

Research Article

Comparing the Performance and Evaluation of Computer Experts and Farmers when Operating Agricultural Robots: A Case of Tangible vs Mouse-Based UIs

Andreas Mallas ¹, Maria Rigou ², and Michalis Xenos ¹

¹Department of Computer Engineering and Informatics, University of Patras, 26504, Greece

²Department of Management Science and Technology, University of Patras, 26334, Greece

Correspondence should be addressed to Maria Rigou; rigou@upatras.gr

Received 16 June 2022; Revised 21 July 2022; Accepted 25 July 2022; Published 17 August 2022

Academic Editor: Zheng Yan

Copyright © 2022 Andreas Mallas et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Tangible user interfaces (TUIs) integrate computer systems with the physical world as they allow interaction through physical objects and the human kinesthetic system. They have been studied in various domains, both by the designing and piloting of innovative applications, as well as by comparatively analyzing TUI against GUI interaction. In the human-robot interaction (HRI) field, TUIs are considered a promising approach as they make effective use of physical object affordances, but research is rather inconclusive about whether TUIs offer improved efficiency, lower error rates, increased intuitiveness, engagement, and user satisfaction. In this study, two prototype UIs were designed and evaluated to compare a TUI with a 2D mouse-based UI for remotely operating two agricultural robots used for vineyard spraying. Two different user groups played the role of the operator: computer experts and farmers leading to an experiment with a 2×2 setup, where two different types of UIs were evaluated by two different user groups. The formulated research questions concern the efficiency, accuracy, and user evaluation for each UI and for each group individually and combined. Analysis has shown that there were no statistically significant differences comparing the alternative UIs for each group in terms of time to complete the task, even though computer experts were faster, as expected. Also, the number of collisions, as well as the percentage of unsprayed and double sprayed area, revealed no significant differences, either for user groups or UIs. The TUI received more positive evaluations in terms of user preference, and users reported lower perceived error rates, especially in the case of farmers, who were also more willing to use the TUI in their daily job.

1. Introduction

In all interactive systems, the design of the graphical user interface (GUI) and the selection of the interaction mode in terms of input and output devices are crucial for determining the usability as perceived by users. This also stands for human-robot interaction (HRI) and the field of remotely controlling multiple autonomous robots in scenarios with a high human operator intervention ratio. There have been empirical studies concerning design guidelines for HRI applications [1] targeted at increasing operator awareness of robots and their surroundings, and more recent studies also address the use of various alternative input devices (mouse, haptic, gestures, tangibles,

etc.) [2, 3] in an effort to investigate specific design guidelines while assuring their compliance with more general HCI design guidelines that have long proved their value and practical applicability for guaranteeing effectiveness, efficiency, satisfaction, error tolerance, and learnability [4].

Remote operator control in HRI settings has traditionally been based on 2D user interfaces (UIs) using mouse and keyboard for input, but there are certain limitations inherent in this approach (despite the familiarity of the average computer user with these devices): The motor skills required for efficient use of a mouse and mainly a keyboard are not intuitive to learn and it takes considerable practice and effort to type fast without looking at the keys [5]. In such

cases, there is a substantial amount of time that the attention of the operator is drifted away from the robot control task at hand hindering overall performance. Moreover, the typical 2D representation used on trivial UIs limits people's spatial abilities when controlling robots that move in the 3D environment or interacting with three-dimensional objects. One of the prevailing approaches to overcome these limitations of traditional UIs is tangible user interfaces (TUIs). A TUI (initially referred to as "graspable user interface") is defined as "... a UI in which a person interacts with digital information through the physical environment ... taking advantage of the human ability to grasp and manipulate physical objects and materials" [5]. The position of a physical object in relation to its surroundings along with the spatial orientation gives the human operator intuitive interaction insight and task awareness: We easily interact with physical objects, and there is no need for instruction, training, specific knowledge, or memorization to be able to move and manipulate a physical object in a physical environment [6, 7]. TUIs in a way allow for merging the computational and physical worlds as they make it possible to interact with the digital artifacts through the human kinesthetic system [8]. TUIs bring interesting potential to various domains including the human-robot interaction (HRI) field which is the focus of this study. Our aim is to compare a TUI and its mouse-based variation in terms of efficiency, accuracy, and perceived user satisfaction. The pilot domain of the study is vineyard spraying with remotely operated agriculture robots. To gain more insight and look into the possible effect of domain expertise and computer experience, we have tested the two UI variations with two groups of users: computer experts and domain experts (i.e., farmers in our case). The group of computer experts has no prior experience with the farming domain and vineyards, while the farmers had very limited experience with computers.

2. Literature Review

TUIs have been investigated from several perspectives: from designing and piloting innovative applications to comparative analysis of interacting with TUIs against interacting with GUIs. Zuckerman and Gal-Oz compared similar TUI and GUI versions of a modeling and simulation system [9]. Results showed that most users preferred the TUI version over the GUI version even though TUI version was inferior to the GUI version in terms of usability and both versions were equivalent in task completion time and performance quality. The TUI version was preferred due to its physical interaction, rich feedback, and realism that was highly stimulating and enjoyable. Besançon et al. [2] evaluated the comparative performance and usability of mouse-based, touch-based, and tangible interaction for manipulating objects in a 3D virtual environment. Melcer et al. presented a comparison between the efficacy of tangible and mouse design approaches for improving key learning factors in educational programming games. The results showed that while both game versions were successful at improving programming self-beliefs, the tangible version was considered as more enjoyable [10]. In fact, many research approaches focus on using TUIs for educational purposes based on the assumption that they can provide hands-

on experience, which may have positive learning outcomes [11–14], for instance, when manipulating 3D chemical molecules [15] or learning heart anatomy [16]. Sapounidis and Demetriadis made a comparative study of children's preferences regarding the use of a tangible and an isomorphic graphical interface to program a robot, and they concluded that the tangible interface was more attractive especially for girls, more enjoyable, and easier to be used by younger children. On the contrary, older children, who were more experienced with computers, considered the graphical system as easier [17].

Tsimplinas et al. compared two isomorphic interfaces (graphical/tangible) in the domain of introductory programming for children [18]. Results showed that although no difference between the two interfaces recorded, students' perceived impression on retention was in favor of the tangible interface, which was also perceived as more playful by all students and more appropriate for collaborative work by elder students and girls. Nathoo et al. explored the use of tangible user interfaces for teaching concepts related to internet of things focusing on usability and learning effectiveness [19]. Results revealed a positive score for the usability of the TUI solution, and knowledge gains were significantly higher for students who learnt IoT concepts through the TUI-based system. Nathoo et al. [20] study the usability of a TUI system for teaching basic java programming concepts, by evaluating a developed prototype through the system usability scale (SUS) with results revealing that the system is acceptable despite the identified limitations [20]. Nevertheless, there are reservations about the features of TUIs that offer a learning advantage over a virtual material equivalent [13].

TUIs have also been investigated in the gaming domain as they can provide new levels of immersion and intuitiveness of interaction. Campbell and Carandang compared the TUI and GUI versions of the same tower defense game and reported that users performed better with the GUI and found it easier to use, but the TUI was more interesting and enjoying [21]. Menestrina et al. examined tangible interfaces applied to video games as compared to graphical interfaces. Results suggested that tangible interfaces provide a higher level of sensory and imaginative immersion, competence, positive affect, and experience. On the other hand, there was no significant impact on flow, challenge, negative affect, and tiredness. More recently, there have been research efforts to deploy TUIs in more complex and abstract domains [22]. De Raffaele et al. proposed an active TUI framework for teaching and learning artificial intelligence in universities. The comparison of the TUI approach with previously adopted educational software highlighted the potential of the TUI framework to augment students' gain in knowledge and understanding of abstracted threshold concepts in higher education [23]. Another domain for experimentation of TUIs is interacting with museum exhibits. For instance, Ma et al. compared the behavior of museum visitors at an interactive exhibit that used physical versus virtual objects to explore the visualized distribution of phytoplankton in oceans on a multi-touch table. The findings suggest that the physical controls (rings) better afforded touching and manipulations, which were prerequisites to further exploration, but they detected no measurable differences in

the thoroughness of visitors' interactions, the questions they asked, or on-topic talk with others at the exhibit [24].

TUIs have been investigated in the usability domain as a promising way to foster increased accessibility to groups of users that impose additional requirements and restrictions to interaction engineering. Spreicer discussed the potentials of using TUIs to narrow the digital divide and improve the technology acceptance of older adults [25], and Koushik et al. documented the design of a tangible block-based game that enables blind programmers to learn basic programming concepts by creating audio stories [26]. In their pilot setting, groups of teachers, Braille experts, and students worked together to create accessible stories, and their feedback offers insights for the future development of accessible, tangible programming tools. The role of tangible interfaces in accessible computing is thoroughly examined in [8], where authors presented three projects that demonstrate how tangible interfaces can be used to improve the computing experiences for the visually impaired community.

In the HRI field, TUIs are considered a promising approach as they make effective use of physical object affordances [27], but research is rather inconclusive on the benefits of tangible interfaces compared with traditional ones (mouse and keyboard). Tangible interfaces have been found to improve efficiency (navigation time), accuracy (fewer user mistakes) and user satisfaction [5], or simply improve efficiency but not accuracy [28, 29], while some studies [12] found that the quantitative results were inconclusive with only positive qualitative results. Lucignano et al. compared a tangible tool with a GUI implementation using eye-tracking data for gaining more insight into the user experience and concluded that TUIs require lower mental effort suggesting some cognitive advantages in them [30]. Adams performed user testing of heterogeneous mobile ground-based robots where participants tested one-robot, two-robot, and four-robot tasks and reported that perceived user workload significantly increased, while performance decreased during the four-robot task [31]. Merrad et al. concluded that tangible interface outperformed touch interaction in effectiveness, efficiency, and usability, in a task of remote control of one and two robots, but it did not significantly lower the users' workload [32]. Oppl and Stary investigated the effect of tangible explicit articulation (i.e., eliciting and refining mental models by means of tangible interaction), and their finding indicated the usefulness for explaining mental models and usability in terms of intuitiveness and engagement [33–35].

In this study with the purpose of investigating the effectiveness of a TUI compared with a 2D mouse-based UI for HRI remote operation, we designed and evaluated two prototype UIs based on data from real agricultural (used for vineyard spraying) robots used in the fields. In both UIs, the operator focused only on a simplified field view, monitoring the positions of two robots, without having to monitor output from sensors and cameras. Two different user groups played the role of the operator: computer experts and farmers. Therefore, the experiments presented in this paper follow a 2×2 structure, where two different types of UIs were evaluated by two different user groups. The research questions were related

to the effectiveness and user's evaluation for each UI and for each group individually and combined.

3. Materials and Methods

Investigating the effectiveness of various UI types before moving to implementation is important to avoid costly mistakes, especially after developing TUIs that are difficult to modify. Towards this goal, we designed two prototype interfaces: The first one was a typical computer UI, in which the operator used the mouse as input device to control the robots and a single display screen for output (hereinafter will be called "mouse-based UI"). The second one was a simulated tangible TUI, in which the operator moved miniatures on a surface to control the robots and monitored their movements on a surface screen (hereinafter will be called "tangible UI"). In both interfaces, the operator monitors two robots spraying one vineyard field simultaneously. For the experiment design, we assumed that each robot starts at one of the opposite corners of the field and moves toward the center. Both robots were in autonomous mode, meaning that they could identify the vineyard limits and spray the appropriate areas. During early on-field experiments [36], this was not always easy, since there might be obstacles in the line that the robot aims to move and the operator might need to intervene and control the robot to avoid them, but for simplicity of the experiments, we assumed that such issues have been resolved before the experiment.

The robots' speed in the simulated experiments is based on real data from on-field experiments from two robots. The first robot is an agricultural robot used for spraying grapes with an attached stable electric sprayer, as shown in Figure 1(a). For this robot, various UIs have been tested including a PS3 or a typical mouse/keyboard interface for controlling it and a head mounted display (HMD) or a typical screen for monitoring it [6]. This robot is spraying a large area of each field and is moving quite faster than the second, newer robot. The second robot has a robotic arm with a sprayer attached on it, as shown in Figure 1(b). This feature allows the robot to identify the grapes, then use the arm to reach them, and spray the exact area that is needed. Therefore, this robot was a lot slower when moving at the vineyard field. Several tests were conducted also for this robot [37] for evaluating performance and effectiveness.

Based on the results from our previous on-field experiments [37], we found that operators had some difficulties controlling a single robot and we wanted to investigate the case of one operator being responsible to monitor two robots. The experiment was simplified to exclude cases of robots failing to spray a specific grape or failing to move correctly on the vineyard field, so to be able to focus on specific research questions related to the comparison of the mouse-based UI with the tangible UI. In this version of both UIs, the operator focused only on the field view, monitoring the positions of the two robots, without having to monitor output from sensors and cameras.

Additionally, two different groups of users played the role of the operator, using both UI versions. The first group, called hereinafter "computer experts," was comprised of participants that had a relatively good experience in mouse-based UIs. The requirement for participating was to "use a computer and

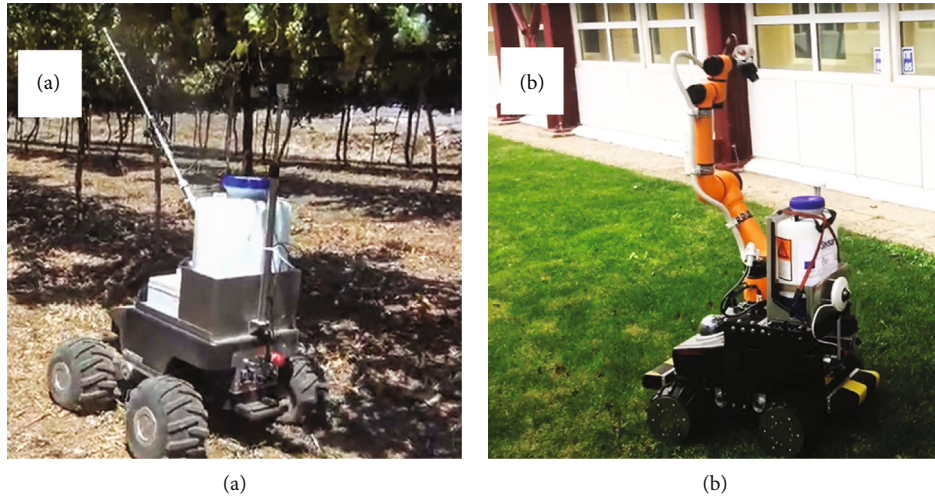


FIGURE 1: (a) Agricultural robot with a fixed electric sprayer. (b) Agricultural robot with a robotic arm and a sprayer attached on it.

work with a mouse-based interface daily for a period longer than 3 years,” but based on the participants’ demographics, most participants were either students of computer science or relative fields or working in computer companies. The second group called hereinafter “farmers” was recruited during a seminar for new farmers, and based on participants’ demographics, they were all farmers with limited (or no) experience in mouse-based interfaces. Therefore, we conducted a 2×2 experiment as shown in Figure 2, allowing us to investigate both interfaces using two different user groups. For presenting the data in further detail, we will call the experiments as “computer experts experiment” and “farmers experiment.”

The research questions, based on this 2×2 experiment structure, are related to comparing both user groups’ performances and different types of UI either combined for one category or separately. The conditions investigated were efficiency, spraying accuracy, and user’s evaluation regarding the UIs. Efficiency is measured using as metrics the time to complete the task and the number of collisions between the robots. Spraying accuracy is measured using as metrics the percentage of the vineyard left without spraying and the percentage of the vineyard that was sprayed with both robots. User evaluation is measured using a simple inquiry method, based on a short questionnaire. The relationship between the measurements and the research questions can be seen in Table 1.

Therefore, the research questions are as follows:

- (i) RQ₁: Are there any differences in efficiency (measured by time to complete the task) for each user group and each interface?
- (ii) RQ₂: Are there any differences in efficiency (measured by number of collisions) for each user group and each interface?
- (iii) RQ₃: Are there any differences in spraying accuracy (measured by the percentage of the vineyard left unsprayed) for each user group and each interface?

- (iv) RQ₄: Are there any differences in spraying accuracy (measured by the percentage of the vineyard that was sprayed with both robots) for each user group and each interface?

- (v) RQ₅: Are there any differences in users’ evaluation (measured by a 5-item custom questionnaire) for the two UIs for each user group?

The results were tested to identify statistically significant differences for each group individually and combined. For example, to evaluate efficiency using the “time” metric, we evaluated the time all users needed while operating the mouse-based compared to the time all users needed while operating the tangible UI, as well as the time each one of the user categories needed for the same interfaces, leading to 5 different data sets, as shown in Figure 2: (a) “all users + mouse-based UI” compared to “all users + tangible UI,” (b) “computer experts + mouse-based UI” compared to “computer experts + tangible UI,” (c) “farmers + mouse-based UI” compared to “farmers + tangible UI,” (d) “computer experts + mouse-based UI” compared to “farmers + mouse-based UI,” and (e) “computer experts + tangible UI” compared to “farmers + tangible UI.”

3.1. Research Design. Before starting the two experiments, we run a pilot phase involving eight users. The goals of this pilot phase were to test and calibrate the equipment, evaluate the tools, and familiarize the evaluators with the experiment requirements. Since the experiment setting is not a typical one, this phase lasted for almost two months, using multiple pilot cases with the same subjects, and led to changes to the experiment design. A change from the initial experiment design was to relocate the operator to the same room with the experiment participants. Initially, the operator was in a different room and controlled the software while watching the experiment from the monitor. Having the operator in the same room with the subjects allowed to directly observe—in

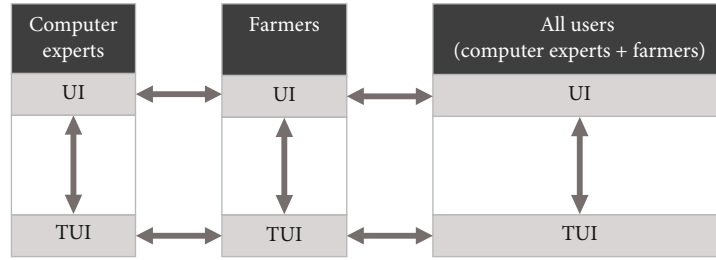


FIGURE 2: UIs and user groups evaluated during the 2x2 experiment structure.

TABLE 1: The relationship between measurements and RQs.

Research question	Examines	Measurements
RQ ₁	Efficiency	Time to complete the task
RQ ₂	Efficiency	Number of collisions
RQ ₃	Accuracy	Percentage of the vineyard left unsprayed
RQ ₄	Accuracy	Percentage of the vineyard sprayed with both robots
RQ ₅	Users' evaluation	5-item custom questionnaire

close proximity—the participant’s actions instead of viewing a fixed angle video stream of the participant, thus achieving faster response time and fewer operator errors. This was measured during the pilot phase, and it was also confirmed during the two experiments. The final design of the experiment was to include 3D-printed robots as a representation of the real ones that the user would move to operate the robots and use toy miniatures for the pilot phase (since the 3D-printed robots were not ready yet). During the pilot phase, the users loved using the toy miniatures, and we decided to abandon the idea of the 3D-printed robots and use the toy miniatures instead for the next phase. Using the miniature toys allowed the participants to hold and place them on the surface more effortlessly than the (more difficult to handle) robot miniatures. Furthermore, the participants had fun using these toy miniatures, and they felt more engaged working with them, which is a goal in HCI studies [38].

Following the pilot phase, we conducted the two experiments at the Software Quality and Human-Computer Interaction Laboratory of the Computer Engineering and Informatics Department of Patras University. The experiment protocol was the same in both experiments, with the only divergence that we did not use eye-tracking glasses during the second experiment. The initial design used eye-tracking glasses to record user actions and to investigate where the users gazed. Since the eye-tracking glasses were disconnected during some cases of the pilot phase, we also used a camera to record the experiments and keep track of user actions and time. Based on the analysis of the first experiment’s eye-tracking data, nothing worth mentioning was observed. To avoid the issue of disconnections that slowed down the experiment, we decided to rely entirely on camera recordings for the second experiment. Camera recordings were proven to be very reliable for both experiments.

3.2. Experiments Setup, Equipment, and Tools

3.2.1. The Lab. The laboratory offered two separate areas, one for the participants and one for the researchers/operators, but as already mentioned, after the pilot phase, we decided to locate the operator in the same room with the users. Therefore, we used one area solely for the mouse-based interface and the other area for the tangible interface. The users worked on a typical desktop computer for the mouse-based interface, and their data were recorded using screen capturing software. For the tangible interface, an overhead projector was used to project the screen on a white table surface, while the users could control the robots’ movements by picking up the toy miniatures and placing them on the surface. Each movement of the miniatures on the surface corresponded to a simulated movement of the robots. Additionally, the user could stop any robot at any time by picking up its miniature. All the users’ actions were performed by the operator using the Wizard of Oz technique [39]. For recording user actions with the tangible interface, we used both a camera and eye-tracking glasses for the first experiment and only the camera for the second experiment.

The operator that adopted the role of the Wizard of Oz remained the same for the entire duration of each experiment and underwent detailed training during the pilot phase. Although the operator was in the same area as the participants, we did not record any interaction between the operator and the participants, apart from the error cases. We excluded from the data any case where the operator (Wizard of Oz) failed to simulate the participants’ actions, which is a common practice in this method [40, 41]. While in the first experiment, two such cases were identified; for the second experiment, we had five operator errors, and we had to remove all these cases. Data from all these seven cases were removed from the experiments. The reason for having

more errors by the operators in the second experiment was that it co-existed with a seminar (that lasted 3 days), and the operator had to perform for multiple participants on the same day with less time to rest between participants. Contrary to the second experiment, in the first experiment, there were fewer participants each day with enough time in between (the first experiment lasted 2 weeks).

3.3. Experiments Protocol and Participants. Both experiments followed a typical within-subject design, with each participant being exposed to both experimental conditions. For randomizing, we created a list defining the starting order (either mouse-based first followed by the tangible interface or vice versa), and when the participants entered the laboratory, using a random function, they were assigned to the list. Using this method, neither the researchers nor the participants could know the order of the interface they will use before starting the experiment.

When each participant arrived at the laboratory, the researcher welcomed them and offered them a water bottle. The participants had time to relax, before starting the experiment. For the first experiment, the researcher informed each participant individually, while for the second experiment, the researcher gave a short oral presentation to all participants during the seminar. In both cases, the researcher was meticulous not to reveal any crucial information about the experiment process. To offer all participants the same information, a short video with information about the experiment process was prepared, and the participants watched it, while the researcher remained silent. Following this short orientation, the participants filled in an appropriate consent form and a questionnaire about demographic information. Then, they started their first session using the interface that was randomly assigned to them. At the end of the first session, the participants filled in a short 5-question questionnaire to evaluate the UI they just interacted with [42]). The questionnaire aimed to record participants' evaluation of how easy, efficient, accurate, and enjoyable the interface was and included the following items:

- (I) I found the interface easy to use
- (II) I considered satisfactory the time needed to complete the task
- (III) The interface helped me avoid mistakes
- (IV) I liked the interface
- (V) I could use the interface daily if required for my job

Then, the participants started the second session using the other interface (based on the random assignment), and at the end of the second session, they filled the same questionnaire for the second interface and left the laboratory.

For the second experiment only (since it was feasible to have all participants in the same room right after the end of the experiment), the researchers explained the process and the experiment in detail. For the computer experts' experiment, since it was conducted in various days, a detailed "thank you for your participation" mail, further explaining the process, was sent to all participants after the end of the experiment.

3.3.1. Computer Experts' Experiment. For the first experiment [43], participants were recruited using a call that was distributed via mailing lists. The goal of this user group was to engage users experienced in mouse-based interfaces. Therefore, the mail asked for participants who were using a computer daily for the last three years and had large experience in using mouse-based UIs. Overall, 38 users participated in the first experiment. To evaluate the effectiveness of the changes adopted in the pilot phase with new unbiased users, the first four of them were considered pilot users as well. After this short pilot phase, minor adjustments in the process were made using their data. From the remaining 34 users, two were removed from the dataset since the Wizard of Oz failed to perform the participants' actions accurately. Therefore, thirty-two participants ($n = 32$), 14 females and 18 males, accounted for the first experiment. Their ages ranged from 18 to 34 years old (mean = 24.28, median = 22.5, and SD = 3.86).

The computer experts experiment lasted for two weeks, and the participants had specific appointments in the lab (Figure 3), allowing the operator to have ample time to rest after each session and to be well-prepared. Half of the participants that remained in the dataset started using the mouse-based interface first, and half started using the tangible interface first.

3.3.2. Farmers' Experiment. The second experiment offered an opportunity and presented us with a challenge. A seminar for farmers took place that offered us the opportunity to recruit a more specific user group of potential users of such applications. The challenge was that the seminar was only for three days and allowed us a very restricted time to conduct the experiments during the breaks since many participants had travelled to attend it and were available only during the seminar. These time restrictions were the main reason for recording more errors in the Wizard of Oz method, since the operator had to work on long sessions, under time pressure, and without breaks.

After explaining the experiment to all attendees of the seminar, all 42 of them volunteered to participate (probably since the topic was thought-provoking and related to their work). Of these participants, five were removed from the dataset since the Wizard of Oz failed to perform the participants' actions accurately, and one was removed due to an error in the recording device (the camera was unplugged during the use of the tangible interface). Therefore, thirty-six participants ($n = 36$), 6 females and 30 males, accounted for the farmers' experiment. All were farmers, and their ages ranged from 28 to 42 years old (mean = 34.06, median = 33.5, and SD = 3.84). Half of the participants that remained in the dataset started using the mouse-based interface first (Figure 4), and half started using the tangible interface first.

3.3.3. Data Cleaning and Statistical Analysis. For the computer experts' experiment, eye-tracking glasses were used to identify cases of intentional or unintentional interaction between a user and the operator. This was a concern after locating the operator in the same room with the participants, but it was proven that no such issue was recorded even after a detailed analysis of all participant gaze plots. Therefore, we

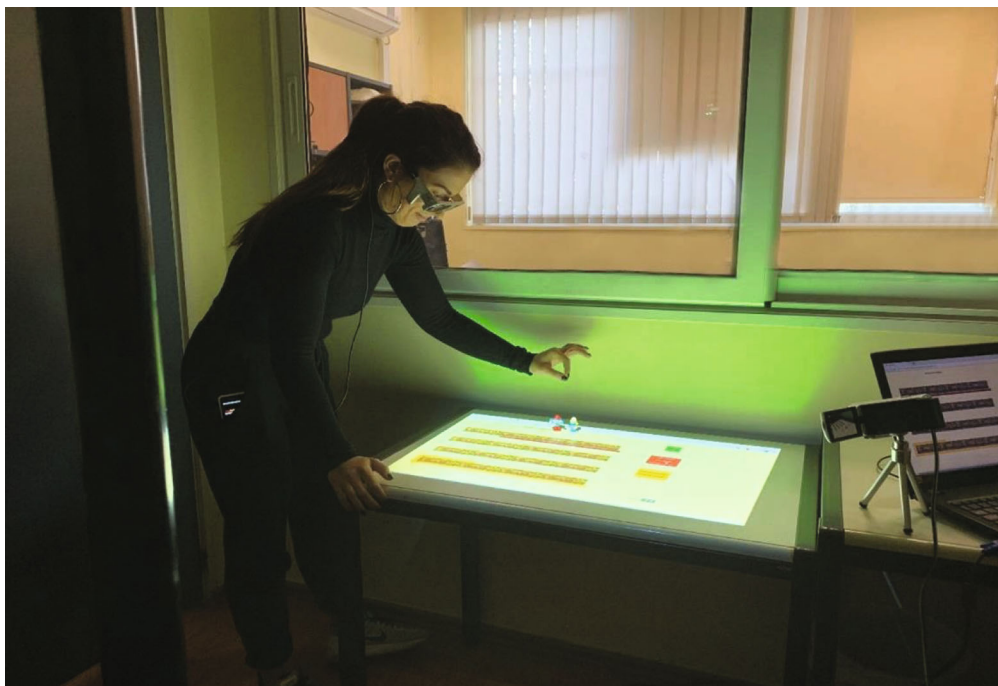


FIGURE 3: The lab that the experiment was conducted with a computer expert using the tangible UI.

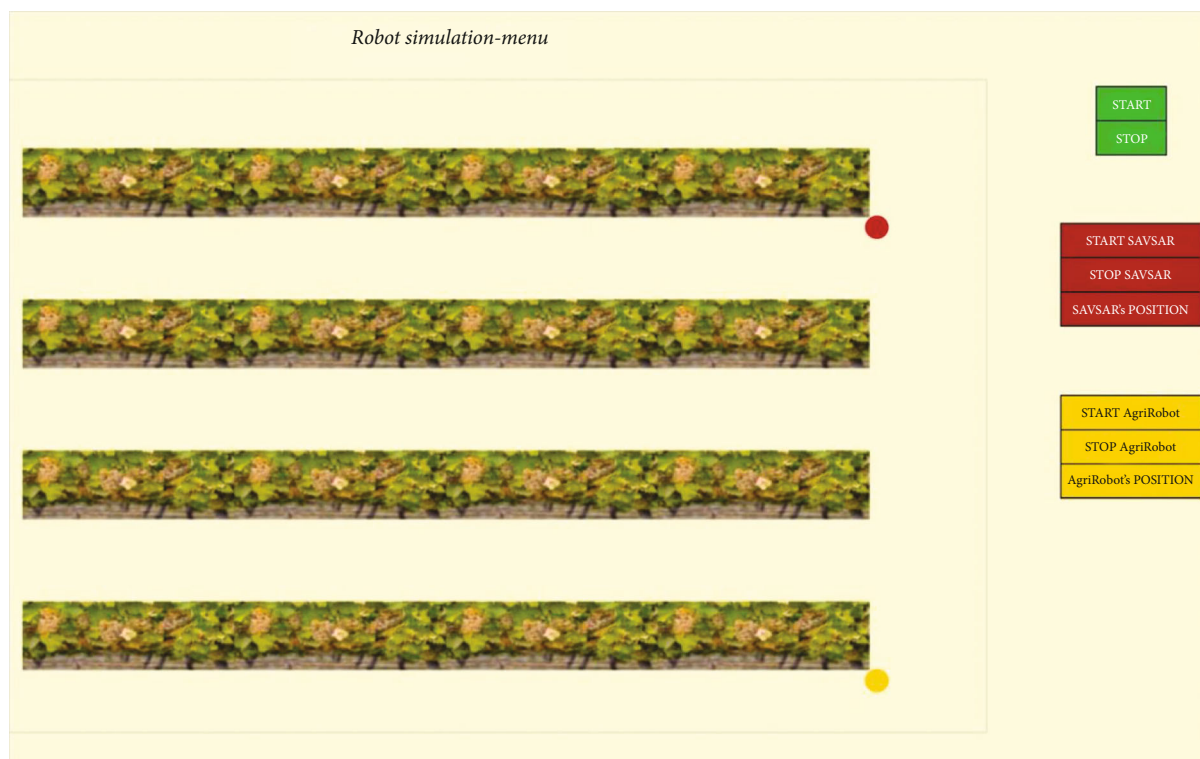


FIGURE 4: The mouse-based UI.

chose to use only the data from the camera for the second experiment which also never revealed cases that needed to remove from the data. Nevertheless, from both experiments, 7 participants were removed from the dataset due to the

Wizard of Oz operator's failure to respond to user actions accurately and promptly. Usually, in cases where the Wizard of Oz method is used, there is tolerance from the users towards the operator, but in this experiment, we adopted a

zero-tolerance policy, since the operator errors are related to the user's performance.

Following the data collection and the data cleaning, the final dataset ($n = 68$) contained an equal number of participants starting with each interface. The collected data were organized and preprocessed using Microsoft Excel 365 Pro Plus and were analyzed using IBM SPSS Statistics v27.0. The results from the statistical analysis are presented in the following section.

4. Results

This section presents the results for all research questions, with RQ_1 - RQ_4 focusing on measured data from the experiments and RQ_5 on the users' evaluation. Each subsection focuses on a specific research question and presents the results of the statistical analysis for all comparisons related to this research question; for computer experts, a comparison between the mouse-based and tangible UI; for farmers, a comparison between the mouse-based and tangible UI; comparisons between computer experts and farmers on mouse-based and tangible UI, respectively; and for all users grouped together, a comparison between the mouse-based and tangible UI. For every category, the assumptions of normality were examined, and then the appropriate statistical tests were conducted; for comparisons found to have a significant difference, a graph visualizing the results was included. The results for each research question and each comparison are presented hereinafter.

4.1. Time to Complete the Task. Tables 2 and 3 present the descriptive statistics for the completion times (in seconds) for the user groups: computer experts ($n = 32$), farmers ($n = 36$), and all users combined ($n = 68$) for each interface. Regarding the computer experts, we see comparable values when comparing the mouse-based with the tangible UI in median and mean, with differences occurring in max (261 vs 175), SD (46.86 vs 34.35), and min (40 vs 47) values, indicating that while they have similar central tendencies, the data of the mouse-based UI are more spread out. Similarly, farmers' data portray a similar picture but with smaller differences in the max value (262 vs 230) and an overall tendency to have more elevated values compared to computer experts for both UIs.

In the following sections, we present the results of the statistical analysis for all comparisons related to efficiency in terms of time to complete the task (RQ_1).

4.1.1. Mouse-Based Compared to Tangible UI for the Computer Experts. A Shapiro-Wilk test showed a significant departure from normality on time measurements for both the mouse-based ($W(32) = 0.765$, $p < 0.001$) and the tangible ($W(32) = 0.851$, $p < 0.001$) UIs, and thus, only nonparametric tests were conducted. A Wilcoxon signed rank test showed that there was no significant difference in completion time for computer experts between the mouse-based (min = 40, max = 261, median = 71.5, mean = 84.47, and SD = 46.86) and the tangible UI (min = 47, max = 175, median = 71.5, mean = 83.97, and SD = 34.35) ($n = 32$, $Z = 0.505$, and $p = 0.614$). The results indicate that regarding

TABLE 2: Time to complete the task (in seconds) for the mouse-based UI.

	All users	Computer experts	Farmers
N	68	32	36
Min	40	40	58
Max	262	261	262
Median	93	71.5	107
Mean	106.24	84.47	125.58
SD	55.76	46.86	56.38

TABLE 3: Time to complete the task (in seconds) for the tangible UI.

	All users	Computer experts	Farmers
N	68	32	36
Min	47	47	67
Max	230	175	230
Median	97	71.5	111.5
Mean	104.32	83.97	122.42
SD	43.05	34.35	42.27

RQ_1 for the computer experts, there is no significant difference in users' time to complete a specific task while using each interface.

4.1.2. Mouse-Based Compared to Tangible UI for the Farmers. For the farmers regarding RQ_1 , only nonparametric tests were conducted as a Shapiro Wilk test showed a significant departure from normality on time measurements for both mouse-based ($W(36) = 0.896$ and $p = 0.003$) and tangible UIs ($W(36) = 0.894$ and $p = 0.002$). A Wilcoxon signed rank test showed that there was no significant difference for farmers in completion time between the mouse-based (min = 58, max = 262, median = 107, mean = 125.58, and SD = 56.38) and the tangible UIs (min = 67, max = 230, median = 111.5, mean = 122.42, and SD = 42.27) ($n = 36$, $Z = -0.228$, and $p = 0.82$). The results indicate that regarding RQ_1 for the farmers, there is no significant difference in users' time to complete a specific task while using each interface.

4.1.3. Comparison of the Computer Experts and the Farmers on the Task Completion Time with the Mouse-Based UI. As shown above, the normality assumption was violated for both computer experts and farmers, and only nonparametric tests were conducted. A Mann-Whitney U test showed that there was a significant difference in completion time between computer experts ($M_{\text{rank}} = 24.91$) and farmers ($M_{\text{rank}} = 43.03$) ($U = 883$, $z = 3.773$, $p < 0.001$, and $r = 0.46$), when using the mouse-based UI. The results for RQ_1 regarding the mouse-based UI indicate that there is a significant difference in users' time to complete a specific task with the computer experts needing less time than the farmers (Figure 5).

4.1.4. Comparison of the Computer Experts and the Farmers on the Task Completion Time with the Tangible UI. As discussed above, the normality assumption was violated, and

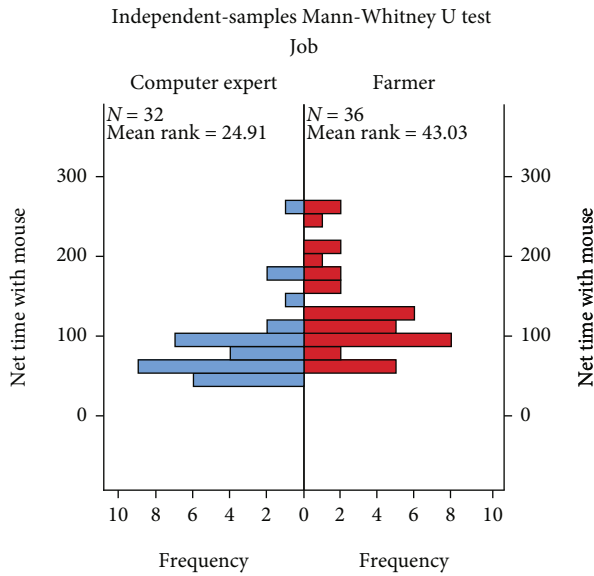


FIGURE 5: Comparison of the frequency graphs between computer experts and farmers in completion time (in seconds) with the mouse-based UI.

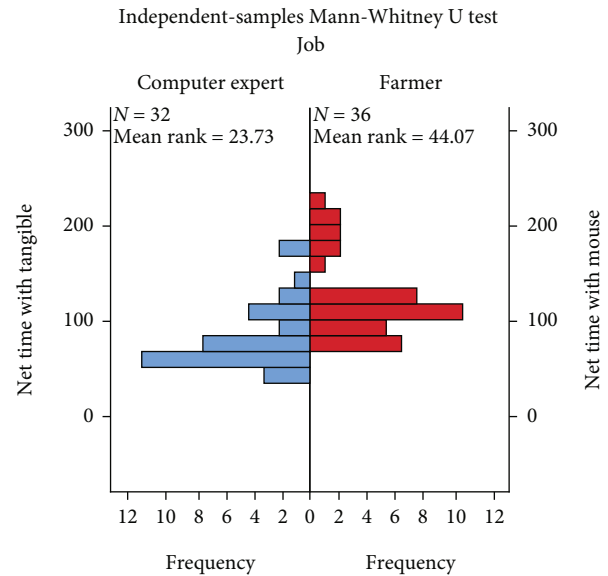


FIGURE 6: Comparison of the frequency graphs between computer experts and farmers in completion time (in seconds) with the tangible UI.

only nonparametric tests were conducted. A Mann–Whitney U test showed that there was a significant difference in completion time between computer experts ($M_{\text{rank}} = 23.73$) and farmers ($M_{\text{rank}} = 44.07$) ($U = 920.5$, $z = 4.234$, $p < 0.001$, and $r = 0.51$), when using the tangible UI. The results for RQ_1 regarding the tangible UI indicate that there is a significant difference in users' time to complete a specific task with the computer experts needing less time than the farmers (Figure 6).

4.1.5. Mouse-Based Compared to Tangible UI for All Users. For RQ_1 , a Kolmogorov-Smirnov test indicated that time measurements do not follow a normal distribution, for both mouse-based ($D(68) = 0.185$, $p < 0.001$) and tangible UIs ($D(68) = 0.12$, $p = 0.016$), and thus only nonparametric tests were conducted. A Wilcoxon signed rank test showed that there was no significant difference in completion time between the mouse-based (min = 40, max = 262, median = 93, mean = 106.24, and SD = 55.76) and the tangible UIs (min = 47, max = 230, median = 97, mean = 104.32, and SD = 43.05) ($n = 68$, $Z = 0.061$, and $p = 0.951$). The results indicate that for RQ_1 , there is no significant difference when grouping all users together (computer experts and farmers) in completion time while using each interface.

4.2. Number of Collisions. The descriptive statistics on the number of collisions for each user group are presented in Tables 4 and 5 for the mouse-based and the tangible UI, respectively. There were only a small number of collisions even though the scenario of the experiment was such that a collision was inevitable if the user had not intervened on time with the appropriate action. The low number of collisions suggests that the design of the UIs was effective in helping the users avoiding the collisions. A notable observation is that the farmers had a few users with 2 collisions in

both types of UIs, whereas the maximum number of collisions for the computer experts was 1 for the mouse-based and tangible UI.

In the following sections, we present the results of the statistical analysis for all comparisons related to efficiency in terms of number of collisions (RQ_2).

4.2.1. Mouse-Based Compared to Tangible UI for the Computer Experts. For RQ_2 , the number of collisions was measured. A Shapiro-Wilk test revealed that the normality hypothesis was violated for both the mouse-based ($W(32) = 0.511$ and $p < 0.001$) and the tangible ($W(32) = 0.540$ and $p < 0.001$) interfaces. A Wilcoxon signed rank test showed that there was no significant difference in number of collisions between the mouse-based (min = 0, max = 1, median = 0, mean = 0.22, and SD = 0.42) and the tangible UIs (min = 0, max = 1, median = 0, mean = 0.25, and SD = 0.44) ($n = 32$, $Z = 0.333$, and $p = 0.739$). The results indicate that regarding RQ_2 for the computer experts, there is no significant difference in preventing the collision of the robots while using each interface.

4.2.2. Mouse-Based Compared to Tangible UI for the Farmers. For RQ_2 , only nonparametric tests were conducted as a Shapiro Wilk test showed a significant departure from normality on the number of collisions for both mouse-based ($W(36) = 0.675$ and $p < 0.001$) and tangible UIs ($W(36) = 0.717$ and $p < 0.001$). A Wilcoxon signed rank test showed that there was no significant difference on the number of collisions between the mouse-based (min = 0, max = 2, median = 0, mean = 0.42, and SD = 0.55) and the tangible UIs (min = 0, max = 2, median = 0, mean = 0.5, and SD = 0.61) ($n = 36$, $Z = 0.688$, and $p = 0.491$). The results indicate that regarding RQ_2 for the farmers, there is no significant difference in preventing the collision of the robots while using each interface.

TABLE 4: Measured number of collisions with the mouse-based UI.

	All users	Computer experts	Farmers
N	68	32	36
Min	0	0	0
Max	2	1	2
Median	0	0	0
Mean	0.32	0.22	0.42
SD	0.5	0.42	0.55

TABLE 5: Measured number of collisions with the tangible UI.

	All users	Computer experts	Farmers
N	68	32	36
Min	0	0	0
Max	2	1	2
Median	0	0	.00
Mean	0.38	0.25	0.50
SD	0.55	0.44	0.61

4.2.3. *Comparison of the Computer Experts and the Farmers on the Measured Number of Collisions with the Mouse-Based UI.* As discussed above, the normality assumption was violated, and only nonparametric tests were conducted. A Mann–Whitney U test showed that there was not a significant difference in measured number of collisions between computer experts ($M_{\text{rank}} = 31.33$) and farmers ($M_{\text{rank}} = 37.23$) ($U = 677.5$, $z = 1.553$, and $p = 0.12$), when using the mouse-based UI. The results for RQ_2 regarding the mouse-based UI indicate that there is no significant difference in the measured number of collisions between the computer experts and the farmers.

4.2.4. *Comparison of the Computer Experts and the Farmers the Measured Number of Collisions with the Tangible UI.* As discussed above, the normality assumption was violated and only non-parametric tests were conducted. A Mann–Whitney U test showed that there was not a significant difference in measured number of collisions between computer experts ($M_{\text{rank}} = 30.75$) and farmers ($M_{\text{rank}} = 37.83$) ($U = 696$, $z = 1.768$, and $p = 0.077$), when using the tangible UI. The results for RQ_2 regarding the tangible UI indicate that there is no significant difference in the measured number of collisions between the computer experts and the farmers.

4.2.5. *Mouse-Based Compared to Tangible UI for All Users.* For RQ_2 , a Kolmogorov-Smirnov test indicated that the measurements of collisions do not follow a normal distribution, for both mouse-based ($D(68) = 0.432$ and $p < 0.001$) and tangible UIs ($D(68) = 0.405$ and $p < 0.001$), and thus only nonparametric tests were conducted. A Wilcoxon signed rank test showed that there was no significant difference in the measurements of collisions between the mouse-based (min = 0, max = 2, median = 0, mean = 0.32, and SD = 0.5) and the tangible UIs (min = 0, max = 2, median = 0, mean = 0.38, and SD = 0.55) ($n = 68$, $Z = 0.756$, and $p = 0.45$). The results indi-

cate that regarding RQ_2 for all users, there is no significant difference in the measurements of collisions while using each interface.

4.3. *Percentage of the Vineyard Left without Spraying.* Tables 6 and 7 present the descriptive statistics of the percentage of the vineyard left without spraying by the users. The experiment’s scenario necessitated the user’s intervention in the robots’ courses, which could result in parts of the vineyard being left unsprayed. Of the 32 computer experts, 28 successfully sprayed every area of the vineyard with both mouse-based and tangible UIs, and of the 36 farmers, 28 and 27 had no unsprayed areas for the mouse-based and tangible UIs, respectively. The values presented in the following tables paint a similar picture for both computer experts and farmers, with the only noteworthy exception being the considerable difference between the Max values of the mouse-based compared to tangible UI (12.5 vs 0.7) for the computer experts. This low max value (0.7), combined with the small number of users that have left the vineyard unsprayed, resulted in the reduced values of tangible compared to the mouse-based UI for mean (0.06 vs 0.51) and SD (0.17 vs 2.22).

In the following sections, we present the results of the statistical analysis for all comparisons related to the percentage of the vineyard left without spraying (RQ_3).

4.3.1. *Mouse-Based Compared to Tangible UI for the Computer Experts.* For RQ_3 , the percentage of unsprayed area was measured. A Shapiro-Wilk test showed that the percentage of the unsprayed area is not normally distributed for both the mouse-based ($W(32) = 0.242$ and $p < 0.001$) and the tangible ($W(32) = 0.411$ and $p < 0.001$) UIs. A Wilcoxon signed rank test showed that there was no significant difference for computer experts in percentage of unsprayed areas between the mouse-based (min = 0, max = 12.5, median = 0, mean = 0.6, and SD = 2.22) and tangible UIs (min = 0, max = 0.7, median = 0, mean = 0.51, and SD = 0.17) ($n = 32$, $Z = -1.352$, and $p = 0.176$). The results reveal that regarding RQ_3 for the computer experts, there is no significant difference in the percentage of the vineyard left without spraying while using each interface.

4.3.2. *Mouse-Based Compared to Tangible UI for the Farmers.* For RQ_3 , only nonparametric tests were conducted as a Shapiro Wilk test showed a significant departure from normality on the percentage of unsprayed areas for both mouse-based ($W(36) = 0.511$ and $p < 0.001$) and tangible UIs ($W(36) = 0.573$, $p < 0.001$). A Wilcoxon signed rank test showed that there was no significant difference in percentage of unsprayed areas between the mouse-based (min = 0, max = 12.5, median = 0, mean = 1.18, and SD = 2.7) and the tangible UIs (min = 0, max = 10, median = 0, mean = 1.62, and SD = 3.16) ($n = 36$, $Z = 0.527$, and $p = 0.598$). The results indicate that regarding RQ_3 for the farmers, there is no significant difference in the percentage of the vineyard left without spraying while using each interface.

4.3.3. *Comparison of the Computer Experts and the Farmers on the Percentage of Unsprayed Area with the Mouse-Based UI.* As discussed above, the normality assumption was violated, and

TABLE 6: Unsprayed percentage with the mouse-based UI.

	All users	Computer experts	Farmers
N	68	32	36
Min	0	0	0
Max	12.5	12.5	12.5
Median	0	0	0
Mean	0.86	0.51	1.18
SD	2.49	2.22	2.7

TABLE 7: Unsprayed percentage with the tangible UI.

	All users	Computer experts	Farmers
N	68	32	36
Min	0	0	0
Max	10	0.70	10
Median	0	0	0
Mean	0.89	0.06	1.63
SD	2.42	0.17	3.16

only nonparametric tests were conducted. A Mann–Whitney U test showed that there was not a significant difference on the percentage of unsprayed areas between computer experts ($M_{\text{rank}} = 32.48$) and farmers ($M_{\text{rank}} = 36.29$) ($U = 640.5$, $z = 1.193$, and $p = 0.233$), when using the mouse-based UI. The results for RQ_3 regarding the mouse-based UI indicate that there is not a significant difference in the measured percentage of the unsprayed areas between the computer experts and the farmers.

4.3.4. Comparison of the Computer Experts and the Farmers on the Percentage of Unsprayed Area with the Tangible UI. As discussed above, the normality assumption was violated, and only nonparametric tests were conducted. A Mann–Whitney U test showed that there was not a significant difference on the percentage of unsprayed areas between computer experts ($M_{\text{rank}} = 31.69$) and farmers ($M_{\text{rank}} = 37$) ($U = 666$, $z = 1.612$, and $p = 0.107$), when using the tangible UI. The results for RQ_3 regarding the tangible UI indicate that there is a not significant difference in the measured percentage of the unsprayed areas between the computer experts and the farmers.

4.3.5. Mouse-Based Compared to Tangible UI for All Users. For RQ_3 , a Kolmogorov–Smirnov test indicated that the measurements of the percentage of unsprayed areas do not follow a normal distribution, for both mouse-based ($D(68) = 0.459$ and $p < 0.001$) and tangible UIs ($D(68) = 0.452$ and $p < 0.001$), and thus only nonparametric tests were conducted. A Wilcoxon signed rank test showed that there was no significant difference on the percentage of unsprayed areas between the mouse-based (min = 0, max = 12.5, median = 0, mean = 0.86, and SD = 2.49) and the tangible UIs (min = 0, max = 10, median = 0, mean = 0.89, and SD = 2.42) ($n = 68$, $Z = -0.262$, and $p = 0.794$). The results indicate that regarding

RQ_3 for all users, there is no significant difference on the percentage of the unsprayed areas while using each interface.

4.4. Percentage of the Vineyard That Was Sprayed by Both Robots. Tables 8 and 9 present the descriptive statistics of the percentage of the vineyard area that was sprayed by both robots for both types of UIs tested. Due to the user’s intervention in the robots’ path to avoid a prearranged collision, a user choosing the wrong evasive action could result in some areas being sprayed by both robots. The following tables show many similarities in values for both user groups and UI types. Of the 32 computer experts, 16 and 17 managed to not spray a vineyard area a second time with both mouse-based and tangible UIs, respectively. Additionally, of the 36 farmers, 19 and 21 did not have any double sprayed areas for the mouse-based and tangible UIs, respectively.

In the following sections, we present the results of the statistical analysis for all comparisons related to the percentage of the double sprayed area (RQ_4).

4.4.1. Mouse-Based Compared to Tangible UI for the Computer Experts. For RQ_4 , the percentage of the area sprayed by both robots (double sprayed) was measured. A Shapiro–Wilk test showed that the percentage of the area sprayed by both robots is not normally distributed for both the mouse-based ($W(32) = 0.695$ and $p < 0.001$) and the tangible ($W(32) = 0.629$ and $p < 0.001$) interfaces. A Wilcoxon signed rank test showed that there was no significant difference in percentage of double sprayed areas for the mouse-based (min = 0, max = 25, median = 0.3, mean = 3.66, and SD = 5.61) and the tangible UIs (min = 0, max = 37.5, median = 0, mean = 5.03, and SD = 8.73) ($n = 32$, $Z = 0.7$, and $p = 0.484$). Based on the results concerning RQ_4 for computer experts, there is no significant difference in the percentage of the vineyard that was sprayed by both robots while using each interface.

4.4.2. Mouse-Based Compared to Tangible UI for the Farmers. For RQ_4 , only nonparametric tests were conducted as a Shapiro Wilk test showed a significant departure from normality on the percentage of the area sprayed by both robots for both mouse-based ($W(36) = 0.609$ and $p < 0.001$) and tangible UIs ($W(36) = 0.601$ and $p < 0.001$). A Wilcoxon signed rank test showed that there was no significant difference in percentage of double sprayed areas between the mouse-based (min = 0, max = 40, median = 0, mean = 4.47, and SD = 7.44) and the tangible UIs (min = 0, max = 40, median = 0, mean = 4.42, and SD = 8.19) ($n = 36$, $Z = -0.269$, and $p = 0.788$). The results indicate that regarding RQ_4 for farmers, there is no significant difference in the percentage of the vineyard that was sprayed by both robots while using each interface.

4.4.3. Comparison of the Computer Experts and the Farmers on the Percentage of Doubled Sprayed Areas with the Mouse-Based UI. As discussed above, the normality assumption was violated, and only nonparametric tests were conducted. A Mann–Whitney U test showed that there was not a significant difference on the percentage of double sprayed areas between the computer experts ($M_{\text{rank}} = 34.53$

TABLE 8: Percentage of double sprayed area with the mouse-based UI.

	All users	Computer experts	Farmers
N	68	32	36
Min	0	0	0
Max	40	25	40.00
Median	0	0.3	0
Mean	4.09	3.65	4.47
SD	6.61	5.61	7.44

TABLE 9: Percentage of double sprayed area with the tangible UI.

	All users	Computer experts	Farmers
N	68	32	36
Min	0	0	0
Max	40	37.5	40
Median	0	0	0
Mean	4.7	5.03	4.42
SD	8.39	8.73	8.19

) and the farmers ($M_{\text{rank}} = 34.47$) ($U = 757$, $z = -0.013$, and $p = 0.989$), when using the mouse-based UI. The results for RQ_4 regarding the mouse-based UI indicate that there is no significant difference in the measured percentage of the double sprayed areas between the computer experts and the farmers.

4.4.4. Comparison of the Computer Experts and the Farmers on the Percentage of Doubled Sprayed Areas with the Tangible UI. As discussed above, the normality assumption was violated, and only nonparametric tests were conducted. A Mann–Whitney U test showed that there was not a significant difference on the percentage of double sprayed areas between computer experts ($M_{\text{rank}} = 35.42$) and farmers ($M_{\text{rank}} = 33.68$) ($U = 546.5$, $z = -0.399$, and $p = 0.69$), when using the tangible UI. The results for RQ_4 regarding the tangible UI indicate that there is no significant difference in the measured percentage of double sprayed areas between the computer experts and the farmers.

4.4.5. Mouse-Based Compared to Tangible UI for All Users. For RQ_4 , a Kolmogorov–Smirnov test indicated that the measurements of the percentage of double sprayed areas do not follow a normal distribution, for both mouse-based ($D(68) = 0.268$ and $p < 0.001$) and tangible UIs ($D(68) = 0.288$ and $p < 0.001$), and thus only nonparametric tests were conducted. A Wilcoxon signed rank test showed that there was no significant difference on the percentage of double sprayed areas between the mouse-based (min = 0, max = 40, median = 0, mean = 4.09, and SD = 6.61) and the tangible UIs (min = 0, max = 40, median = 0, mean = 4.7, and SD = 8.39) ($n = 68$, $Z = 0.231$, and $p = 0.817$). The results indicate that regarding RQ_4 for all users, there is no

significant difference on the percentage of double sprayed areas while using each interface.

4.5. User Evaluation of the UIs

4.5.1. Reliability of Questionnaires. The users’ perception of the UIs tested (mouse-based and tangible) was evaluated with an identical questionnaire for the computer experts and the farmer participants. Reliability analysis of the questionnaires was carried out by calculating Cronbach’s α for the participants of both experiments ($n = 68$), which showed that both questionnaires had acceptable reliability ($\alpha > 0.7$). Specifically, the mouse-based questionnaire had $\alpha = 0.791$, while the tangible had $\alpha = 0.814$.

When each experiment is viewed individually, the reliability analysis of the questionnaires for the computer expert participants ($n = 32$) showed acceptable reliability for the mouse-based ($\alpha = 0.799$) and marginal reliability for the tangible UI questionnaire ($\alpha = 0.682$). The reliability analysis of the questionnaires in the experiment with the farmer participants ($n = 36$) showed acceptable reliability for the mouse-based ($\alpha = 0.765$) and the tangible UI questionnaire ($\alpha = 0.914$).

The mean score of each participant’s answers for all questions on each interface was calculated (descriptive statistics of mean scores are presented in Tables 10 and 11), and comparisons—per user group and interface—were made and are presented in the following sections. Additionally, each questionnaire item was examined individually for statistical differences, and only those with such differences are presented for each comparison.

In the following sections, we present the results of the statistical analysis for all comparisons related to the users’ evaluation of the UIs (RQ_5).

4.5.2. Mouse-Based Compared to Tangible UI for the Computer Experts. For the mean scores of the computer experts, only nonparametric tests were conducted as a Shapiro–Wilk test showed a significant departure from normality on the answers for both mouse-based ($W(32) = 0.929$ and $p = 0.036$) and tangible UIs ($W(32) = 0.927$ and $p < 0.032$). A Wilcoxon signed rank test showed that there was not a significant difference on the mean scores between the mouse-based (min = 3.2, max = 7, median = 6, mean = 5.77, and SD = 0.99) and the tangible UIs (min = 4.4, max = 7, median = 6.2, mean = 6, and SD = 0.76) ($n = 32$, $Z = 1.120$, and $p = 0.263$). The results indicate that regarding RQ_5 , there is no significant difference in the users’ evaluation of each interface for the computer experts.

Regarding the users’ evaluation of the UIs for computer experts, the data of items I, II, III, and V of the questionnaires showed no statistical differences between mouse-based and tangible UIs. In contrast to the previous findings, for the questionnaire item IV, a Shapiro–Wilk test showed that the questionnaire data were not following the normal distribution for both the mouse-based ($W(32) = 0.796$ and $p < 0.001$) and the tangible ($W(32) = 0.736$ and $p < 0.001$) UIs. Consequently, a Wilcoxon signed ranked test revealed that there is a significant difference in perceived user satisfaction for the mouse-based (min = 2, max = 7, median = 6,

TABLE 10: Mean score of users' answers with the mouse-based UI.

	All users	Computer experts	Farmers
<i>N</i>	68	32	36
Min	3.2	3.2	3.8
Max	7	7	7
Median	6.2	6	6.3
Mean	5.95	5.77	6.12
SD	0.88	0.99	0.75

TABLE 11: Mean score of users' answers with the tangible UI.

	All users	Computer experts	Farmers
<i>N</i>	68	32	36
Min	3.6	4.4	3.6
Max	7	7	7
Median	6.4	6.2	6.6
Mean	6.2	6	6.38
SD	0.8	0.76	0.9

mean = 5.94, and SD = 1.24) and the tangible UIs (min = 4, max = 7, median = 7, mean = 6.41, and SD = 0.8) ($n = 32$, $Z = 2.239$, $p = 0.025$, and $r = 0.4$). Based on the results concerning RQ₅ for computer experts, we can conclude that there is a significant difference in users' satisfaction for each interface and the opinion of the users was more positive about the tangible interface compared to the mouse-based interface (Figure 7).

4.5.3. Mouse-Based Compared to Tangible UI for the Farmers. For the mean scores of the farmers, only nonparametric tests were conducted as a Shapiro Wilk test showed a significant departure from normality on the answers for both mouse-based ($W(36) = 0.883$ and $p = 0.001$) and tangible UIs ($W(36) = 0.72$ and $p < 0.001$). A Wilcoxon signed rank test showed that there was a significant difference on the mean scores between the mouse-based (min = 3.8, max = 7, median = 6.3, mean = 6.12, and SD = 0.75) and the tangible UIs (min = 3.6, max = 7, median = 6.6, mean = 6.38, and SD = 0.9) ($n = 36$, $Z = 2.758$, $p = 0.006$, and $r = 0.46$). The results indicate that regarding RQ₅ for farmers, there is a significant difference in the users' evaluation of each interface, and the opinion of the users was more positive about the tangible interface compared to the mouse-based interface (Figure 8).

Regarding the users' evaluation of the UIs, for farmers the data of items I, II, IV, and V of the questionnaires showed no statistical differences between mouse-based and tangible UIs. For the questionnaire item III, only nonparametric tests were conducted as a Shapiro Wilk test showed a significant departure from normality on the answers for both mouse-based ($W(36) = 0.675$ and $p < 0.001$) and tangible UIs ($W(36) = 0.717$ and $p < 0.001$). A Wilcoxon signed rank test showed that there was a significant difference on the answers between the mouse-based (min = 4, max = 7, median = 6,

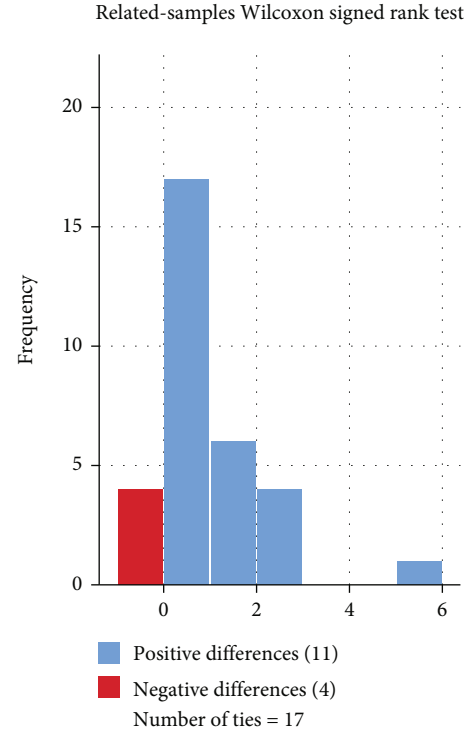


FIGURE 7: Frequency graph on differences between the mouse-based and the tangible UI concerning question IV for computer experts.

mean = 5.83, and SD = 1.06) and the tangible UIs (min = 3, max = 7, median = 7, mean = 6.31, and SD = 0.98) ($n = 36$, $Z = 2.088$, $p = 0.037$, and $r = 0.35$). The results indicate that regarding RQ₅ for farmers, there is a significant difference in the perceived error rate in controlling the robots while using each interface with users reporting less perceived error rate with the tangible compared to the mouse-based UI (Figure 9).

4.5.4. Comparison of the Computer Experts and the Farmers on the Users' Evaluation of the Mouse-Based UI. As discussed above, the normality assumption was violated, and only nonparametric tests were conducted. A Mann-Whitney U test showed that there was not a significant difference on the mean scores between the computer experts ($M_{\text{rank}} = 30.86$) and the farmers ($M_{\text{rank}} = 37.74$) ($U = 692.5$, $z = 1.438$, and $p = 0.15$), when using the mouse-based UI. The results for RQ₅ regarding the mouse-based UI indicate that there is no significant difference in the users' evaluation between the computer experts and the farmers.

4.5.5. Comparison of the Computer Experts and the Farmers on the Users' Evaluation of the Tangible UI. As discussed above, the normality assumption was violated, and only nonparametric tests were conducted. A Mann-Whitney U test showed that there was a significant difference on the mean scores between computer experts ($M_{\text{rank}} = 28$) and farmers ($M_{\text{rank}} = 40.28$) ($U = 784$, $z = 2.583$, $p = 0.01$, and $r = 0.31$), when using the tangible UI. The results for RQ₅ regarding the tangible UI indicate that there is a significant difference

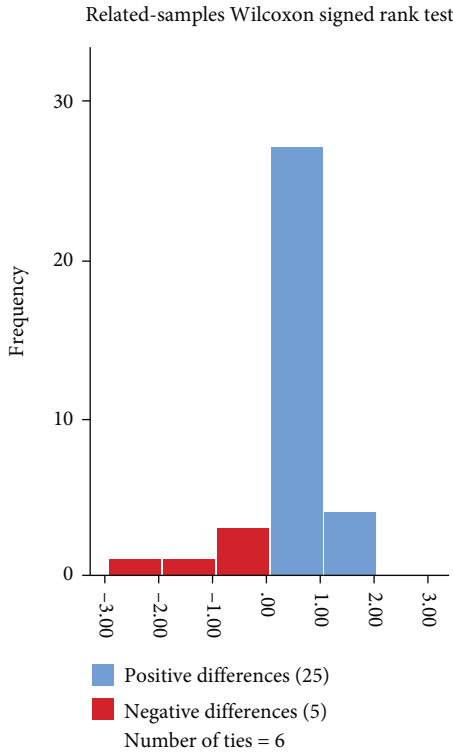


FIGURE 8: Frequency graph on differences between the mean scores of farmers' evaluation for the mouse-based and the tangible UI.

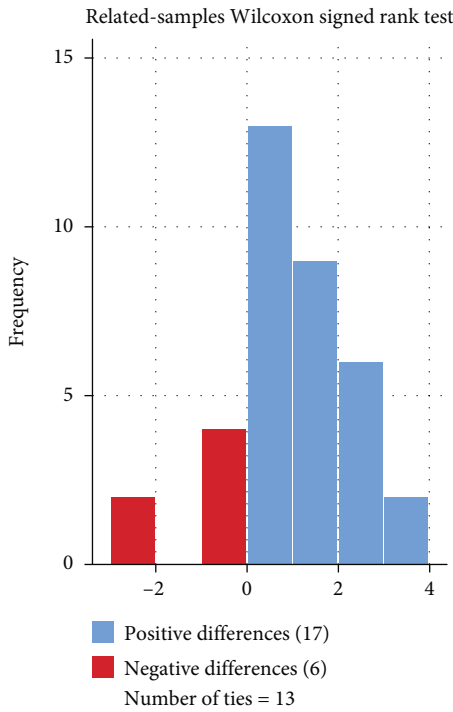


FIGURE 9: Frequency graph on differences between the mouse-based and the tangible UI concerning question III for farmers.

in the users' evaluation and the opinion of the farmers was more positive compared to the computer experts (Figure 10).

Regarding the users' evaluation of the UIs between computer experts and farmers, the data of items I, II, and IV of the questionnaires showed no statistical differences regarding the tangible UI. For questionnaire item III, a Shapiro Wilk test indicated that the data do not follow a normal distribution, for both computer experts ($W(32) = 0.848$ and $p < 0.001$) and farmers ($W(36) = 0.699$ and $p < 0.001$), and thus only non-parametric tests were conducted. For the questionnaire item III, a Mann-Whitney U test showed that there was a significant difference on the answers between computer experts ($M_{\text{rank}} = 28.89$) and farmers ($M_{\text{rank}} = 39.49$), ($U = 755.5$, $z = 2.349$, $p = 0.019$, and $r = 0.28$), when using the tangible UI. The results for RQ₅ regarding the tangible UI indicate that there is a significant difference in the perceived error rate when controlling the robots, with the farmers reporting less perceived error rate compared to the computer experts (Figure 11).

For questionnaire item V, a Shapiro Wilk test indicated that the data do not follow a normal distribution, for both computer experts ($W(32) = 0.859$ and $p = 0.001$) and farmers ($W(36) = 0.759$ and $p < 0.001$), and thus only nonparametric tests were conducted. A Mann-Whitney U test showed that there was a significant difference on the answers between computer experts ($M_{\text{rank}} = 29.77$) and farmers ($M_{\text{rank}} = 38.71$) ($U = 727.5$, $z = 1.982$, $p = 0.047$, and $r = 0.24$), when using the tangible UI. The results regarding the tangible UI indicate that there is a significant difference in users' opinions in wanting to use the interface daily if it was required by their job, with farmers reporting be more willing to use the tangible UI daily compared to the computer experts (Figure 12).

4.5.6. Mouse-Based Compared to Tangible UI for All Users. For RQ₅, a Kolmogorov-Smirnov test indicated that the mean scores for users' evaluation of the UIs do not follow a normal distribution, for both mouse-based ($D(68) = 0.166$ and $p < 0.001$) and tangible UIs ($D(68) = 0.175$ and $p < 0.001$), and thus only nonparametric tests were conducted. A Wilcoxon Signed Rank Test showed that there was a significant difference on the mean scores between the mouse-based (min = 3.2, max = 7, median = 6.2, mean = 5.95, and SD = 0.88) and the tangible UIs (min = 3.6, max = 7, median = 6.4, mean = 6.2, and SD = 0.8) ($n = 68$, $Z = 2.541$, $p = 0.011$, and $r = 0.31$). The results indicate that regarding RQ₅ for all users, there is a significant difference in the users' evaluation of each interface, and the opinion of the users was more positive about the tangible UI compared to the mouse-based UI (Figure 13).

For the questionnaire item III, only nonparametric tests were conducted as a Kolmogorov-Smirnov test showed a significant departure from normality on the answers for both mouse-based ($D(68) = 0.178$ and $p < 0.001$) and tangible UIs ($D(68) = 0.279$ and $p < 0.001$). A Wilcoxon signed rank test showed that there was a significant difference on the answers between the mouse-based (min = 1, max = 7, median = 6, mean = 5.56, and SD = 1.29) and the tangible UIs (min = 1, max = 7, median = 6, mean = 5.96, and SD = 1.28) ($n = 68$, $Z = 2.321$, and $p = 0.02$). The results indicate that regarding RQ₅ for all users, there is a significant difference in the perceived error rate in controlling the robots

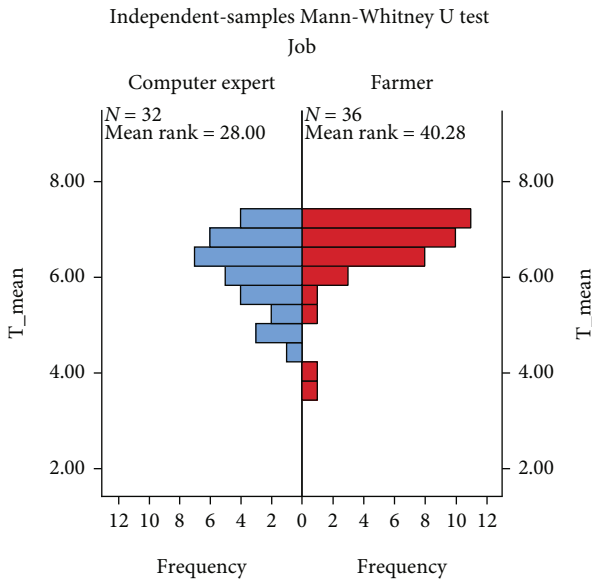


FIGURE 10: Comparison of the frequency graphs of the mean scores between computer experts and farmers with the tangible UI.

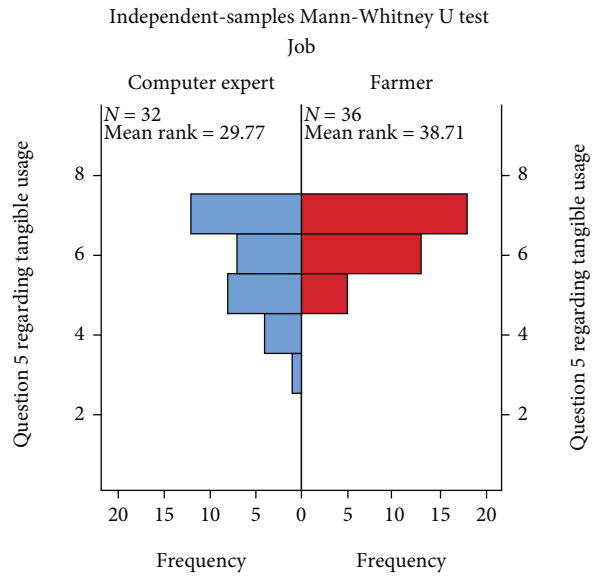


FIGURE 12: Comparison of the frequency graphs regarding question V between computer experts and farmers for the tangible UI.

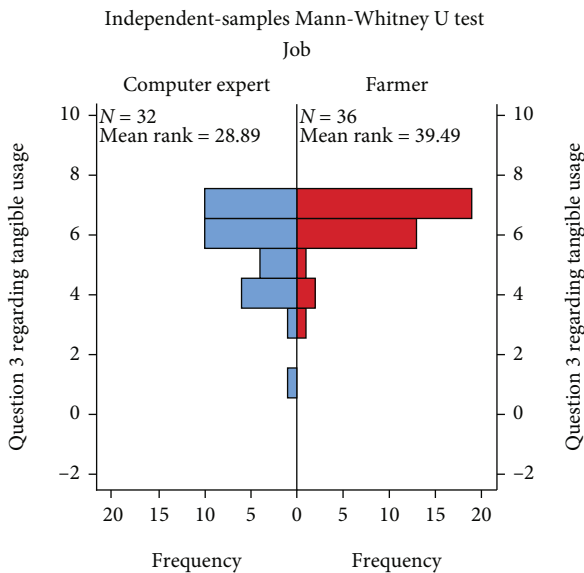


FIGURE 11: Comparison of the frequency graphs regarding question III between computer experts and farmers for the tangible UI.

while using each interface, with users reporting less perceived error rate with the tangible compared to the mouse-based UI (Figure 14).

For the questionnaire item IV, only nonparametric tests were conducted as a Kolmogorov-Smirnov test showed a significant departure from normality on the answers for both mouse-based ($D(68) = 0.241$ and $p < 0.001$) and tangible UIs ($D(68) = 0.337$ and $p < 0.001$). A Wilcoxon signed rank test showed that there was a significant difference on the answers between the mouse-based (min = 2, max = 7, median = 6, mean = 6.04, and SD = 1.14) and the tangible UIs (min = 3, max = 7, median = 7, mean = 6.37, and SD = 0.95) ($n = 68$, Z

= 2.563, and $p = 0.01$). The results indicate that regarding RQ_5 for all users, there is a significant difference in the perceived users' satisfaction for each interface, with users reporting being more satisfied with the tangible compared to the mouse-based UI (Figure 15).

5. Discussion

Based on the analysis that preceded, there were no statistically significant differences regarding the time to complete the task comparing the alternative UIs for each group. There was though a difference between computer experts and farmers in both mouse-based and tangible UIs with the computer experts being faster in both types of UI. Also, the number of collisions and the percentage of the unsprayed and double sprayed area revealed no significant differences, either for user groups or UIs.

Concerning users' evaluation for each UI, mean scores were compared (taking user answers to all questions for a specific UI), and then individual questions were examined for potential differences. For the computer experts group, there were no significant differences when comparing the mean score of each interface. When analyzing individual questions, there was a significant difference in perceived user satisfaction (Q4: I liked the interface), as computer experts liked the tangible UI more [43]. Answers from the farmers group indicate that there is a significant difference in users' evaluation of each interface (mean scores) and the user opinion was more positive about the tangible interface compared to the mouse-based. When examining individual questions, there was a significant difference in the perceived error rate in controlling the robots using each interface with farmers reporting lower perceived error rates with the tangible compared to the mouse-based UI (Q3 "The interface helped avoid mistakes").

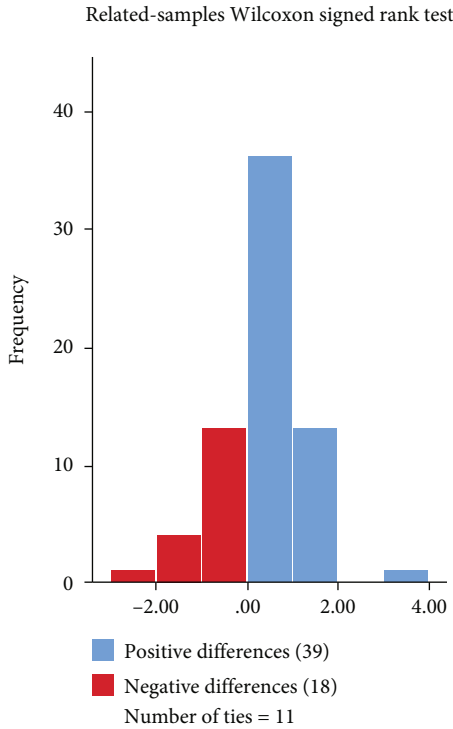


FIGURE 13: Frequency graph on differences between the mean scores of all users' evaluation for the mouse-based and the tangible UI.

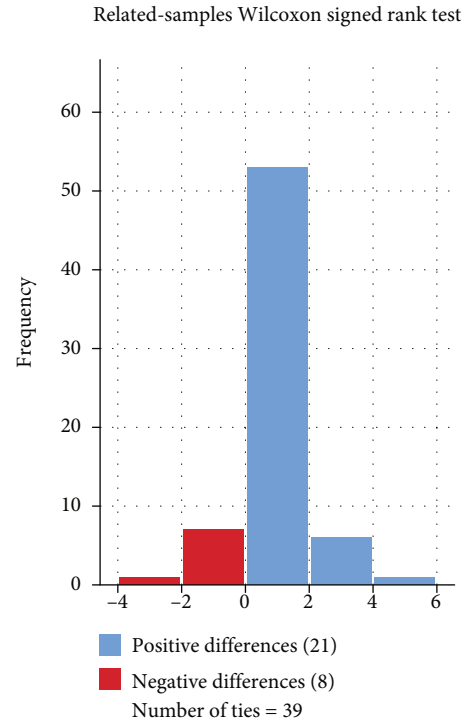


FIGURE 15: Frequency graph on differences between the mouse-based and the tangible UI concerning question IV for all users.

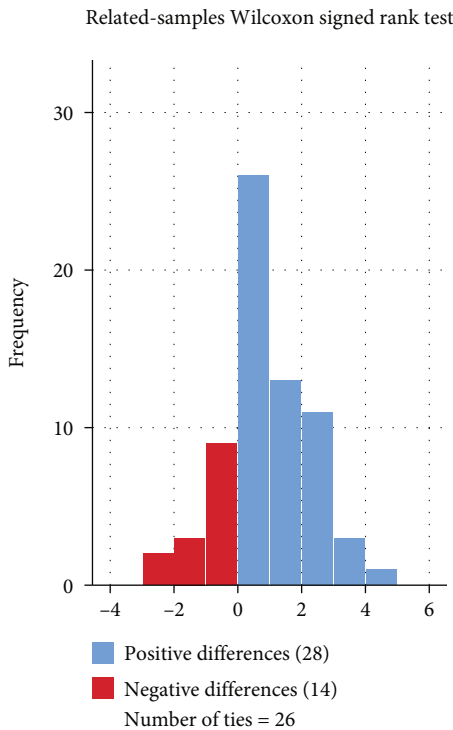


FIGURE 14: Frequency graph on differences between the mouse-based and the tangible UI concerning question III for all users.

When examining differences between the mean scores of the questionnaires, no statistically significant difference was identified in users' evaluation between computer experts and farmers with the mouse UI. There was though a significant difference regarding the tangible UI indicating that users' evaluation and the opinion of the farmers were more positive compared to the computer experts. Also, there was a significant difference in the perceived error rate when controlling the robots with the tangible UI, with farmers reporting lower perceived error rates compared to the computer experts (Q3: "The interface helped me to avoid making mistakes"). Moreover, farmers were significantly more willing to use daily the tangible UI if required by their job compared to computer experts (Q5: "I could use the interface daily if that was required for my job").

When comparing the two UIs for all users, there was a significant difference in the mean scores of the users' evaluation when using each interface, and the opinion of users was more positive about the tangible interface compared to the mouse-based. Also, for all users, there was a significant difference in the perceived error rate in controlling the robots, with users reporting a lower perceived error rate with the tangible compared to the mouse-based UI (Q3 "The interface helped me to avoid making mistakes"). In addition, for all users, there was a significant difference in the perceived users' satisfaction, with users declaring more satisfied with the tangible compared to the mouse-based UI (Q4 "I liked the interface").

In summary, the main findings of this study seem to support the inconclusiveness discussed in the Introduction section [12] where the authors conclude that their quantitative

results are inconclusive in suggesting whether our TUI approach is “better” than the non-TUI alternative. Also, we found no evidence that the tangible UI improved user efficiency or accuracy [28] [29]. Moreover, our findings agree with the observation in [12] that there is no correlation between user performance and UI perceived preference with Guo and Sharlin claiming that users seem to be more satisfied with the tangible UI [5] and this stands for both user groups. It is evident that the issue of TUIs requires further investigation and larger sample populations to confirm their effect (if any) on user efficiency, accuracy, and satisfaction. Larger sample populations and longer experiment durations would allow for studying the learnability of TUIs (given their intuitiveness) and how effectively they can be integrated into the everyday needs of domain users. In addition, it would be interesting to conduct comparative studies of TUIs with mouse or touch UI variations to determine whether TUIs offer tangible advantages, to what degree, and under which conditions (domain, type of users, or tasks).

6. Conclusions

As claimed by Baldwin et al. [8], TUIs in a way allow for merging the computational and physical worlds as they make it possible to interact with the digital artifacts through the human kinesthetic system. Their intuitiveness brings great potential to several domains to allow for more effective, less error-prone, and more satisfying user interactions lowering the required mental effort [30, 33–35]. In this study, two prototype UIs were designed and evaluated with the purpose of comparing a TUI with a 2D mouse-based UI for operating remotely two agricultural robots used for vineyard spraying. Two different user groups played the role of the operator: computer experts and farmers leading to an experiment with a 2×2 setup, where two different types of UIs were evaluated by two different user groups. The formulated research questions were related to the efficiency, accuracy, and user’s evaluation for each UI and for each group individually and combined. Analysis has shown that there were no statistically significant differences between the two UIs for each group in terms of time to complete the task (with computers experts being faster, as expected), number of collisions, as well as the percentage of unsprayed and double sprayed area. The TUI prototype was preferred by both user groups, and also, users stated that the TUI prototype helped them make fewer errors (a claim that is not supported by our observations). All users were willing to use the TUI in their daily jobs with farmers being more positive about this possibility.

Concerning the limitations of the study, the number of female users in the farmers group is low, and this could lead to reservations of gender skewness, even though the sample is quite representative of the actual female farmers population in the country. Another limitation relates to the simulation deployed for the tangible UI. The Wizard of Oz technique is used in many studies on tangible UI design and assessment as the effort and resources required to implement such a UI are prohibitively high and this simulation approach allows for comparative studies. Still, using an implemented tangible

UI might affect some of the metrics as it would probably have lower task completion times (due to lower response times compared to the human operator) and higher error rates due to the increased speed of interaction (responsiveness). These assumptions can only be verified with an implemented TUI. Finally, the significant advantage that the tangible UI had over its mouse-based variation in terms of user preference and satisfaction may be due to its innovative nature. Users often state that they prefer something new and interesting to something that is familiar but trivial. This novelty effect may also explain the results of similar studies on tangible UIs and should be considered in the interpretation of findings.

Data Availability

The primary data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank all users that participated in the experiments and Mrs. Dimitra Fountzoula and Mr. Panayiotis Kordakis for aiding in the experiments by performing the role of the Wizard of Oz.

References

- [1] M. A. Goodrich and D. R. Olsen, “Seven principles of efficient human robot interaction,” in *SMC’03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme-System Security and Assurance (Cat. No. 03CH37483)*, pp. 3942–3948, Washington, DC, USA, 2003.
- [2] L. Besançon, P. Issartel, M. Ammi, and T. Isenberg, “Usability comparison of mouse, touch and tangible inputs for 3D data manipulation,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 4727–4740, New York, 2017.
- [3] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch, “Human-aware robot navigation: a survey,” *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1726–1743, 2013.
- [4] J. Nielsen, *Usability engineering*, Morgan Kaufmann, United States, 1994.
- [5] C. Guo and E. Sharlin, “Exploring the use of tangible user interfaces for human-robot interaction: a comparative study,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 121–130, New York, 2008.
- [6] G. Adamides, C. Katsanos, Y. Parment et al., “HRI usability evaluation of interaction modes for a teleoperated agricultural robotic sprayer,” *Applied Ergonomics*, vol. 62, pp. 237–246, 2017.
- [7] H. Ishii, “Tangible user interfaces,” in *Human-Computer Interaction: Design Issues, Solutions, and Applications*, pp. 141–157, CRC Press Boca Raton, FL, USA, 2007.

- [8] M. S. Baldwin, R. Khurana, D. McIsaac et al., “Tangible interfaces,” in *Web Accessibility: A Foundation for Research*, Y. Yesilada and S. Harper, Eds., pp. 715–735, Springer, 2019.
- [9] O. Zuckerman and A. Gal-Oz, “To TUI or not to TUI: evaluating performance and preference in tangible vs. graphical user interfaces,” *International Journal of Human-Computer Studies*, vol. 71, no. 7–8, pp. 803–820, 2013.
- [10] E. F. Melcer, V. Hollis, and K. Isbister, “Tangibles vs. mouse in educational programming games: influences on enjoyment and self-beliefs,” in *Proceedings of the 2017 CHI conference extended abstracts on human factors in computing systems*, New York, 2017.
- [11] M. S. Horn, E. T. Solovey, and R. J. Jacob, “Tangible programming and informal science learning: making TUIs work for museums,” in *Proceedings of the 7th International Conference on Interaction Design and Children*, pp. 194–201, New York, 2008.
- [12] P. Lapidis, E. Sharlin, and M. Costa Sousa, “Three dimensional tangible user interface for controlling a robotic team,” in *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, pp. 343–350, New York, 2008.
- [13] P. Marshall, “Do tangible interfaces enhance learning?,” in *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, pp. 163–170, New York, 2007.
- [14] L. Xie, A. N. Antle, and N. Motamedi, “Are tangibles more fun? Comparing children’s enjoyment and engagement using physical, graphical and tangible user interfaces,” in *Proceedings of the 2nd International Conference on Tangible and Embedded Interaction*, pp. 191–198, New York, 2008.
- [15] A. Gillet, M. Sanner, D. Stoffler, and A. Olson, “Tangible interfaces for structural molecular biology,” *Structure*, vol. 13, no. 3, pp. 483–491, 2005.
- [16] A. Skulmowski, S. Pradel, T. Kühnert, G. Brunnett, and G. D. Rey, “Embodied learning using a tangible user interface: the effects of haptic perception and selective pointing on a spatial learning task,” *Computers & Education*, vol. 92, pp. 64–75, 2016.
- [17] T. Sapounidis and S. Demetriadis, “Tangible versus graphical user interfaces for robot programming: exploring cross-age children’s preferences,” *Personal and Ubiquitous Computing*, vol. 17, no. 8, pp. 1775–1786, 2013.
- [18] A. Tsimplinas, E. Berki, J. P. Valtanen, and T. Sapounidis, “Teaching computational thinking via tangible and graphical interfaces - an overview through reflective and critical lenses,” *Global Connectivity and Learning across the Generations*, no. 55, pp. 55–68, 2019.
- [19] A. Nathoo, G. Bekaroo, T. Gangabissoon, and A. Santokhee, “Using tangible user interfaces for teaching concepts of internet of things: Usability and Learning Effectiveness,” *Interactive Technology and Smart Education*, vol. 17, no. 2, pp. 133–158, 2020.
- [20] A. Nathoo, T. Gangabissoon, and G. Bekaroo, “Exploring the use of tangible user interfaces for teaching basic Java programming concepts: a usability study,” in *2019 Conference on Next Generation Computing Applications (NextComp)*, Mauritius, 2019.
- [21] J. Campbell and X. Carandang, “Comparing Graphical and Tangible User Interfaces for a Tower Defense Game,” *AMCIS 2012 Proceedings*, vol. 11, 2012.
- [22] Z. Menestrina, A. Conci, A. Siesser, R. Masu, M. Bianchi, and A. De Angeli, “Tangible and Graphical Game Interfaces: An Experimental Comparison,” in *In Games User Research*, M. A. Garcia-Ruiz, Ed., pp. 119–143, A K Peters/CRC Press, FL, USA, 2016.
- [23] C. De Raffaele, S. Smith, and O. Gemikonakli, “An active tangible user interface framework for teaching and learning artificial intelligence,” in *23rd International Conference on Intelligent User Interfaces*, pp. 535–546, New York, 2018.
- [24] J. Ma, L. Sindorf, I. Liao, and J. Frazier, “Using a tangible versus a multi-touch graphical user interface to support data exploration at a museum exhibit,” in *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 33–40, New York, 2015.
- [25] W. Spreicer, “Tangible interfaces as a chance for higher technology acceptance by the elderly,” in *Proceedings of the 12th International Conference on Computer Systems and Technologies*, pp. 311–316, New York, 2011.
- [26] V. Koushik, D. Guinness, and S. K. Kane, “Storyblocks: a tangible programming game to create accessible audio stories,” in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, New York, 2019.
- [27] S. Faisal, P. Cairns, and B. Craft, “Infovis Experience Enhancement through Mediated Interaction,” in *ICMI’05: Workshop on Multimodal Interaction for the Visualization and Exploration of Scientific Data*, Trento, Italy, 2005.
- [28] G. Randelli, M. Venanzi, and D. Nardi, “Evaluating tangible paradigms for ground robot teleoperation,” in *2011 ROMAN*, pp. 389–394, Atlanta, GA, USA, 2011a.
- [29] G. Randelli, M. Venanzi, and D. Nardi, “Tangible interfaces for robot teleoperation,” in *Proceedings of the 6th International Conference on Human-Robot Interaction*, pp. 231–232, New York, 2011b.
- [30] L. Lucignano, S. Cuendet, B. Schwendimann, M. Shirvani Boroujeni, and P. Dillenbourg, “My hands or my mouse: comparing a tangible and graphical user interface using eye-tracking data,” in *Proceedings of the FabLearn conference 2014*, Stanford, CA, USA, 2014.
- [31] J. A. Adams, “Multiple robot/single human interaction: effects on perceived workload,” *Behaviour & Information Technology*, vol. 28, no. 2, pp. 183–198, 2009.
- [32] W. Merrad, A. Héloir, C. Kolski, and A. Krüger, “RFID-based tangible and touch tabletop for dual reality in crisis management context,” *Journal on Multimodal User Interfaces*, vol. 16, no. 1, pp. 31–53, 2022.
- [33] S. Oppl and C. Stary, “Effects of a tabletop Interface on the co-construction of concept maps,” in *IFIP Conference on Human-Computer Interaction*, pp. 443–460, Springer, Berlin, Heidelberg, 2011a.
- [34] S. Oppl and C. Stary, “Towards informed metaphor selection for TUIs,” in *Proceedings of the 3rd ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, pp. 247–252, New York, 2011b.
- [35] S. Oppl and C. Stary, “Facilitating shared understanding of work situations using a tangible tabletop interface,” *Behaviour & Information Technology*, vol. 33, no. 6, pp. 619–635, 2014.
- [36] G. Adamides, C. Katsanos, G. Christou, M. Xenos, N. Kostaras, and T. Hadzilacos, “Human-robot interaction in agriculture: usability evaluation of three input devices for spraying grape clusters,” *Innovations*, vol. 24, 2013.
- [37] G. Adamides, C. Katsanos, I. Constantinou et al., “Design and development of a semi-autonomous agricultural vineyard

- sprayer: human-robot interaction aspects,” *Journal of Field Robotics*, vol. 34, no. 8, pp. 1407–1426, 2017.
- [38] J. Lazar, J. H. Feng, and H. Hochheiser, *Research Methods in Human-Computer Interaction*, Morgan Kaufmann, United States, 2017.
- [39] J. F. Kelley, “An iterative design methodology for user-friendly natural language office information applications,” *ACM Transactions on Information Systems*, vol. 2, no. 1, pp. 26–41, 1984.
- [40] N. Dahlbäck, A. Jönsson, and L. Ahrenberg, “Wizard of Oz studies – why and how,” *Knowledge-Based Systems*, vol. 6, no. 4, pp. 258–266, 1993.
- [41] S. Hudson, J. Fogarty, C. Atkeson et al., “Predicting human interruptibility with sensors: a wizard of Oz feasibility study,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 257–264, New York, 2003.
- [42] M. Xenos and D. Christodoulakis, “Software quality: the user’s point of view,” in *Software Quality and Productivity: Theory, Practice, Education and Training*, M. Lee, B.-Z. Barta, and P. Juliff, Eds., pp. 266–272, Springer, US, 1995.
- [43] A. Mallas, M. Xenos, and M. Rigou, “Evaluating a mouse-based and a tangible interface used for operator intervention on two autonomous robots,” in *Human-Computer Interaction. Multimodal and Natural Interaction*, M. Kurosu, Ed., pp. 668–678, Springer International Publishing, 2020.