Age group differences in performance using diverse input modalities: insertion task evaluation

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

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

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

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

ABSTRACT
Novel input modalities such as touch, tangibles or gestures try to exploit human’s innate skills rather than imposing new learning processes. However, no work has been reported that systematically evaluates how these interfaces influence users’ performance, that is, assesses if one interface can be more or less appropriate for interaction regarding: (1) different age groups; and (2) different basic tasks, as content insertion or manipulation. This work presents itself as an exploratory evaluation about whether or not the users’ efficiency is indeed influenced by different input modalities and age. We conducted a usability evaluation with 60 subjects to understand how different interfaces may influence the speed and accuracy of three specific age groups (children, young adults and older adults) when dealing with a basic content insertion task. Four input modalities were considered to perform the task (keyboard, touch, tangibles and gestures) and the methodology was based on usability testing (speed, accuracy and user preference). Overall, results show that there is a statistically significant difference in speed of task completion between the age groups, and there may be indications that the type of interface that is used can indeed influence efficiency in insertion tasks, and not so much other factors like age. Also, the study raises new issues regarding the “old” mouse input versus the “new” input modalities.

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
Human-Computer Interaction, Natural User Interfaces, Interaction Paradigms, Input Modalities, Mouse, Keyboard, Touch, Tangibles, Gestures, Insertion Task.

1. INTRODUCTION
Interaction paradigms are shifting from an era of WIMP interfaces (Windows, Icons, Menus, Pointing device) to Post-WIMP [11], where the systems adopt a user-oriented and task-oriented approach, attempting to improve the usability of the interface [27] and allowing users to take advantage of recognition-based technologies that understand complex human behaviors, such as speech, eye gaze, body language or gestures.

Over the past decade, many input modalities have emerged to help users take advantage of their innate skills and avoid new learning processes. Interfaces based on gesture recognition, touch or tangible objects are considered natural and intuitive due to their low cognitive load requirements [3]. In addition, they have become increasingly more flexible regarding the users’ abilities and/or preferences, such as age, skill level, cognitive profiles, sensory and motor impairments, native language, or temporary illness [28]. However, although there has been a continuous addition of new input modalities in our daily lives, there is not yet a full conscience as to which could be the most adequate for different age groups in their everyday tasks.

Also, the set of actions that can be performed by the user in a digital environment encompasses selection, insertion and manipulation of content, and each of these can be influenced by their supported sense of naturalism, efficiency and degrees of freedom (DOF) provided to the user [6]. Thus, a systematic study that reveals the relation between input modalities and interaction tasks for different age groups is of paramount importance, i.e. understanding which interfaces could be considered best or worst for specific tasks.

Our deeper research approach aims to understand at which point the constant appearance and addition of novel input modalities leads to a weaker performance by the users. In a previous study of ours [7], we believe we were able to provide insights on how three
groups of users (children, young adults and older-olds) complete selection tasks using different input modalities, and which ones hold the best results in terms of usability. As we could see, the results fluctuated and we could understand the need for using specific inputs when dealing with selection operations. As such, we believe in the importance of continuing this analysis and take yet another step in our research: in this case, insertion tasks.

In this paper, we intend to recognize what is the most efficient interaction paradigm for specific target-audiences when performing one of the most basic tasks previously stated – data insertion. After presenting a brief outline of the background on the different input modalities and presenting examples of related work, we describe the methodology used for this preliminary study and discuss its results. Finally, we conclude this paper by revealing some of our future work in the field.

2. BACKGROUND

Human-computer interaction (HCI) is an area of research that began initially as a specialty area in the field of computer science, embracing cognitive science and human factors engineering [33].

Indeed, HCI research is concerned with the design, implementation and evaluation of interfaces on different contexts regarding the users’ task and work at hand, aiming at an easy and efficient use of the interactive systems. The main goal of this field is to understand how interfaces can allow users to take advantage of high performance computing without retracting their focus from their work, rather than making them learn technical details necessary to handle the interfaces and complete the tasks [23].

In recent years, interactive computer-based systems have become tools for communication, collaboration and social interaction amongst diverse user populations with different abilities, skills, disorders, requirements and preferences in a variety of contexts of use [10]. As such, the needs of the users are becoming increasingly important and computers are considered integrated environments that should be accessible and usable by anyone, anytime, anywhere, ensuring the safety, utility, effectiveness, efficiency, accessibility and usability of interactive systems by all [34]. There cannot be only a concern of usability of interfaces by elderly people and individuals with disabilities or special needs, but it should have a more ample connotation, embracing all individuals with different levels of abilities, skills, requirements and preferences to be able to access information technologies [10].

However, published studies have not yet provided an understanding of how different user groups perceive distinct elemental tasks and if their performance is directly influenced by the interaction modality. It thus becomes important to acknowledge that different user groups interact with technology in different ways, and thereby various means of interaction might be needed and desired.

The human being has passed several technological revolutions over time, not continuously in periods of cyclical evolution, but irregularly. Currently, society witnesses a new technological revolution regarding the interaction paradigm: natural interfaces. These allow for a direct interaction, taking advantage of human senses for a better human-computer communication. Instead of having a technology-centered context, we now turn to a human-centered one, where technology can “understand” the user and his/her context of use [18].

Contrarily to previous technological generations of interfaces, as “point-and-click” WIMP (graphical user interfaces based on window, icon, menu, pointing device), this paradigm does not have a pointing device as a demand, such as the computer mouse / keyboard. Instead, it is based on users’ innate behaviors and abilities, such as gestures or touch, for specification and execution of commands and tasks. This new era is also known as Post-WIMP, due to its status post-graphical interfaces, and focuses on the users’ needs and skills, facilitating the usability of the interface [4]. Natural interfaces encourage the use of the users’ own senses to communicate with the machine. The system should prioritize the most basic means of communication people learn since birth (speaking, gesturing, facial expression, and other forms of human communication), to the detriment of an interaction that requires third-party devices, unfamiliar to our innate skills, and forces new learning processes [17].

However, there seems to be a constant offering of new modes of interaction without the proper awareness as to which is 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 distinct types of tasks (e.g. elemental or compound). Indeed, little is known about how the different interfaces affect users’ performance when it comes to age-related issues. It is important to understand these distinct interfaces individually, i.e. their characteristics, benefits and limitations, in order to grasp the advantages they may bring to the community in virtue of a better human-computer interaction.

Work has been developed in understanding how different natural interfaces affect performance, but not as a systematic approach. There are no transversal comparisons of different age groups in one same study where more than one natural interface is evaluated for performing specific tasks, except for a previous exploratory study on discrete selection activities [8], where the research had focus on understanding which input modality was the most efficient for a specific target-audience in a selection task. Here, three groups of users were considered (children, young adults and older-olds), but only discrete tasks were performed and no other elemental activity was evaluated other than selection.

The majority of studies that embraces more than one modality [5, 12, 16, 31] reach the conclusion that touch presents better results, and gestural performance is worse than the other interfaces regarding specific niche of participants, but these interfaces are not compared between more than two groups of users, and most compare interaction performances when using traditional mouse inputs or touchscreens, but do not make comparisons with other natural recognition-based interfaces.

Indeed, studies are still immature when it comes to stating differences in terms of efficiency, effectiveness, and performance between users with distinct attributes, namely age.

Therefore, this situation is our point of action, where we intend to conduct a systematic study in order to understand the differences in performance of different age groups.

3. CASE STUDY

The aim of this work is to understand if diverse input modalities may cause significant differences in interaction regarding elemental insertion tasks. As such, we focused on a usability evaluation considering the participants’ effectiveness (error rate), efficiency, and user preference. We intend to recognize: (1) the relation users may have with the different interfaces on a content
insertion activity; and (2) if age may affect interaction with the system.

In this study we work with the keyboard, touch, gestures and tangibles as input modalities. First and foremost, we chose to work with the graphical interface because it was our plan to understand if this was still suitable for any elemental task. Also, we had interest in knowing if our target-audience would present similar performances with regard to this traditional interface. We selected touch as another interaction modality, since it has proven its benefits considering different contexts of use and fields of study, whether it is in education [21, 24, 38], health [25, 39], working with users with special needs [15] or elderly users [9, 26]. Regarding the tangible interface, it could be interesting to analyse how different users from distinct age groups behave using physical objects that are able to couple digital information and consequently eliminate the conceptual gap from input / output of data [35]. Lastly, we considered gesture-based interfaces for this study because of their potential when it comes to human natural communication. Indeed, a lot of information transmitted amongst humans is passed through gestures [13] and this interaction modality can thus be explored by different fields of study and regarding several purposes: attentive and immersive environments [29, 36], education [19] and as alternative communication systems for people with disabilities or impairments [20, 30].

3.1 Participants

60 participants, who were naïve to the purpose of the experiment, took part in the experiment and were divided into three age groups:

• Children: 20 participants of 9 to 12 years old ($M = 10.5; SD = 1.04$).
• Young adults: 20 participants of 20 to 30 years old ($M = 25.6; SD = 4.09$).
• Older adults: 20 participants of 45 to 60 years old ($M = 49.2; SD = 4.19$).

For the first group, we chose children attending the basic school and worked with three different schools in the city of Vila Real. Regarding the group of young adults, we restricted our sample to participants that were students and used the computer on a daily basis. As such, we selected subjects attending courses related to computer science. It was our purpose to limit the age brackets, but we did not intend to include people of other fields that would not require everyday computer use. This case helped us control our human sample in terms of digital literacy. Finally, for the third group, we chose older adults that worked with the computer mouse as the primary professional activity and thus contemplated workers of secretariat departments of two different schools in the city of Vila Real. Of all of the participants, four were left-handed.

3.2 Experiment Design

The experiment used a complex design: regarding the evaluation within groups using the different input modalities, we used a repeated-measures within-participants design, whereas regarding the evaluation between age groups we resorted to a repeated-measures between-participants design. This is due to this study having two independent variables to study: age group and interaction modality. Altogether there were 720 experimental trials during the study. Each of the 60 participants completed four required tasks randomly, three times each, which gives a total of 12 trials per subject. The aim of each task was to use one of the four input modalities available (keyboard, touch, tangibles, gestures) to insert the requested content. For the usability testing, we followed two methods of evaluation: a quantitative one with specific performance metrics – effectiveness (number of errors) and efficiency (time to successfully complete the task) – and a qualitative one through user preference and observational analysis on the participants’ behaviour. At the beginning of each test, we performed a minor survey to determine the participants’ previous experience with the interfaces at issue. Also, at the end of each test we proposed a questionnaire with closed-end questions, qualitative Likert Scales (Likert, 1932) and ranking lists, in order to understand the users’ preferences and their views regarding: ease of use of the input modalities, ease of learning, fatigue effect, naturalism of interaction, level of user comfort / frustration, and user’s degree of presence and concentration.

3.3 Apparatus

We conducted the study in a closed room, and the tests were performed in a specific setup assembled for the purpose of this research. The system (Figure 1) consisted of: a 22” touchscreen placed in front of the user, with a resolution of 1280x800 pixels; a Leap Motion sensor placed on top of the desk, between the user and the screen, and facing upwards; a physical keyboard placed on top of the desk and in front of the user, next to the leap motion; a webcam; and 10 tangible pieces located in front of the screen, following the specific order of the numbers (from 1 – 0) and with even spaces in-between them.

![Figure 1: Setup assembled for the insertion task](image)

A purpose-built application was created with support for all modes of interaction, in order to keep the testing environment as coherent as possible. The tasks’ software was developed in Python with the support of: the Kivy Framework, the open-source computer-vision Framework reactIVision (to track the fiducials) and the Leap Motion SDK.

Also, in order to reduce the impact of the system’s feedback across the different input modalities, adjustments were made to the application: the gestural-based pointing illustration was a target badge (Figure 2) grounded on the validated “point and wait” strategy for selection [32]; and the touch input underwent a slight adjustment made to the cursor position on the screen in order to avoid the occlusion effect [37]. Therefore, for touch inputs we calibrated the system to register a contact point at a corrected position of minus 1 mm on the x axis, a slight change that enhanced the accuracy of the users for low objects’ widths, as was understood after a preliminary pilot study.
3.4 Procedure

First, in the beginning of each test, we provided the users with an overview about the range of available input modalities: keyboard, touch, tangibles and gestures. Participants were instructed about the goal of each activity at hand: correctly insert a code provided on the screen in the beginning of each trial.

We divided the task into four random activities, each making use of one interaction modality. Over the course of the experiment, the researcher was always present in order to explain any doubts that could occur and respond to any questions the participants could have. Also, we conceded training trials in order for the participants to adapt to the interface and understand their reaction time.

Each activity would start with a countdown from 3 to 1 to prepare the participant for the test, and then an input field would appear on the screen, along with a numerical code. When using touch or gestures for interaction, a soft-keyboard would also appear on the bottom of the screen (Figure 3b).

At this moment, participants were requested to insert the provided code using one of four input modalities, according to what was requested, as quickly as possible while avoiding any errors. Also, the four activities were randomly placed.

As such, touch and gestures made use of the soft-keyboard in the screen to insert the numbers. On the other hand, regarding the graphical interface, users would insert the code with the physical keyboard, and with the tangible interface they used the numbered pieces on top of the table. In this case, when a participant made an error and inserted the wrong digit, he / she could delete it using a specific black piece. All the other pieces were white, as illustrated in Figure 1.

When a digit was inserted, it would appear inside the text field and the user would keep typing until the three numbers were inserted. The user did not need to select any buttons to submit, as we wanted this task to remain elemental and not composed in order to control the variables of the experiment. When the last digit was inserted, the system would automatically check if the code was correct. If so, it would automatically advance to the next trial, until it reached the third and last trial, and then a “successfully completed” message would appear on the screen, after which the user could continue to the next activity. If the code was incorrect, the digits inside the text field would turn red and the trial would not advance until the code was corrected. For that, participants could use the “delete” button on the soft and physical keyboards. We registered an error each time the delete button was used.

There were no time constraints to successfully complete the tasks, since this could cause stress and influence performance. However, the participants were asked to complete the activities as fast as they could. Also, the tests could be paused at any moment between the tasks in order for the participants to rest. At the end of the tests, participants were asked to fill in a survey about their preferences.

Each experimental session lasted approximately 15 minutes.

4. RESULTS AND DISCUSSION

Below are presented and analysed the results regarding the usability evaluation of a basic insertion task: effectiveness (error rate and completion success), efficiency (completion time) and satisfaction.

4.1 Effectiveness

We considered an error when the wrong number was inserted in the system. The group of the older-adults registered the highest total number of errors, but especially when using gestures, which may imply that over the years, their dexterity decreases.

Figure 4 presents the errors registered during the experiment. In total, we registered 102 errors: 3 with the physical keyboard as the input device, 6 using touch, 10 using tangibles and 83 resorting to gestures for content insertion, being most of these errors committed by older-adults.

4.2 Efficiency

We analyzed the completion time within each group of participants and also between them, considering the different input modalities. We detected the presence of outliers per input modality on the values regarding time spent to complete the tasks, and thus these results were removed in order to prevent distortion of estimates in the statistical analysis. We followed the outlier-labeling rule for this analysis [2, 14].
4.2.1 Overall results

The task completion time was measured between the appearance of the input field displayed on-screen and the successful insertion of the code. The input device that registered the fastest mean results throughout all the three groups was the physical keyboard: children (1.28 s); young adults (0.61 s); and older-adults (0.95 s). On the other hand, the gestural interface registered the highest mean completion times: children (2.91 s); young adults (1.81 s); and older-adults (6.09 s). The lowest time recorded was by young adults using the keyboard (0.21 s) and the highest time was registered by older-adults using gestures (19.85 s). Notice also that the mean minimum time recorded using gestures (0.69 s) was actually very close to the mean maximum value documented using the keyboard (0.96 s).

Figure 5 presents the tasks’ mean completion times recorded during the experimental tests with the different interfaces by each age group.

![Figure 5: Mean time taken to complete the task (in seconds)](image)

Overall, the children’s group presented the highest mean times in most of the interfaces, comparing to the other groups, except regarding the gestural input modality. In contrast, the group of young adults performed better with all interfaces and thus showed the lowest mean completion times. It may be also important to highlight the completion time recorded by the older-adults with the gestural interface, which was about twice the one of the children’s group.

Nevertheless, the mean insertion time for the several input modalities did follow a pattern throughout the three user groups: (1) the gesture-based interface was not as efficient as the rest, especially considering the group of older-adults; and (2) using the physical keyboard was the fastest. It should be noticed that the soft-keyboard was the second best regarding completion times, but for an elemental insertion task the use of a physical keyboard was more efficient amongst the three groups in terms of mean time.

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 the size of our sample. Except for the gesture-based interface used by the older-adults (p = .008), the data were normally distributed. Indeed, there was a statistically significant difference between all the groups, and for each interface, as determined by one-way ANOVA: graphical interface (F(2, 57) = 17.539, p = .000); touch interface (F(2, 57) = 12.526, p = .000); tangible interface (F(2, 57) = 10.821, p = .000); gestural interface (F(2, 57) = 16.053, p = .000).

Regarding the use of the physical keyboard, a Tukey post-hoc multiple comparisons test showed that the mean completion times were significantly different between children and young adults (p = .000), children and older-adults (p = .012), young adults and older-adults (p = .013). As for the soft-keyboard, the results were also significantly different: children versus young adults (p = .000), children and older-adults (p = .037), young adults and older-adults (p = .042). Here, we can see that the difference between children versus young adults using the touch interface is higher than for the other combinations, which may imply that these groups tend to have more disparate performance results regarding interaction.

However, not all the groups showed a significant difference in the results considering the tangible and gestural interfaces: the tangibles only presented a significant difference between children and young adults (p = .000), and young adults versus older-adults (p = .032); and a not significant difference between the children and older-adults (p = .110). As for the gesture-based interface, the results were significantly different the group of children versus children and older-adults (p = .000), and young adults versus older-adults (p = .000). On the other hand, the children’s group and the young adults did not present statistically significant differences regarding the use of the gestural interface (p = .349). To sum up, there are not significant differences between the children versus older-adults regarding the tangible interface, and between the children versus the young adults regarding the gestural interface.

We have reached the conclusion that for elemental insertion tasks the usage of different interfaces may influence the users’ performance and preference. The pattern for the interface that enables the fastest or the slowest completion times appears to be constant (being the physical keyboard the first and gestures the latter), although displaying significantly different times depending on the interface. This may imply that the type of interface that is used can influence efficiency in insertion tasks, and not so much other factors like age, as seen with this experiment. Additionally to comparing the performance between each group regarding the different interfaces, we also analysed the performance within each group resorting to a within-participant repeated-measures design.

4.2.2 Children

A repeated measures ANOVA with a Greenhouse-Geisser correction determined that there was a significant difference between the mean time taken by children to complete the task in each interface (F(2.393, 45.470) = 23.493, p < .000). The input device with the fastest results was the physical keyboard (1.28 s), followed by touch (1.61 s), tangibles (2.41 s) and finally gestures (2.91 s). Post-hoc pairwise comparisons tests using the Bonferroni correction showed that, except for the physical versus soft keyboard (p = .305) and tangibles versus gestures (p = .534), the mean times were statistically significantly different throughout the other comparisons. Using the graphical interface against tangibles (p = .000) and gestures (p = .000) was indeed significantly faster; as was the touch interface against these same interfaces: tangibles (p = .000) and gestures (p = .000). To sum up, the physical and soft keyboards display significantly faster results for insertion tasks regarding the group of children, but do not show much of a discrepancy between each other. Although the graphical interface is indeed faster for insertion tasks (1.28 s), this difference in mean time is not significant comparing to touch.
4.2.3 Young adults
As the Mauchly’s sphericity was not assumed, like in the children’s group, we applied the Greenhouse-Geisser adjustment to the repeated measures ANOVA. This group also presented statistically significantly different completion times ($F_{2, 381, 45.241} = 95.497, p < .000$). Post hoc tests using the Bonferroni correction revealed that the mean times returned by the young adults using the tangible interface compared to the gesture-based one (1.62s and 1.81s, respectively) showed no statistical difference ($p = .491$). However, there were significant differences in using the other interfaces ($p = .000$), being the graphical interfaces the fastest for this insertion tasks (0.61 s), followed by touch (0.92 s), as observed in the group of children.

4.2.4 Older-adults
The older-adults’ efficiency also achieved a statistically significant difference between the mean times in each interface ($F_{1.034, 19.643} = 26.766, p < .000$). Although the analysis returns a statistically significant difference regarding the use of tangibles versus gestures ($p = .001$) and the other interfaces ($p = .000$), using the physical or soft keyboard is near to not being significantly different ($p = .455$). In this situation, more tests would have to be performed to really understand this correlation between the soft and physical keyboards. The graphical interface scored the lowest mean time (0.95 s), followed by touch (1.26 s), tangibles (2.06 s) and gestures (6.09 s). Indeed, the gesture-based interface presented much higher results, proving to be the worst option for older-adults when it comes to insertion tasks.

4.3 Participants’ Preferences
At the end of the experiment, we collected the participants’ preferences with a questionnaire. In terms of ease of use, all of the groups reported they preferred touch, followed by the keyboard, which they mentioned to be extremely easy to use. Also, children thought tangibles were extremely easy, and all groups stated that gestures were relatively difficult to handle. According to the concentration necessary to complete the tasks, children thought that tangibles were the most challenging, but they chose it to be their favorite input modality.

All groups shared the same opinion: touch was the modality they preferred and gestures was the least liked modality.

5. CONCLUSIONS AND FUTURE WORK
With this study, we wanted to understand which input modality could induce better results considering basic content insertion. Indeed, our findings showed that: (1) more errors were detected when using gestures, for all age groups; (2) the physical keyboard proved to be more efficient, followed by the soft-keyboard using touch; (3) although tangibles proved not to be as efficient as other interaction modalities, children had a great response in terms of preference, which may suggest that this type of input could be interesting to use with this age group not because of efficiency results, but because this group had such an empathy with the pieces, which could eventually influence and improve their interest in the task at hand.

Again, this is a work in progress, and more tests will be conducted to further attest these findings, but we believe that we have indications that the type of interface that is used can influence efficiency in insertion tasks, and not so much other factors like age. The physical keyboard was the best in this specific insertion task and all groups performed better with the physical keyboard.

In the future, we intend to conduct more tests and also broaden our research to another elemental task (manipulation) using the different interfaces and observe performance considering our targeted age groups. Indeed, we believe to be important to understand how the different age groups relate to linear and circular manipulation tasks. Also, aside from the usability evaluation of these interfaces, we intend to follow other evaluation methods for tracking performance evaluation of these devices and test both the Fitts’ and Steering Laws [1, 22].

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).

This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT – Fundação para a Ciência e Tecnologia within the Project Scope: UID/CEC/00319/2013.

7. REFERENCES
‘Connecting’ to the information society: a European perspective. Technology and Disability. IOS Press.


