with the Leap Motion (IP = 2.03 bits/s); followed by the young adults with an IP of 2.98 bits/s with the Kinect, and 2.47 bits/s with the Leap Motion. The group of older-adults presented the highest throughput regarding gestural interaction for target acquisition tasks with both devices: the Kinect with an IP of 3.03 bits/s and the Leap Motion with 3.25 bits/s.

A Shapiro-Wilk Test showed us that the data were not normally distributed and, as such, we performed a Mann-Whitney U test to understand if there were significant differences between each group's performance concerning the device used. None of the groups presented a significant difference between the performance using each device: children ($U = 34$, $p = .226$); young adults ($U = 38.50$, $p = .384$); older-adults ($U = 44$, $p = .650$).

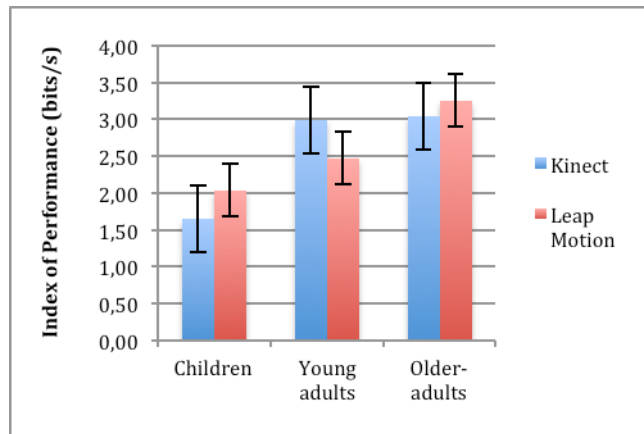


Figure 4. Groups' Index of Performance concerning both devices

An independent-samples Kruskal-Wallis 1-way ANOVA test was also run to determine if there was a statistically significant difference in terms of performance between the groups. Considering the Kinect, the performance of the children versus young adults ($p = .031$) as well as versus older-adults ($p = .013$) was significantly different. Contrarily, the group of young adults and older-adults did not register significant differences ($p = 1.000$). On the other hand, considering the Leap Motion sensor, the only significant difference presented was between the groups of children versus older-adults ($p = .034$), and not between children and young adults ($p = .824$) or young adults versus older-adults ($p = .453$).

We might reach the conclusion that, in terms of performance, neither the Kinect nor the Leap Motion sensor directly influences gestural interaction. However, they appear to have different results regarding users of different age brackets, which implies that these sensors could affect accuracy of interaction depending on the user's age. The group that revealed the best performance results was the older-adults' one, which has already shown to have worse results when it came to reaction times during selection tasks. Here, we may infer that since other pointing devices like the mouse or even touch inputs may request more accuracy than gestural commands, these allow older-adults to overcome some problems of dexterity that may exist in this age bracket, and thus demonstrate improved performances over other groups of users. However, we will not further explore this assumption on the course of this paper. In contrast, children may need more time practicing to overcome the learning curve to better understand the relation between the physical environment of the gesture itself and its visual feedback.

4.3 Participants' Behavior and Feedback

During the experiment, participants shared the same behavior concerning both devices. They were attentive to the task: trying to maintain a certain pattern with the arm's motion and trying to avoid any sudden movements. The most perceptible behavior we noticed was that some of the participants would push the arm forward in order to try to elevate the target badge on the screen, instead of lifting it upwards. This situation would lead the participants to a higher level of frustration with either device because the badge would not go up unless the arm was raised.

Also, at the end of the experiment the participants answered a questionnaire about their preferences. We discuss below the responses with the majority of the preferences. In terms of ease of use, 50% of the children focused their choice in that the Kinect was relatively difficult to use, being the other answers scattered throughout the Likert scale. In contrast, the Leap Motion had 37% of the children saying it was relatively easy. The older-adults also shared this opinion (40% signed the Kinect as relatively difficult; and 45% said the Leap Motion was very easy). Although the results do not show this tendency, some participants felt that the Leap Motion sensor seemed easier to use than the Kinect. The young adults stated that both devices were very easy to use (50% of the responses regarding each device).

On the other hand, in terms of ease of learning and improvement, both children and young adults thought both devices were relatively easy, but the older-adults argued that the Leap Motion was easier to learn than the Kinect. Regarding fatigue, all of the groups thought both devices were relatively demanding except the older-adults' group that considered the Leap Motion less challenging than the Kinect. Finally, all groups considered the Kinect as a device that requires further concentration to complete the trials, as opposed to the Leap Motion that was simpler in terms of attention to the task.

5. CONCLUSIONS AND FUTURE WORK

This investigation was intended to understand (1) if there are age-related differences regarding gestural selection tasks when using one device to the detriment of the other; (2) if two different gestural sensors, the Kinect and the Leap Motion, could differently influence the users' interaction performance and target acquisition times.

We are aware that conducting further tests may help ascertain whether the devices do truly influence the participant's movement times and respective index of performance. Also, it is possible that by continuing the experiment the groups' index of performance would increase, and we therefore need to understand if practice could be considered as a way to attenuate the learning curve, and consider running more tests to set aside potential hardware problems on detection that could be negatively affecting accuracy.

Nevertheless, we acknowledge it appears to be evidence that for selection and target acquisition tasks a gestural interface may not be the best approach, as it shows low indices of performance throughout the three groups. We may presume that for selection tasks that require a higher level of precision neither device displayed proof that it influenced the user to complete the trials with better performance. In fact, the devices' accuracy did not appear to have a direct relation to improved performance.

As stated earlier, mean movement times should not be considered the only variable when exploring the users' performance, and here we can relate to why: for example, when comparing the

movement times and respective index of performance, children exhibited mean times similar to the other groups, but not nearly as good performance as the others. Children have as good response times as other groups, but they do not present a consistent interaction, which may reveal that they are faster but not as precise when interacting with gesture-based interfaces. On the other hand, the group of older-adults was indeed the one that exhibited the best results concerning performance with both devices. However, it was also the one group to present higher movement times, which may indicate that they have a consistent interaction. Regardless of their movement times being not as fast as the other groups, they tend to be more precise throughout the several levels of difficulty of each set of trials.

This is a work in progress, and more tests will be conducted to further attest these findings, but we may suggest that: (1) in terms of index of performance the groups displayed statistically different results: the group of older-adults held the best gestural performance results, and the children the worst, when compared between each other; (2) both devices behave in a similar manner for selection tasks and there are no statistically significant differences concerning their accuracy when comparing each of the three groups individually. However, in terms of user preference, some participants felt that the Leap Motion sensor was easier to use and less demanding in terms of concentration than the Kinect.

Aside from more tests to be conducted, in the future we also intend to broaden our research to other elemental tasks (as insertion and manipulation) using gestural interfaces and observe performance considering our targeted age groups. Also, we intend to follow other evaluation methods and test the 7 measurements of McKenzie [14] and Steering Law [1] for tracking performance evaluation for these pointing devices.

6. ACKNOWLEDGMENTS

We would like to acknowledge the support and contribution of “Universidade de Trás-os-Montes e Alto Douro” and the schools that took part in our study: “Monsenhor Jerónimo do Amaral”, “Escola Secundária Morgado de Mateus”, and the studies center “Super-Heróis”, all in Vila Real, Portugal.

Diana Carvalho has a PhD fellowship granted by FCT – Fundação para a Ciência e a Tecnologia (SFRH/BD/81541/2011).

7. REFERENCES

- [1] Accot, J. and Zhai, S. 1999. Performance evaluation of input devices in trajectory-based tasks. *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI '99* (New York, New York, USA, May 1999), 466–472.
- [2] Bobeth, J. et al. 2014. Tablet, gestures, remote control? *Proceedings of the 2014 ACM international conference on Interactive experiences for TV and online video - TVX '14* (New York, New York, USA, Jun. 2014), 139–146.
- [3] Carvalho, D. et al. 2014. Different interaction paradigms for different user groups: an evaluation regarding content selection. *Proceedings of the XV International Conference on Human Computer Interaction - Interaccion '14* (Puerto de la Cruz, Tenerife, Spain, Sep. 2014), 40.
- [4] Cockburn, A. et al. 2012. Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. *International Journal of Human-Computer Studies*. 70, 3 (Mar. 2012), 218–233.
- [5] Dam, A. van 1997. Post-WIMP user interfaces. *Commun. ACM*. 40, 2 (1997), 63–67.
- [6] Findlater, L. et al. 2013. Age-related differences in performance with touchscreens compared to traditional mouse input. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (New York, New York, USA, Apr. 2013), 343.
- [7] Fitts, P.M. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*. 47, 6 (1954), 381–391.
- [8] Goldberg, K. et al. 2013. *A New Derivation and Dataset for Fitts' Law of Human Motion*.
- [9] Inai, Y. 2013. Age-related differences in pointing movements in restricted visual tasks and their design implication. *2013 ICME International Conference on Complex Medical Engineering* (May 2013), 439–443.
- [10] ISO 2007. *ISO 9241-400:2007 - Ergonomics of human--system interaction -- Part 400: Principles and requirements for physical input devices: International Organization for Standardization*.
- [11] ISO 2000. *ISO 9241-9 International standard: Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 9: Requirements for non-keyboard input devices: International Organization for Standardization*.
- [12] Lemos, G.C.E.M.P. de 2006. *Habilidades cognitivas e rendimento escolar entre o 5.º e 12.º anos de escolaridade*. Universidade do Minho.
- [13] Likert, R. 1932. A technique for the measurement of attitudes. *Archives of Psychology*. 22 140, (1932), 55.
- [14] MacKenzie, I.S. et al. 2001. Accuracy measures for evaluating computer pointing devices. *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '01* (New York, New York, USA, Mar. 2001), 9–16.
- [15] MacKenzie, I.S. 1992. Fitts' Law as a Research and Design Tool in Human-Computer Interaction. *Human-Computer Interaction*. 7, 1 (Mar. 1992), 91–139.
- [16] Mayer, R.E. and Moreno, R. 2003. Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*. 38, 1 (Mar. 2003), 43–52.
- [17] Nielsen, J. 1993. Noncommand user interfaces. *Communications of the ACM*. 36, 4 (Apr. 1993), 83–99.
- [18] Osborne, J.W. and Overbay, A. 2004. The power of outliers (and why researchers should always check for them). *Practical Assessment, Research & Evaluation*. 9, 6 (2004).
- [19] Oviatt, S.L. 2007. Multimodal Interfaces. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, Second Edition*. A. Sears and J.A. Jacko, eds. CRC Press. 413–432.
- [20] Ren, Z. et al. 2011. Robust hand gesture recognition with kinect sensor. *Proceedings of the 19th ACM international conference on Multimedia - MM '11* (New York, New York, USA, Nov. 2011), 759.
- [21] Sambrooks, L. and Wilkinson, B. 2013. Comparison of gestural, touch, and mouse interaction with Fitts' law. *Proceedings of the 25th Australian Computer-Human Interaction Conference on Augmentation, Application, Innovation, Collaboration - OzCHI '13* (New York, New York, USA, Nov. 2013), 119–122.

- [22] Schapira, E. and Sharma, R. 2001. Experimental evaluation of vision and speech based multimodal interfaces. *Proceedings of the 2001 workshop on Perceptive user interfaces - PUI '01* (New York, New York, USA, Nov. 2001), 1.
- [23] Soukoreff, R.W. and MacKenzie, I.S. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *International Journal of Human-Computer Studies*. 61, 6 (Dec. 2004), 751–789.
- [24] Weichert, F. et al. 2013. Analysis of the accuracy and robustness of the leap motion controller. *Sensors (Basel, Switzerland)*. 13, 5 (Jan. 2013), 6380–93.
- [25] Zhai, S. et al. 2004. Speed–accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. *International Journal of Human-Computer Studies*. 61, 6 (Dec. 2004), 823–856.