Multimodal Technologies and Interaction

MDPI

Article

Human Attitudes in Robotic Path Programming: A Pilot Study of User Experience in Manual and XR-Controlled Robotic Arm Manipulation

Oscar Escallada [1,*], Nagore Osa [1], Ganix Lasa [1], Maitane Mazmela [1], Fatih Doğangün [2], Yigit Yildirim [2], Serdar Bahar [2] and Emre Ugur [2]

[1] Design Innovation Center (DBZ), Faculty of Engineering, Mondragon Unibertsitatea, Loramendi, 4, 20500 Mondragón, Spain; nosa@mondragon.edu (N.O.); glasa@mondragon.edu (G.L.); mmazmela@mondragon.edu (M.M.)

[2] Department of Computer Engineering, Bogazici University, Istanbul 34342, Turkey; fatih.dogangun@std.bogazici.edu.tr (F.D.); yigit.yildirim@bogazici.edu.tr (Y.Y.); serdar.bahar@std.bogazici.edu.tr (S.B.); emre.ugur@bogazici.edu.tr (E.U.)

* Correspondence: oescallada@mondragon.edu

Abstract: Extended reality (XR) and collaborative robots are reshaping human–robot interaction (HRI) by introducing novel control methods that enhance user experience (UX). However, human factors such as cognitive workload, usability, trust, and task performance are often underexplored. This study evaluated UX during robotic manipulation tasks under three interaction modalities: manual control, XR-based control at real-time speed (RS), and XR-based control at reduced safety speed (SS). Twenty-one participants performed a series of tasks across three scenarios, where we measured usability, workload, flow state, trust, and agency using a subjective questionnaire adapted from SUS, NASA-TLX, FSS, SoAS, and Trust in Industrial Human–Robot Collaboration Questionnaire, and objective task metrics (completion time, errors, and attempts). Our results reveal that RS-based control modes significantly reduced physical workload and improved usability compared to manual control. RS control at real-time speed enhanced task efficiency but increased error rates during complex tasks, while SS mode mitigated errors at the cost of prolonged completion times. Trust and agency remained stable across all modalities, indicating extended reality technologies do not undermine user confidence. These findings contribute to the field of human–robot collaboration by offering insights regarding efficiency, accuracy, and UX. The results are particularly relevant for industries seeking to optimize safety, productivity, and human-centric robotic systems.

Keywords: human–robot interaction; extended reality; Industry 5.0; user experience

Academic Editor: Mu-Chun Su

Received: 21 January 2025
Revised: 5 March 2025
Accepted: 6 March 2025
Published: 10 March 2025

Citation: Escallada, O.; Osa, N.; Lasa, G.; Mazmela, M.; Doğangün, F.; Yildirim, Y.; Bahar, S.; Ugur, E. Human Attitudes in Robotic Path Programming: A Pilot Study of User Experience in Manual and XR-Controlled Robotic Arm Manipulation. *Multimodal Technol. Interact.* **2025**, *9*, 27. https://doi.org/10.3390/mti9030027

1. Introduction

In recent years, the technological shifts defined by Industry 5.0 have reshaped modern manufacturing, with the incorporation of advanced robotic systems playing an essential role in enhancing operational efficiency, flexibility, and productivity [1]. These advances in robotics, along with associated digital technologies, are not only reshaping production lines but also paving the way for more responsive and adaptable industrial ecosystems [2]. Building upon this foundation, the development of Industry 5.0 marks a transformative shift in the interaction between human operators and emerging technologies such as robotics and extended reality (XR). Industry 5.0 places emphasis on human-centric approaches, highlighting the synergy between advanced technologies and human creativity [3]. This paradigm seeks to enhance the quality of human–machine collaboration by integrating tools that foster seamless, intuitive, and efficient interactions [4–6]. Robots and XR technologies have emerged as pivotal components in this ecosystem, offering novel methods for operators to interact with complex systems while maintaining control and oversight [7].

In this context, XR technologies (which encompass augmented reality (AR), virtual reality (VR), and mixed reality (MR)) have gained prominence for their ability to bridge the gap between digital and physical environments. By overlaying digital information onto the real world or immersing users in fully virtual spaces, XR enhances situational awareness, reduces cognitive load, and facilitates complex task execution [8,9]. Recent studies have shown that such immersive XR-based interfaces can not only improve accuracy and reduce mental workload in industrial robot programming but also increase user satisfaction and perceived control [10,11]. Specifically, Jiang et al. [11] found that interactive AR-based robot programming reduced training time for novice users, while Walker et al. [10] demonstrated that providing in situ visualizations and intuitive input methods fostered stronger operator engagement. Furthermore, Chen et al. [12] highlighted how allowing direct demonstration of trajectories in AR environments leads to a heightened sense of ownership over the outcome. Collectively, these findings underscore that well-designed XR systems significantly enhance user experience in human–robot interaction, offering both safer workflows and more effective collaboration. For instance, in industrial environments, XR can be employed to guide operators in real time, enabling them to visualize robotic movements, assess system states, and execute intricate operations with greater precision [13].

Despite these advancements, the effective integration of robots and XR technologies into industrial workflows necessitates a nuanced understanding of user experience (UX). User-centric evaluation methods are essential to identify barriers, optimize interaction paradigms, and ensure that these technologies align with human capabilities and limitations [14,15]. Enhancing the UX not only improves operational efficiency but also fosters user acceptance, trust, and satisfaction—critical factors for the successful adoption of Industry 5.0 technologies [16,17].

This study aims to explore the interplay between operators and robots mediated by XR, focusing on methods that enhance interaction quality and productivity. By analyzing UX, this research seeks to provide actionable insights into the design and implementation of human–robot interaction (HRI) systems within the Industry 5.0 framework.

2. Key UX Factors for XR in Human–Robot Collaboration

The integration of XR and robotics in industrial environments has garnered significant attention in recent years, driven by the increasing demand for flexible, efficient, and user-friendly systems. Human–robot interaction (HRI) examines how humans and robots communicate, collaborate, and coexist within shared environments. In practice, HRI can span a continuum ranging from traditional cage-based setups, where the robot is isolated for safety, to full collaboration, in which robots and humans work on the same task in unison. Collaborative robots—cobots—facilitate these closer interactions by incorporating advanced sensing, force-limiting capabilities, and intuitive interfaces, enabling safe and efficient cooperation without the need for traditional safety barriers [18,19].

Human–robot collaboration (HRC) refers to these settings, where humans and robots work together in real time to achieve common objectives, with robotic systems continuously adapting to human actions and requirements [20]. Such closely integrated interactions are often considered the most hazardous, owing to extensive shared workspaces and direct contact between human operators and robotic systems. Traditionally, safety in human–machine interaction has been predominantly focused on the physical safety [21–23]. However, in recent years, the psychological aspects of safety, such as experience and familiarity, comfort, predictability, transparency, and trust have garnered significant attention in manufacturing safety reviews [24–27].

Recent advancements have focused on developing interaction paradigms that balance system autonomy with human control, thereby fostering trust and improving task performance [28]. This evolution opens opportunities to integrate XR technologies, enabling more seamless and intuitive interaction paradigms within human–robot collaboration.

Wang et al. [29] provide a systematic review of XR-enabled remote human–robot interaction systems, highlighting current trends and future directions in this field. Their review also identifies key challenges (including system latency and the need for adaptive, multimodal, user-centered designs) that must be addressed to further optimize XR-enabled remote HRI. For instance, Karpichev et al. [30] propose a human-in-the-loop framework that utilizes XR to facilitate intuitive communication and programming between humans and robots, enhancing adaptability and task generalization. However, to ensure that such interactions are truly effective from the human perspective, especially in complex robotic environments, it is essential to consider both their strengths and limitations. Factors such as usability, workload, flow state, agency, and trust play pivotal roles in shaping the success of XR-based human–robot interaction.

Usability serves as the foundation for evaluating XR and robotic systems, as it directly influences user acceptance and task performance. Standardized frameworks such as the system usability scale (SUS) [31] have been widely used to measure usability. In XR environments, where interface design, latency, and responsiveness play a significant role, ensuring high usability is critical to delivering a seamless UX [17]. Research by Kim et al. [32] highlights that adaptive and intuitive interfaces tailored to varying skill levels not only improve usability but also enhance task performance and improve learnability in industrial settings.

While usability simplifies interaction, workload, measured using the NASA Task Load Index (NASA-TLX) [33], reflects the cognitive and physical effort required from users during task execution. XR systems, when properly designed, have the potential to alleviate workload by streamlining task-related information delivery. For instance, Dünser et al. [34] demonstrated that extended reality interfaces reduce cognitive load by providing real-time, context-relevant information. However, systems with high latency or calibration errors can significantly increase cognitive demands, leading to user frustration and decreased performance. These findings underscore the importance of balancing usability and task complexity to optimize workload and UX [33].

Reducing workload and enhancing usability creates conditions conducive to achieving the flow state, a concept introduced by Csikszentmihalyi [35]. Flow state describes the experience of complete immersion and engagement in a task, achieved through clear goals, immediate feedback, and an appropriate challenge–skill balance [36–38]. XR systems, as shown by Speicher et al. [8], can facilitate flow through immersive environments that align with these principles, leading to greater task satisfaction and productivity. However, disruptions such as technical glitches or poorly designed interactions can break this state, negatively affecting user performance. Therefore, ensuring the stability and reliability of XR systems is key to maintaining flow state.

Closely linked to usability and flow is the concept of agency, to the subjective experience of controlling one's actions and, through them, external events [39–41]. In XR-based HRI, agency is particularly relevant for maintaining user confidence and engagement. The sense of agency scale (SoAS) [42] has been widely used to measure this construct, with studies indicating that higher agency improves satisfaction and task efficiency. In particular, [42] demonstrate that the SoAS effectively captures individuals' perceived control over their actions, and that stronger perceived agency correlates not only with enhanced performance but also with more positive subjective experiences. This suggests that designing interfaces and interactions to bolster users' sense of agency can lead to better engagement and overall outcomes. [42]. XR systems can strengthen agency through gesture-based interactions and real-time feedback mechanisms. However, delays in system responsiveness or design flaws can erode this sense of control, resulting in disengagement and diminished performance.

Finally, trust plays a crucial role in the acceptance and effective use of XR and robotic systems. Transparent system behavior and predictable performance are key to building trust, as Hoffman [28] notes. Trust is reinforced in XR systems through consistent feedback, robust error handling, and reliable operations, which enable users to depend on the technology even in complex scenarios. Importantly, trust interacts dynamically with other variables—high usability, reduced workload, flow state, and agency all bolster trust, while excessive cognitive demands or low agency can undermine it.

These variables (usability, workload, flow, agency, and trust) collectively shape the UX in XR and robotic systems. Despite their potential, challenges such as latency, calibration issues, and the cognitive demands of immersive environments remain significant.

Aligning these insights with technological capabilities is critical for developing effective and inclusive HRI solutions for Industry 5.0. To address this, the present study aims to conduct a comparative analysis of various interaction methods, evaluating key factors such as usability, workload, flow, agency, and trust. Given that the tested scenario involves direct physical contact and simultaneous human–robot task execution, it falls within the domain of human–robot collaboration (HRC), where safety, transparency, and mutual adaptation are critical. The goal is to determine where XR solutions genuinely enhance human–robot interaction and identify scenarios where conventional methods still outperform XR-based approaches. This analysis will provide a comprehensive understanding of how these technologies can best be utilized to improve collaboration in complex environments.

3. Research Question

The primary objective of this study is to investigate the impact of XR technologies on HRI during robotic path programming. Specifically, the research explores whether XR-based interfaces improve HRI compared to manual programming methods, focusing on both technical performance (e.g., task efficiency, workload, error rates) and human attitudes (e.g., usability, trust, and overall satisfaction). The main hypothesis of this study is as follows:

RQ: *XR technologies enhance human–robot interaction in robotic path programming by improving technical performance while positively influencing operator's experience, such as perceived usability, trust, and workload compared to manual programming methods.*

Additionally, this study considers whether variations in XR interaction speed influence these outcomes. While not treated as a separate research question, this aspect provides an exploratory layer to determine whether differences in execution speed (real-time interaction versus slower/safety mode) affect task performance and user perception. Specifically, we aim to assess whether slower speeds enhance precision and perceived control or whether real-time speeds offer operational advantages despite potential trade-offs.

4. Materials and Methods

4.1. Participants

The participants in this study were primarily from Computer Science and Robotics background, reflecting the industrial context for which collaborative robots (cobots) are often designed. A total of 21 participants from Turkey took part, including 11 women and 10 men, with a mean age of 22.76 years (SD = 4.28 years), ensuring a gender balance.

Importantly, we prioritized diversity within this technically oriented group by including participants with varying levels of experience in both robotics and XR technologies (see Appendix A.1). Additionally, participants were asked whether they required contact lenses (13.6%), glasses (50%), or no visual aids (36.4%), as these factors can sometimes influence interactions with XR glasses. This segmentation is summarized in Table 1 below.

Table 1. Table of the segmentation of participants.

Experience Interacting with	Experience with Extended Reality			Total
	XR Glasses	XR Mobile	Knows What It Is but Never Used	
Experienced	1	1	0	2
Little experience	2	2	2	6
No experience	5	3	5	13
Total	8	8	7	21

Regarding the experience interacting with robots, we considered *No Experience* when the participant has never interacted with a robot. *Little experience* means that the participant has at least interacted once with any kind of robot, and *Experienced* when the participant has experience interacting with robotic arms while executing a task.

A similar segmentation has been performed with XR. *XR glasses* on the table means that the participant has interacted at least once with XR through head-mounted displays (HMD). *XR mobile* means that the participant has never used HMDs but has at least interacted with tablets/mobile phones while experiencing XR. In *Knows what it is but never used* is when the participant has never used any kind of XR but has heard about it. On the questionnaire there was even a lower rating that said: *Never interacted and did not know about it until now*, but no participant chose that option.

4.2. Experimental Setup

This study utilized an experimental setup consisting of a UR10 robot and a 160 × 75 cm table, on which a template with a pre-drawn circuit was placed. Figure 1 shows the experimental setup in which a participant (1) programs the robot's path (2) by drawing an imaginary line on the circuit table (3). The PC screen (4) displays the participant's perspective, where a purple line indicates recorded path. Simultaneously, a virtual robot follows the traced path in real time. Participants were required to perform tasks at three different levels of difficulty, each executed in three distinct modes. The participants programmed the robot in these three modes and then observed how the robot replicated the programmed movements. Subsequent analysis included results derived from observations and a questionnaire described in a later section.

These are the three levels set for the experiment (see Figure 2):

- Level 1 (easy): The robot was required to follow a flat circuit without leaving the lines, limited to motion on the x and y axes, without height differences. This level primarily assessed motion control and served as a familiarization phase for participants.
- Level 2 (intermediate): This level introduced a higher degree of dexterity by incorporating tactile-sensitive obstacles. Both the participants (during manipulation) and the robot (during execution) had to avoid these obstacles. If touched, these lightweight obstacles would either move or fall, providing clear visual feedback.
- Level 3 (advanced): In addition to the challenges of Level 2, this level incorporated motion along the *z*-axis, requiring greater precision. A bridge was added that the robot

had to avoid by moving over it, along with a second bridge featuring a wooden cube. The robot had to push the cube off the platform, necessitating precise manipulation and configuration by the participant. This level integrated the general requirements for robotic manipulation, such as dexterity across all three axes (x, y, z), and precision.

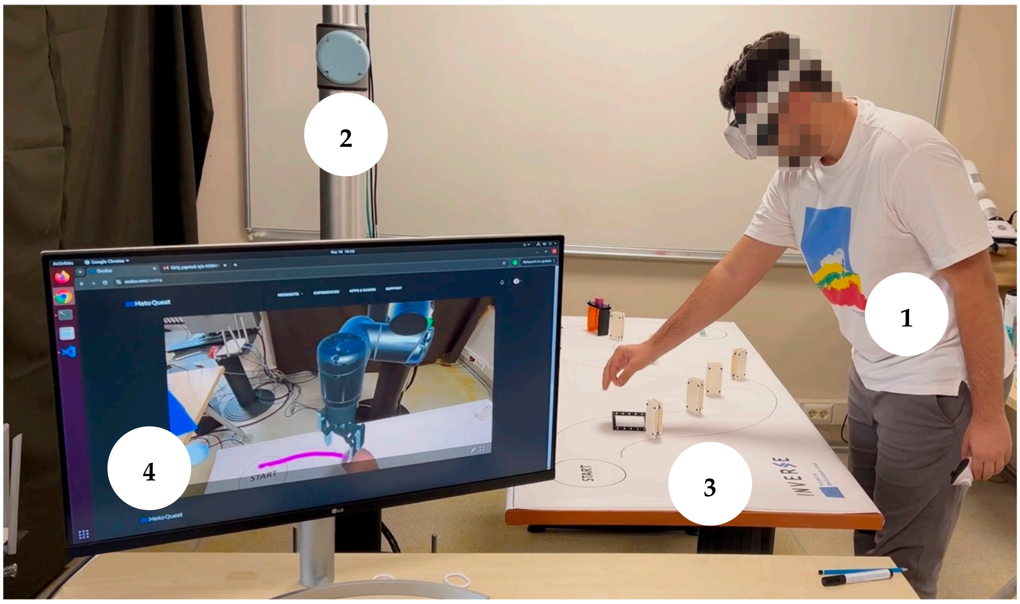

Figure 1. Experimental setup.

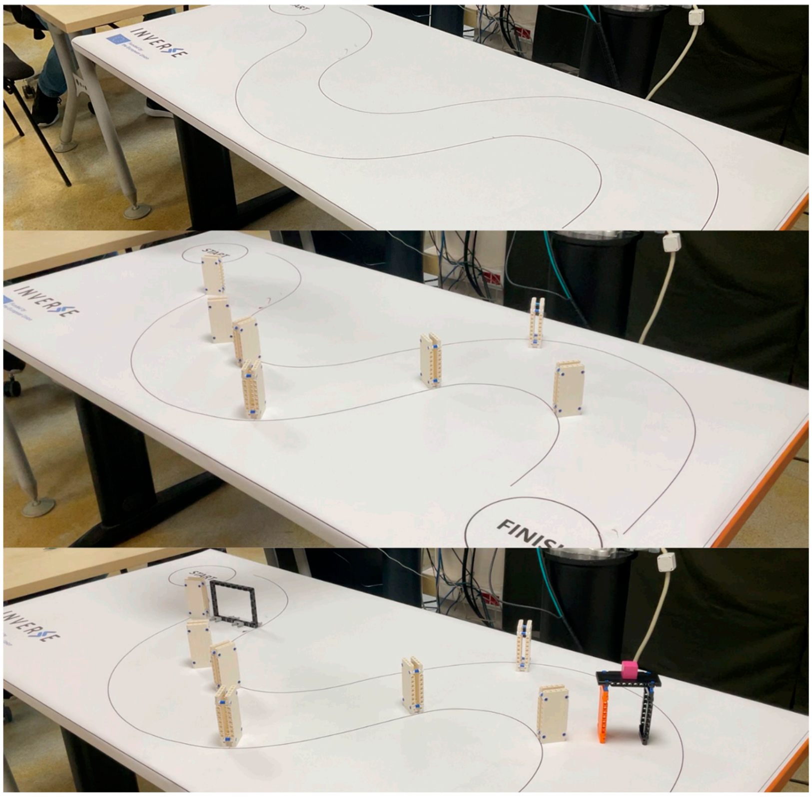

Figure 2. Different levels of difficulty for the setup.

Regarding manipulation modes, each participant performed the task using three distinct methods, with the order of these methods randomized among participants to mitigate task familiarity effects. The manipulation modes were:

- Manual robot manipulation (RE): Participants manually guided the robot to record the trajectory using their hands. After recording, participants observed the robot's execution of the programmed task while standing in front of the table.
- XR manipulation in real-time speed (RS): Participants used Meta Quest 3 XR glasses [43] and the RAMPA (v.1.0) application [44]. The RAMPA application is a software tool designed to assist users in programming robot paths by allowing them to draw imaginary lines on a circuit table. As users draw, the application displays a virtual robot following the traced path in real time, providing immediate visual feedback to facilitate precise path programming. The physical robot then replicates the movements at the same speed as the participant's drawing. If the participant draws quickly, the robot moves quickly, and if they draw slowly, the robot replicates the trajectory at a slower pace.
- XR manipulation in safety speed (SS): Like the RS mode, participants used the XR glasses and application to draw the robot's trajectory. However, in this mode, the physical robot executed the movements at a speed five times slower than the participant's drawing speed.

When conducting the test, the manual vs. RS comparison helps validate our research question regarding the improvement of UX in extended reality. Meanwhile, the RS vs. SS comparison allows us to analyze whether the robot's speed influences the UX. The entire test sequence is presented in Table 2.

Table 2. Task levels and different human-robot interactions.

Interaction with the Robot	Speed of the Robot	Level of the Task
Manual robot manipulation	Same speed as user (RE).	Level 1: Easy, motion control (x,y)
		Level 2: Medium, dexterity (x,y)
		Level 3: Medium, dexterity and precision (x,y,z)
XR-based robot manipulation	Same speed as user (RS).	Level 1: Easy, motion control (x,y)
		Level 2: Medium, dexterity (x,y)
		Level 3: Medium, dexterity and precision (x,y,z)
	Safety speed: X5 times slower than user (SS).	Level 1: Easy, motion control (x,y)
		Level 2: Medium, dexterity (x,y)
		Level 3: Medium, dexterity and precision (x,y,z)

4.3. Variables, Measures and Evaluations

Table 3 summarizes the variables and the corresponding measures. The questionnaire utilized for this study is presented in Appendix A.2. This questionnaire was developed by adapting items from various standardized instruments. The selection of questions was tailored to the specific tasks and evaluation criteria of the study. To simplify the process for participants, the rating scale was normalized to a 21-point Likert scale [45], using the NASA-TLX framework as a basis. Below, the analyzed variables are presented.

Table 3. Items of the used questionnaire.

Variable	Tool	Number of Items
Usability	System usability scale	6
Workload	NASA-TLX	6
Flow state	Flow state scale	4
Agency	Sense of agency scale	2
Trust	Trust in Industrial Human–Robot Collaboration Questionnaire	3

- Usability: For usability, a subset of 6 questions (Q1–Q6) from the system usability scale (SUS) [31] was used. The Cronbach's alpha score of 0.81 demonstrates good internal consistency, indicating that these items effectively capture perceived ease of use and usefulness of the system. Despite using only part of the original SUS scale, the results confirm the reliability of this measure.
- Workload: The workload construct was evaluated using all 6 questions (Q7–Q12) from the NASA-TLX [46]. The resulting Cronbach's alpha of 0.74 indicates acceptable internal consistency. This score suggests that the selected items reliably measure the cognitive and physical demands placed on users during task execution.
- Flow: The flow construct was measured using 4 questions (Q13–Q16) from the flow state scale (FSS) [47]. The Cronbach's alpha for this construct was 0.32, reflecting poor internal consistency. This result may stem from the limited number of items used or variability in participant responses. Future studies should consider revising or expanding the flow-related items to better capture the immersion and engagement experience.
- Agency: For agency, 2 questions (Q17–Q18) from the sense of agency scale (SoAS) [42] were used. The Cronbach's alpha score of 0.73 indicates acceptable internal consistency, suggesting that the selected items adequately measure the user's perceived control and influence within the system.
- Trust: The trust construct was assessed using 3 questions (Q19–Q21) from the Trust in Industrial Human–Robot Collaboration Questionnaire [48]. Cronbach's alpha for this construct was 0.58, indicating relatively low internal consistency. This suggests a need for refining the measurement items or including additional questions to assess user trust in the system comprehensively.
- Technical measurements: In addition to the questionnaire, several technical measurements were recorded during the user testing. These measurements included the number of errors made, the time taken for task execution, and the number of attempts required for each task. These measurements were gathered via observations and video recordings.

Finally, after completing all levels using the three modes of interaction, participants were asked to rank their preference for the modes from 1 to 3, with 1 indicating the most preferred mode and 3 the least preferred.

5. Results

In this section, the findings from the study are presented, focusing on both quantitative and qualitative analyses of user interactions with the system under three modes of control: manual control (RE), XR control at normal speed (RS), and XR control at safety speed (SS). The results include a comparative analysis of questionnaire responses to evaluate user perceptions, qualitative metrics assessing usability, workload, flow, agency, and trust, and performance metrics that capture task efficiency and accuracy. Additionally, the section explores differences among user segmentations and overall preferences for the three modes of system use. These analyses provide a comprehensive understanding of the system's impact on UX, performance, and preferred usage mode.

5.1. RQ: Do XR Technologies Enhance Human–Robot Interaction in Robotic Path Programming by Improving Technical Performance and Operator's Experience?

Here, we present the results of the questionnaire, focusing on user perceptions of the three modes of control: manual control (RE), XR control at normal speed (RS), and extended reality control at safety speed (SS). Mean scores for each questionnaire item were calculated to identify trends and differences in user responses.

A one-way analysis of variance (ANOVA) was also performed for each questionnaire item to assess differences among the three modes. As shown in Figure 3, significant differences were found in 6 out of the 21 items ($p < 0.05$):

- Q1. Desire to use the system frequently: A significant effect of mode on participants' desire to use the system frequently was observed ($F(2,N) = 7.21$, $p = 0.0016$). Both XR control modes elicited a higher desire to use the system compared to manual control.
- Q2. Perceived ease of use: The perceived ease of use differed significantly among modes ($F(2,N) = 15.74$, $p < 0.001$). Participants rated the XR control modes as easier to use than the manual control mode.
- Q5. Learning curve: Beliefs about the ease with which most people could learn to use the system varied by mode ($F(2,N) = 3.88$, $p = 0.026$). The XR control modes were perceived as having a shorter learning curve.
- Q6. Laboriousness of use: Significant differences were found in the perceived laboriousness of the system ($F(2,N) = 23.13$, $p < 0.001$), with the manual control mode rated as more laborious.
- Q8. Physical activity required: The physical effort required differed significantly among modes ($F(2,N) = 135.11$, $p < 0.001$). The manual control mode required more physical activity than the XR control modes.
- Q11. Effort to achieve performance: A significant difference was observed in the effort participants felt they had to exert to achieve their performance level ($F(2,N) = 10.73$, $p = 0.0001$), with less effort reported under XR control modes.

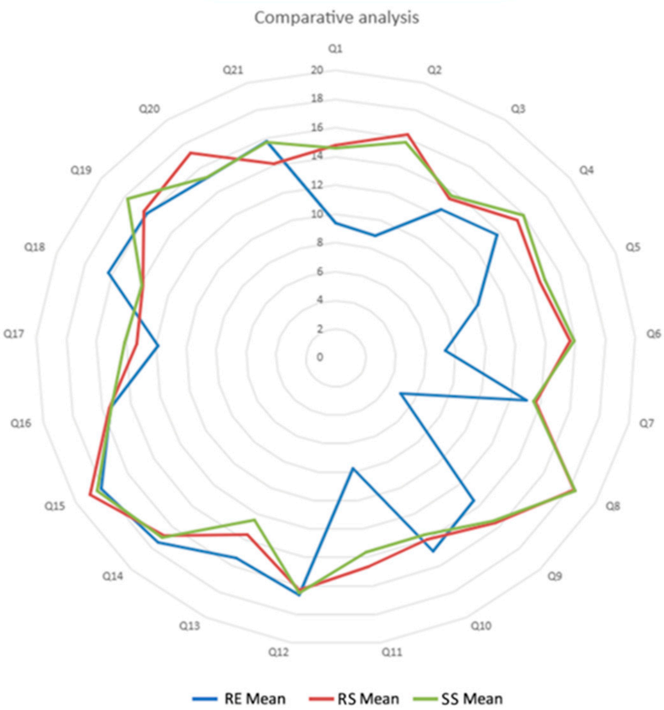

Figure 3. Overall results of the questionnaire.

Post hoc analyses indicate that the significant differences were between the manual control mode and both XR control modes, with no significant differences between RS and SS. This suggests that while XR control enhances UX over manual control, the variation in speed within XR control modes does not significantly impact user perceptions of these items. No significant differences were found for the remaining 15 items ($p > 0.05$), indicating that the mode of control did not significantly affect participants' perceptions in these areas.

5.2. Qualitative Metrics

- **Usability:** The results for usability, measured through six questions (Q1–Q6), reveal a clear advantage for the XR-based control modes (RS and SS) compared to manual control (RE). Participants rated the XR modes significantly higher in perceived ease of use (Q2) and the learning curve (Q5), indicating that XR control was both intuitive and easier to learn. Additionally, the manual mode was perceived as more laborious, as evidenced by the lower scores in Q6, which assessed the effort required to use the system. The radar chart (Figure 4) visually supports these findings, showing consistently higher scores for XR-based modes across all usability items. These results highlight the XR modes' ability to streamline interactions and reduce the friction typically associated with manual control, making them the preferred choice in terms of user-friendliness and efficiency.

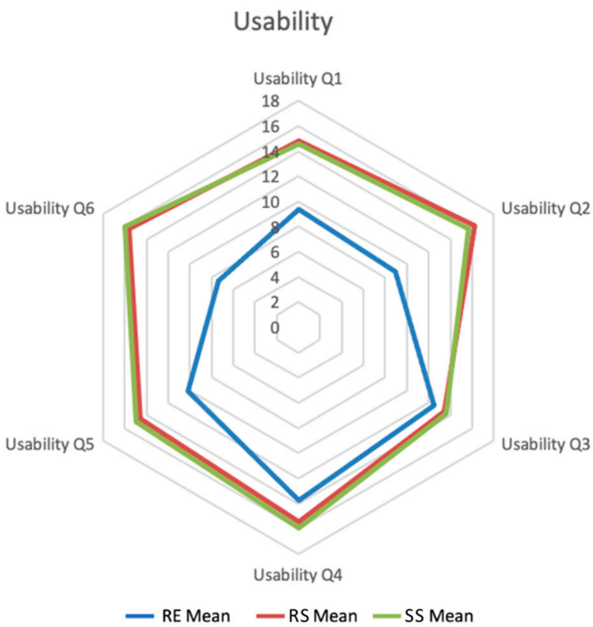

Figure 4. Usability results extracted from the questionnaire.

- **Workload:** Workload was evaluated using six items (Q7–Q12) from the NASA Task Load Index (NASA-TLX), and the results indicate significant differences between the control modes. The manual mode (RE) consistently imposed a higher cognitive and physical workload on participants. For example, participants reported increased mental demands in Q7, as the manual control required greater focus on decision-making, coordination, and precision. Similarly, Q8 highlighted higher physical effort for manual tasks, where controlling the system required strenuous movements. Q11 further reinforced this trend, with users feeling they had to work harder (both mentally and physically) to achieve satisfactory performance.

In contrast, the XR-based modes (especially SS with its safety speed) demonstrated significantly lower workload scores, as can be seen in Figure 5. This suggests that XR technologies effectively reduce task-related demands, improving overall efficiency and reducing user fatigue.

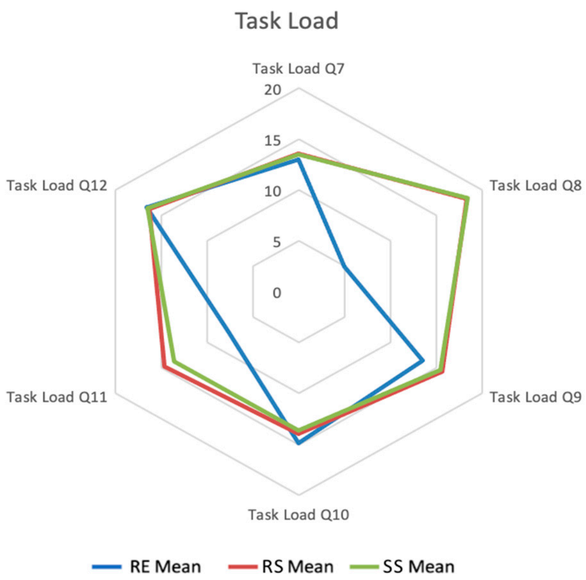

Figure 5. Workload results extracted from the questionnaire.

- **Flow State:** The results for flow state, assessed through items such as Q13 (challenge in relation to skills) and Q15 (focused attention), reveal a relatively balanced perception across the three modes of control: manual (RE), XR at normal speed (RS), and XR at safety speed (SS). Notably, in Q13, the manual control mode (RE) scored slightly higher, suggesting that participants felt the challenge provided by direct interaction with the robot was better aligned with their perceived skills. However, despite potential limitations of the XR modes (such as the intrusiveness of wearing XR glasses, weight of the equipment, and the experience of perceiving reality through a screen) participants rated their ability to stay focused and immersed in the tasks similarly to manual control. This lack of a significant difference can be viewed as a positive indicator: the XR-based interactions, even with their physical and perceptual demands, did not substantially hinder the participants' ability to achieve flow states. These results (see Figure 6) suggest that XR systems can deliver experiences comparable to real-world interactions, demonstrating their potential to support user engagement and focus on human–robot interactions.
- **Agency:** The sense of agency, assessed through questions such as Q17 (sense of control) and Q18 (logical consequences of actions), did not show significant differences across the control modes. Although the XR modes offered intuitive interfaces and a slightly higher perception of control, these differences were not substantial enough to conclude a clear advantage over the manual mode. Overall, participants did not perceive substantial changes in their ability to effectively influence the system, suggesting that the type of control had minimal impact on this dimension. See Figure 7.
- **Trust:** This item, assessed through items such as Q19 (comfort with robot movements) and Q21 (system reliability), did not show substantial differences between the three modes of control as can be seen in Figure 8: RE, RS, and SS. While participants expressed slightly higher levels of trust in the XR modes, the differences were not pronounced enough to suggest a clear preference for one mode over another. This lack of significant variation could be attributed to several factors. First, participants may have had similar expectations for system reliability across all modes, as the robot's fundamental capabilities and behaviors did not change significantly

between modes. They also were in a controlled environment, and this could have helped to avoid feelings of distrust. Additionally, the XR environment, despite introducing perceptual differences (e.g., wearing a headset and viewing the physical world through a screen), likely maintained sufficient predictability and consistency to prevent substantial trust disparities.

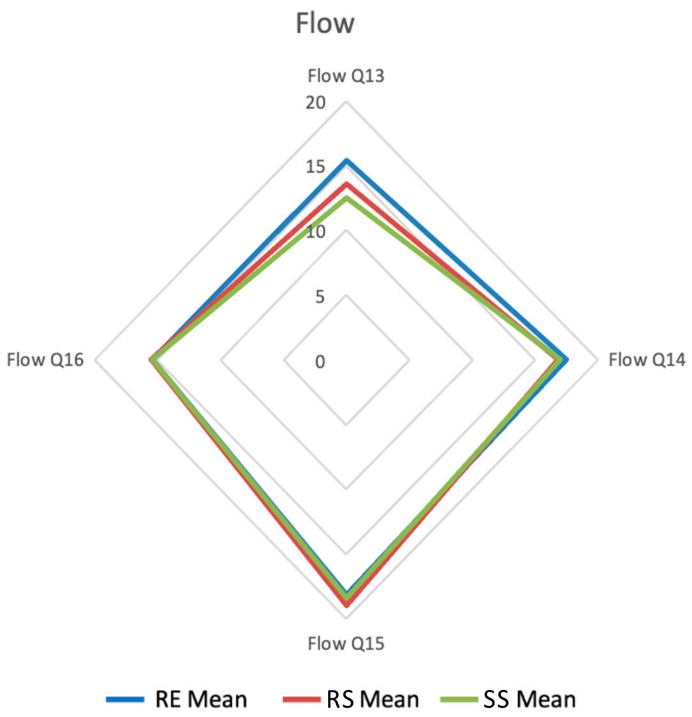

Figure 6. Flow state results extracted from the questionnaire.

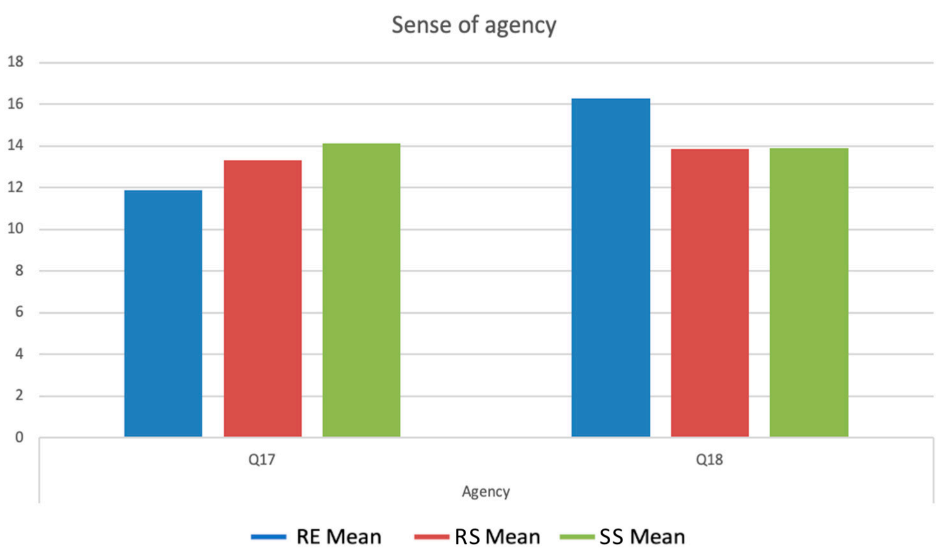

Figure 7. Agency results extracted from the questionnaire.

Another possible cause is the relatively short duration of the tasks, which may have limited participants' ability to fully experience and evaluate differences in system reliability and comfort. Overall, the relatively uniform levels of trust across modes suggest that XR technologies, despite their added complexity, do not negatively impact users' confidence in system performance. However, it also indicates that improvements in XR design, particularly in addressing physical discomfort or perceptual challenges, might help further enhance trust in these systems.

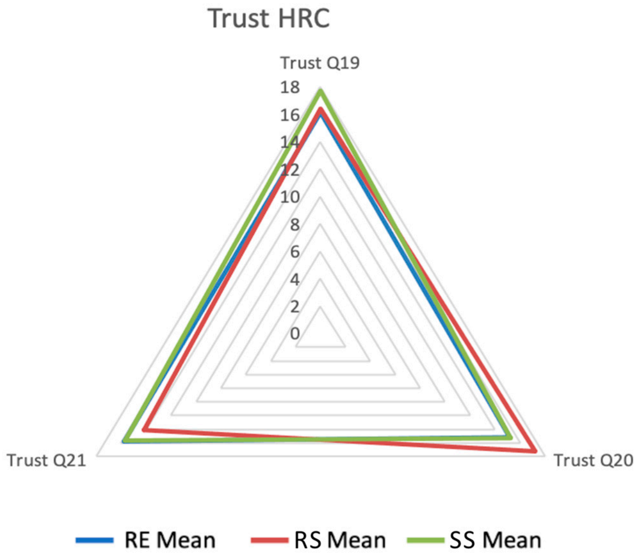

Figure 8. Trust in HRC results extracted from the questionnaire.

- **Correlation among variables:** The following section analyzes the correlation matrices obtained from the questionnaire data under the three experimental conditions (RE, RS, and SS), with a particular focus on identifying meaningful patterns of relationships among the different questions. In case of need, the correlation matrixes of all three cases can be found in the Supplementary Materials. In line with established guidelines for interpreting correlation magnitudes [49] we considered coefficients exceeding 0.70 to represent strong relationships.

When examining the manual control condition (RE), the correlation structure suggests a more fragmented pattern, with fewer strongly interrelated clusters of questions. In RE, some expected relationships between ease of use (Q2), self-sufficiency (Q3), system integration (Q4), and trust (Q21) are weaker and less consistent. This fragmentation implies that participants' perceptions of usability, workload, and trust do not strongly coalesce into a unified, easily interpretable set of dimensions under manual control. Moreover, the most prominent correlation is a negative one: under manual conditions, trust (Q21) demonstrates a strong negative association with perceived unpredictability (Q18) (r ≈ -0.87). This relationship suggests that trust in the system's reliability (Q21) cannot coexist with perceived unpredictability (Q18). The more participants feel the robot's responses make sense, the more they trust it to do what it is supposed to do. Conversely, unpredictable or illogical consequences severely undermine trust.

In contrast, when comparing these patterns to the RS and SS conditions, some shifts appear. Under XR conditions, both at normal and safety speeds, items related to system understanding, performance satisfaction, and trust show intensified clustering and generally higher correlation values. For instance, in RS, Q2 (ease of use) and Q3 (need for technician support) become highly correlated (r $\approx$ 0.75), reinforcing the notion that perceived complexity and self-sufficiency remain a core dimension of the UX. Moreover, Q4 (well-integrated functions) correlates strongly with Q10 (successful performance), Q12 (feeling stressed/irritated), Q17 (full control), Q18 (logical consequences), and Q21 (trust), often with correlations above r $\approx$ 0.75. This suggests that when participants perceive the RS system's functions as coherent and well-integrated, they also report higher satisfaction with their performance, greater feelings of control and reliability, and fewer negative emotional responses, reflecting a more cohesive and psychologically supportive interaction experience.

Similarly, in SS conditions, the patterns mirror those of RS, with slight variations. The correlations remain high between Q2 and Q3, and Q4 also emerges as a strong hub connecting trust, performance, and perceived control. The similarity between RS and SS matrices indicates that the XR-based conditions introduce a more stable, internally consistent perception of the system. Participants seem to integrate their judgments of usability, emotional comfort, and trust into closely linked constructs, in contrast to the more dispersed associations observed under manual control.

These findings suggest that transitioning from manual to XR-based control methods fosters a more cohesive network of user perceptions, ultimately contributing to a more integrated understanding of the system's usability, reliability, and emotional impact.

5.3. Performance Metrics

- **Task Completion Time**

As can be seen in Figure 9, significant differences in task completion times were observed among the three modes across all tasks. For Task 1, participants completed the task significantly faster in the RS mode (M = 45.67 s, SD = 17.65) compared to the RE (M = 89.71 s, SD = 19.41) and SS modes (M = 84.14 s, SD = 27.97), with a *p*-value of 0.0001.

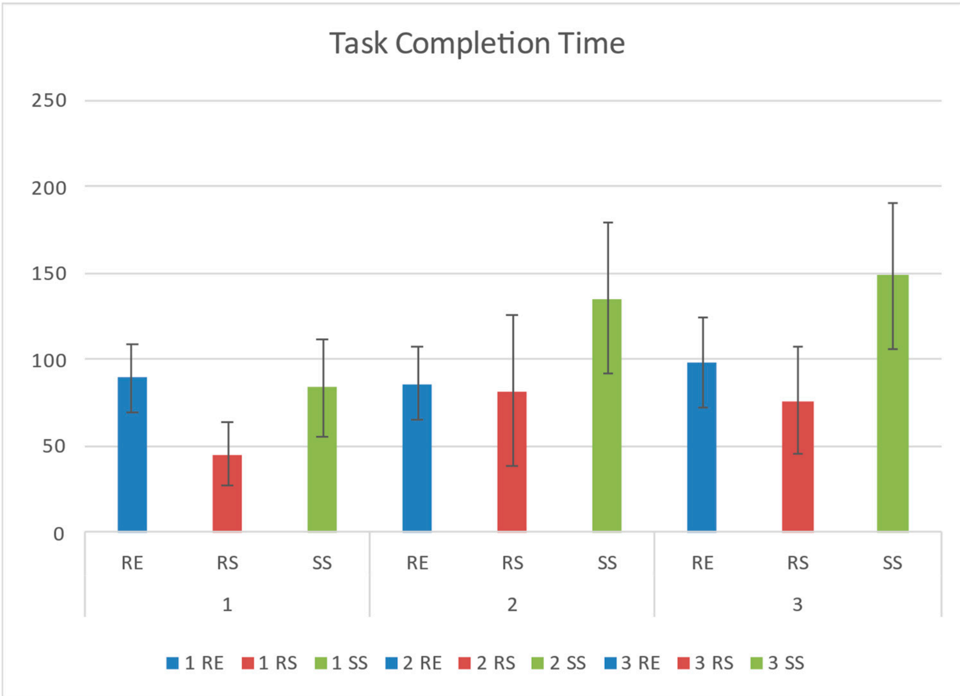

Figure 9. Comparison of task completion times.

In Task 2, the SS mode resulted in longer completion times (M = 135.57 s, SD = 43.31) compared to the RE (M = 86.48 s, SD = 20.88) and RS modes (M = 81.62 s, SD = 43.60), *p* = 0.000025. Similarly, for Task 3, participants were faster in the RS mode (M = 76.19 s, SD = 31.14) than in the RE (M = 98.43 s, SD = 25.90) and SS modes (M = 148.76 s, SD = 42.34), with a *p*-value of 0.0001.

These results indicate that the RS mode generally allowed for quicker task completion, while the SS mode, operating at reduced speed, increased completion times, especially in more complex tasks.

- **Number of Attempts**

The number of attempts required to complete the tasks differed significantly in Tasks 2 and 3 (see Figure 10). In Task 2, participants required more attempts in the RS (M = 2.05, SD = 1.25) and SS modes (M = 1.71, SD = 1.16) compared to the RE mode (M = 1.00, SD = 0.00), with a p-value of 0.0048. For Task 3, a similar pattern was observed: RS (M = 1.90, SD = 1.02) and SS modes (M = 1.90, SD = 1.27) required more attempts than the RE mode (M = 1.05, SD = 0.21), p = 0.013.

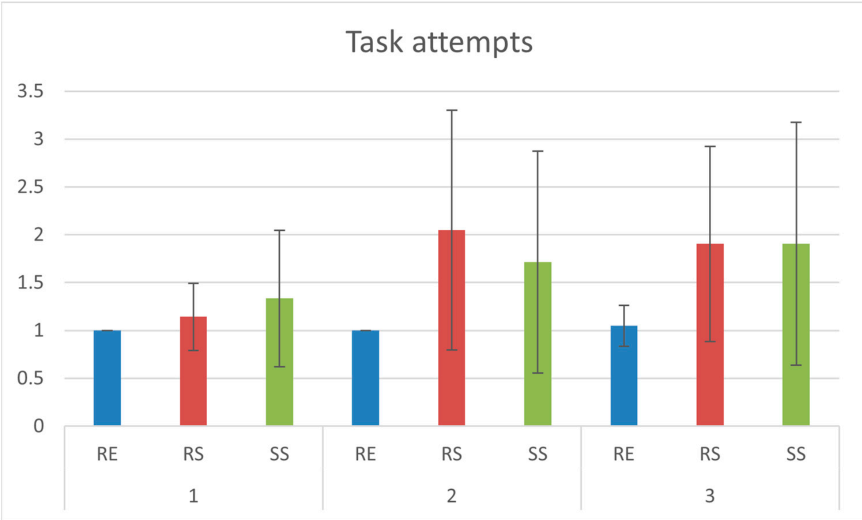

Figure 10. Task attempt comparison.

No significant differences were found in Task 1 (p = 0.591), suggesting that the mode of control did not affect the number of attempts in simpler tasks.

- **Number of Errors**

Significant differences in the number of errors were also found in Tasks 2 and 3 (see Figure 11). In Task 2, participants made more errors in the RS (M = 2.29, SD = 2.14) and SS modes (M = 1.67, SD = 1.64) compared to the RE mode (M = 0.62, SD = 1.09), with a p-value of 0.0026. In Task 3, errors were again higher in the RS (M = 2.33, SD = 0.82) and SS modes (M = 2.14, SD = 1.81) than in the RE mode (M = 0.71, SD = 0.82), p = 0.0001.

Figure 11. Error count in different HRI modes.

For Task 1, no significant differences were observed in the number of errors across modes (p = 0.642), indicating that task simplicity may mitigate error rates regardless of control mode.

The performance analysis reveals that while the RS mode enhances efficiency by reducing task completion times, it may also increase the number of attempts and errors in more complex tasks. The SS mode, designed to improve safety by reducing operational speed, resulted in longer completion times without a significant reduction in errors compared to the RS mode.

In Tasks 2 and 3, which were more complex, participants under RS control modes (both RS and SS) made more errors and required more attempts than under manual control. This suggests that while RS control facilitates faster performance, it may compromise accuracy and precision in tasks that demand higher levels of dexterity or cognitive processing.

5.4. Differences Among Segmentations

Apart from analyzing the results altogether, different participant profiles have been analyzed in search of differences among them.

In the manual control mode (RE), participants without XR experience reported higher perceptions of laboriousness (Q6) and discomfort (Q19, Q20), highlighting the physical and cognitive effort required for manual operation. However, robot-experienced users tended to rate the manual mode more favorably for ease of use (Q2) and learning (Q5). This suggests that while manual control offers more precision for experienced participants, it remains challenging for those less familiar with such systems, particularly those unfamiliar with XR interfaces, as previous experiences affect the overall experience.

In the RS control mode (normal speed), trends again show that participants without XR experience found the system more demanding, reporting higher values for laboriousness (Q6) and discomfort (Q19, Q20). This indicates an initial adaptation barrier to the RS interface for inexperienced users. By contrast, robot-experienced participants reported greater ease of use (Q2) and a shorter learning curve (Q5). Interestingly, concerns regarding speed (Q20) were less pronounced in this group, likely due to their ability to anticipate and manage robotic trajectories more effectively. Despite these minor challenges for novices, RS was generally viewed as reducing physical strain compared to manual control.

The safety speed (SS) mode received mixed feedback. Robot-experienced participants generally adapted well to the slower, safety-focused mode, reporting lower perceptions of effort (Q6) and fewer concerns about speed (Q20). However, XR-inexperienced users reported higher levels of laboriousness and frustration, particularly with the slower speed. While SS successfully reduces physical strain, its perceived lack of responsiveness may have led inexperienced users to view it as inefficient or less intuitive to control.

In summary, we can say that robot-experienced participants consistently adapted better to all modes, benefiting from familiarity with robotic systems, which allowed them to manage both speed and control more effectively. In contrast, XR-inexperienced users faced more challenges, particularly in manual control and slower XR modes, where the learning curve and perceptions of inefficiency contributed to increased discomfort and effort. Although the differences are not substantial, the results suggest that further refinements (such as better calibration, speed responsiveness, and intuitive feedback mechanisms) are needed to optimize XR systems for users with varying experience levels.

5.5. Overall Preference Metrics

Regarding the final question on user preferences among the three control modes, the mean rankings revealed a clear hierarchy of preference. The RS mode received the highest preference with the lowest mean ranking (M = 1.38), followed by the SS mode (M = 2.00), and the RE mode had the highest mean ranking (M = 2.62), indicating the least preference. The median rankings reinforced these findings, with RS having a median of 1, SS having a median of 2, and RE having a median of 3. The modes of the rankings were consistent with the medians: RS (Mode = 1), SS (Mode = 2), and RE (Mode = 3).

Chi-squared goodness-of-fit tests were performed to determine whether the observed ranking frequencies differed significantly from a uniform distribution (i.e., no preference among the modes). The results were statistically significant for all three modes:

- Manual control (RE): $\chi^2 = 14.00$, $p = 0.000456$
- XR control normal speed (RS): $\chi^2 = 14.00$, $p = 0.000456$
- XR control safety speed (SS): $\chi^2 = 7.71$, $p = 0.01056$

The p-values for all modes were below the conventional alpha level of 0.05, indicating that the observed rankings were not due to chance and that there were significant preferences among the modes [50].

The ranking data in Figure 12 reveal a strong user preference for the XR control at normal speed (RS). A majority of participants (15 out of 21) ranked RS as the best mode, underscoring its favorable reception. In contrast, the manual control mode (RE) was predominantly ranked as the worst, with 15 participants assigning it the lowest rank. The RS control at safety speed (SS) was mostly ranked in the middle position, suggesting a moderate preference.

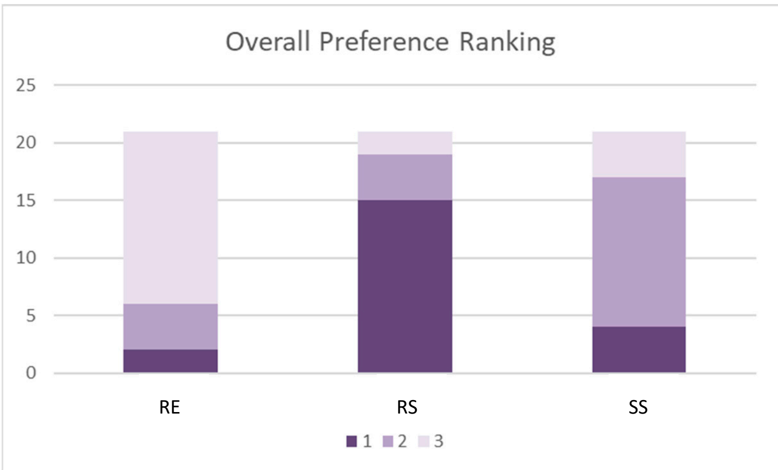

Figure 12. Overall preference ranking of robot manipulation modes.

When incorporating qualitative data from participants' free-text responses, two recurring themes clearly emerge. The first concerns the manual control mode (RE), which numerous participants describe as physically demanding or requiring excessive force. Some participants explicitly state that it was "hard to use due to physical force needed" or that they became "exhausted" when controlling the robot manually. Others note that manual manipulation was "physically challenging" and "laborious", and that it would likely prevent certain demographic groups (e.g., elderly individuals, children, or those with limited physical ability) from using the system comfortably. These comments align with the ranking data, where participants were predominantly assigned the manual control mode as the lowest rank.

A second theme relates to the safety speed in the XR mode (SS), which participants repeatedly describe as "too slow" and "boring". Several comments highlight frustration with waiting for the robot to complete its tasks, with one user stating explicitly that "safety speed was a little too slow" another calling it "unnecessarily slow", and others reporting feelings of annoyance and boredom. These qualitative impressions further support the ranking data, where the safety speed mode was often placed in an intermediate position, reflecting a compromise between control and comfort but not an entirely preferred solution.

5.6. Minimal Detectable Effect (MDE) Analysis

While the above results indicate several significant differences among the three modes (RE, RS, SS), we also conducted an approximate minimal detectable effect (MDE) analysis to gauge the sensitivity of our repeated-measures design. In essence, the MDE represents the smallest true effect that the study, given its sample size and design, can reliably detect.

Using a simplified paired t-test approximation (assuming $\alpha = 0.05$, power = 0.80, and a moderate correlation r = 0.50 among repeated measures), we estimated an MDE corresponding to Cohen's d = 0.60—commonly viewed as a moderate to large effect size. Interpreted practically, this means our study design (N = 21 participants, all of whom experienced each mode) has about an 80% chance of detecting differences of at least 0.60 standard deviations in measures such as task completion time, error rates, and questionnaire ratings.

Consequently, the significant results we observed (e.g., Q1, Q2, Q5, completion times) likely reflect effects that reach or exceed this moderate threshold. The non-significant findings in the remaining 15 questionnaire items may indicate either genuinely negligible differences or differences smaller than our study could reliably detect. Similarly, the pronounced preferences revealed by the chi-squared tests (e.g., strong favoring of RS mode) align with effects that appear larger than our estimated MDE.

In summary, although a sample of 21 participants in a within-subjects design provided a feasible approach for hands-on robotic testing, it remains possible that smaller, subtler effects were not captured.

6. Discussion

This study provides valuable insights into the central research question, which investigates whether XR technologies enhance HRI in robotic path programming by improving technical performance and operator experience. The findings from the questionnaires, performance metrics, and ranking analysis collectively highlight a nuanced understanding of user preferences and the efficacy of XR technologies in robotic manipulation tasks.

The questionnaire results demonstrated a clear preference for XR control modes over manual control. Six out of twenty-one items showed significant differences favoring the XR modes, particularly in aspects related to ease of use, reduced physical effort, and overall desire to use the system frequently. Participants perceived the XR modes as less laborious and more user-friendly, aligning with previous research suggesting that XR interfaces enhance user satisfaction by simplifying interactions and reducing physical strain [9,34].

The ranking analysis further substantiated this preference. Most participants ranked the RS mode as their top choice, indicating a strong inclination towards XR control at normal speed. Despite the SS mode being designed with safety considerations through reduced speed, it was predominantly ranked second, suggesting that while safety is valued, efficiency and ease of use are more influential in determining user preference.

- The performance data presented a more complex picture. While the RS mode significantly reduced task completion times, particularly in simpler tasks, it also resulted in an increased number of attempts and errors in more complex tasks (Tasks 2 and 3). The SS mode, intended to mitigate errors through slower operation, did not significantly reduce errors compared to the RS mode and led to longer completion times. These errors stemmed from both the system and the interaction itself. Due to calibration and the robot's 3D model, the system had a safety offset and an additional offset from a scale error of about 1:1.02. As a result, at certain points in the circuit, the robot would touch an obstacle even when participants executed the task correctly. Nevertheless, the most frequent error was the issue with hand tracking. In some cases, because of the participants' hand position, the headset would lose tracking and the tracing would stop, sometimes requiring them to start over. Since these errors were largely system-based, testing made it possible to identify them, as well as the fact that tracking becomes more complicated with smaller hands. These improvements will be implemented at a later stage.

The discrepancy between user preference and performance outcomes suggests that participants favored the XR modes despite experiencing higher error rates and requiring more attempts in complex tasks. One possible explanation is that users prioritize the intuitive and engaging nature of XR interfaces over objective performance metrics, a phenomenon observed in technology acceptance studies [16]. The immersive experience and reduced physical effort associated with XR control may enhance user satisfaction to the extent that they overlook the increased errors and attempts.

The preference for XR modes, even in the face of suboptimal performance metrics, indicates a strong user acceptance of XR technology in robotic manipulation. However, the increased errors and attempts in complex tasks highlight areas for improvement. Calibration issues, interface design, and system responsiveness may contribute to these performance deficits. Enhancing the precision and reliability of XR systems could mitigate errors, thereby aligning user satisfaction with optimal performance outcomes.

- The findings suggest that users are willing to tolerate certain inefficiencies in exchange for the benefits offered by XR control. This tolerance underscores the importance of user-centered design in XR technologies, where the focus is on creating interfaces that are not only functional but also engaging and easy to use [17]. Developers should consider incorporating adaptive features that adjust to task complexity, providing additional guidance or automation in more challenging tasks to reduce errors.

7. Conclusions

While this study has several limitations, it provides valuable insights into the role of XR technologies in enhancing human–robot interaction for robotic path programming. One primary limitation of this study is the modest sample size ($n = 21$) in a repeated-measures design, which influences the minimal detectable effect (MDE). A post hoc approximation suggested that our design can reliably detect moderate or larger effects (roughly Cohen's $d \geq 0.60$) with about 80% power. Consequently, the significant results we observed for certain questionnaire items and performance metrics likely reflect differences at or above this threshold. However, non-significant findings do not necessarily imply the absence of any effect; smaller differences (below the MDE) may remain undetected due to limited statistical power. Future studies with larger or more diverse samples could help determine whether subtler differences exist, particularly for questionnaire items and performance measures that showed non-significant trends in the present work.

In addition, the study was conducted in a controlled laboratory environment rather than in a real-world industrial workplace. While the controlled setting allowed for precise measurements and minimized extraneous variables, it may not fully capture the complexity, unpredictability, and dynamic interactions present in actual industrial settings. This limitation could affect the generalizability of the results to real-world applications. Furthermore, although our participant group was intentionally selected to reflect the technical background common in industrial contexts, the relatively homogenous sample (primarily Computer Science and Robotics students) may not fully represent the broader population of industrial workers. Finally, potential carryover effects or learning effects in the repeated-measures design, despite counterbalancing, may also influence the observed outcomes.

Future research should address these limitations by employing larger, more diverse samples in field-based studies to validate and extend these findings under real-world conditions.

Besides the aforementioned limitations, this study demonstrates that XR technologies hold significant promise for advancing human–robot interaction in robotic path programming. The results demonstrate a clear preference for XR-based controls, with participants highlighting their ease of use, reduced physical effort, and overall superior usability. Among these, XR control at normal speed (RS) was the most favored due to its balance of efficiency and user-friendliness. This preference underscores the potential of XR technologies to provide intuitive and engaging interaction paradigms in robotic manipulation tasks.

In terms of task performance, XR control significantly reduced task completion times for simpler tasks, showcasing its ability to enhance efficiency. However, in more complex scenarios requiring higher levels of precision and dexterity, XR control led to increased errors and attempts. The safety speed mode mitigated some of these errors but at the cost of prolonged task durations, illustrating the trade-off between operational speed and accuracy. These findings suggest that while XR control improves performance in certain contexts, its application in complex tasks requires further refinement to address challenges such as system calibration and interface responsiveness.

The study also revealed that XR systems excel in usability by reducing both physical and cognitive demands compared to manual control. Participants reported feeling less fatigued and more confident in using XR technologies, which aligns with their broader potential to streamline interactions in industrial settings. Despite these advantages, tasks requiring heightened precision highlighted areas for improvement in XR system design, particularly in achieving greater alignment between user actions and system outputs.

Interestingly, trust and perceived control, or agency, remained relatively consistent across all control modes, suggesting that XR technologies do not inherently compromise user confidence. This finding is significant as it indicates that, even in its current state, XR can support user engagement without introducing uncertainty or discomfort. Nevertheless, further enhancements in the design of XR systems could bolster trust and agency, particularly in scenarios demanding high reliability.

To build on these results, future research should incorporate biometric analyses to complement subjective evaluations. Physiological data can offer objective insights into user experiences, potentially revealing differences between interaction modes that are not fully captured by current metrics. By leveraging these data, researchers can develop adaptive XR systems that optimize usability and performance across a range of task complexities.

Overall, the findings emphasize the promise of XR technologies in fostering human-centric interaction in Industry 5.0. While XR control modes offer distinct advantages in terms of usability and UX, their integration into complex industrial tasks will benefit from additional refinements. Optimizing XR systems to balance efficiency, precision, and reliability will be key to unlocking their full potential in advancing seamless human–robot collaboration in diverse applications.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/mti9030027/s1, Figure S1: Correlation Matrixes of the UX questionnaire.

Author Contributions: Conceptualization, O.E., N.O., G.L. and M.M.; methodology, O.E., N.O., Y.Y. and F.D.; formal analysis, O.E. and N.O.; resources, Y.Y., F.D., S.B. and E.U.; data curation, O.E. and N.O., writing—original draft preparation, O.E. and N.O.; writing—review and editing, O.E., N.O., G.L. and M.M.; supervision, G.L., M.M. and E.U.; funding acquisition, G.L., M.M. and E.U. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Commission, Horizon Europe project INteractive robots that intuitiVely lEarn to inVErt tasks by ReaSoning about their Execution (INVERSE)—GA 101136067.

Institutional Review Board Statement: The study was conducted in accordance with the Research Ethics Committee of Mondragon Unibertsitatea. I.C: IEB-20241118. (18 November 2024). It was approved by Bogazici University's Ethics Committee. E-84391427-050.04-211747.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets presented in this article are not readily available because the data are part of an ongoing study. Requests to access the datasets should be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Questionnaires

Appendix A.1. Socio-Demographic Questionnaire

Next, you will be asked questions of a demographic nature. Please fill in your data.
What gender do you identify with?

- o Woman
- o Man
- o Prefer not to say
- o Other

Indicate your age:
Which is your role?
Do you use eyeglasses or contact lenses?

- o Yes, eyeglasses
- o Yes, contact lenses
- o No

Select the answer that best fits your previous experience with a robot.

- o I have never interacted with a robot.
- o I have previously interacted at least once with a robotic arm while executing a task.
- o I have experience in interacting with a robotic arm while executing a task.

Select the answer that best fits your previous experience with Augmented Reality (AR).

- o I have never interacted with AR and did not know about it until now
- o I have never interacted with AR, but I know what it is.
- o I have interacted at least once with AR via mobile/tablet.
- o I have interacted at least once with AR through glasses.

Appendix A.2. Questionnaire for Qualitative Assessment for UX

Questions should be answered in a 21-point Likert scale.

1. I think that I would like to use this system frequently.
2. I thought the system was easy to use.
3. I think that I would need the support of a technician person to be able to use this system.
4. I found the various functions in this system were well integrated.
5. I would imagine that most people would learn to use this system very quickly.
6. I found the system very laborious to use.

7. The mental and perceptual activity required was high (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.).
8. The physical activity required was high (e.g., pushing, pulling, turning, controlling, activating, etc.).
9. I felt high time pressure due to the rate or pace at which the tasks occurred.
10. I think I was successful in accomplishing the goal of the task set by the experimenter. I am satisfied with my performance.
11. I had to work hard (mentally and physically) to accomplish my level of performance.
12. I felt insecure, discouraged, irritated, stressed and annoyed during the task.
13. I was challenged, but I believed my skills would allow me to meet the challenge.
14. I knew clearly what I wanted to do.
15. My attention was focused entirely on what I was doing.
16. I was aware of how well I was performing.
17. I am in full control of what I do.
18. The consequences of my actions feel like they don't logically follow my actions.
19. The way in which the robot moved made me feel uncomfortable.
20. The speed with which the robot picked and released the components made me feel uneasy.
21. I felt I could rely on the robot to do what it was supposed to do.

** Extra question after all 3 modes were used:

Classify from 1 (most suitable) to 3 (less suitable) the 3 ways you interacted with the robot and explain why. Add any kind of comment regarding the experiment aswell.

Robot controlled manually

Robot controlled with AR (same speed)

Robot controlled with AR (different speed)

Comments:

References

1. Gao, Z.; Wanyama, T.; Singh, I.; Gadhrri, A.; Schmidt, R. From industry 4.0 to robotics 4.0—A conceptual framework for collaborative and intelligent robotic systems. In *Procedia Manufacturing*; Elsevier B.V.: Amsterdam, The Netherlands, 2020; pp. 591–599. [CrossRef]
2. Soori, M.; Dastres, R.; Arezoo, B.; Jough, F.K.G. Intelligent robotic systems in Industry 4.0: A review. *J. Adv. Manuf. Sci. Technol.* **2024**, *4*, 2024007-0. [CrossRef]
3. Xu, M.; David, J.M.; Kim, S.H. The Fourth Industrial Revolution: Opportunities and Challenges. *Int. J. Financ. Res.* **2018**, *9*, 90–95. [CrossRef]
4. Sheridan, T.B. *Human-Robot Interaction*; SAGE Publications Inc.: Thousand Oaks, CA, USA, 2016. [CrossRef]
5. Wang, X.V.; Kemény, Z.; Váncza, J.; Wang, L. Human–robot collaborative assembly in cyber-physical production: Classification framework and implementation. *CIRP Ann. Manuf. Technol.* **2017**, *66*, 5–8. [CrossRef]
6. Kemény, Z.; Beregi, R.; Nacsa, J.; Kardos, C.; Horváth, D. Human–robot collaboration in the MTA SZTAKI learning factory facility at Győr. *Procedia Manuf.* **2018**, *23*, 105–110. [CrossRef]
7. Nahavandi, S. Industry 5.0-a human-centric solution. *Sustainability* **2019**, *11*, 4371. [CrossRef]
8. Speicher, M.; Hall, B.D.; Nebeling, M. What is mixed reality? In Proceedings of the Conference on Human Factors in Computing Systems—Proceedings, Glasgow, Scotland, 4–9 May 2019. [CrossRef]
9. Azuma, R.T. A Survey of Augmented Reality. 1997. Available online: http://www.cs.unc.edu/~azumaW (accessed on 4 December 2024).
10. Walker, M.; Phung, T.; Chakraborti, T.; Williams, T.; Szafir, D. Virtual, Augmented, and Mixed Reality for Human-robot Interaction: A Survey and Virtual Design Element Taxonomy. *ACM Trans. Hum. Robot. Interact.* **2023**, *12*, 1–39. [CrossRef]

11. Jiang, X.; Mattes, P.; Jia, X.; Schreiber, N.; Neumann, G.; Lioutikov, R. A Comprehensive User Study on Augmented Reality-Based Data Collection Interfaces for Robot Learning. In Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction, Boulder, CO, USA, 11–15 March 2024; pp. 333–342. [CrossRef]
12. Chen, J.; Salas, D.A.C.; Bilal, M.; Zhou, Q.; Johal, W., Mr. LfD: A Mixed Reality Interface for Robot Learning from Demonstration. 2024. Available online: https://github.com/CHRI-Lab/MrLfD_Hub (accessed on 4 December 2024).
13. Zhao, F.; Deng, W.; Pham, D.T. A Robotic Teleoperation System with Integrated Augmented Reality and Digital Twin Technologies for Disassembling End-of-Life Batteries. *Batteries* **2024**, *10*, 382. [CrossRef]
14. Steinfeld, A.; Fong, T.; Kaber, D.; Lewis, M.; Scholtz, J.; Schultz, A.; Goodrich, M. Common metrics for human-robot interaction. In Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction, Salt Lake City, UT, USA, 2–3 March 2006; pp. 33–40. [CrossRef]
15. Goodrich, M.A.; Schultz, A.C. Human-robot interaction: A survey. *Found. Trends® Hum. Comput. Interact.* **2007**, *1*, 203–275. [CrossRef]
16. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* **1989**, *13*, 319–339. [CrossRef]
17. Norman, D.A. *The Design of Everyday Things: Revised and Expanded Edition*; Basic Books: New York, NY, USA, 2013.
18. Villani, V.; Pini, F.; Leali, F.; Secchi, C. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics* **2018**, *55*, 248–266. [CrossRef]
19. Haddadin, S.; Croft, E. Physical Human–Robot Interaction. In *Springer Handbooks*; Springer Science and Business Media Deutschland GmbH: Berlin/Heidelberg, Germany, 2016; pp. 1835–1874. [CrossRef]
20. Bauer, A.; Wollherr, D.; Buss, M. Human-robot collaboration: A survey. *Int. J. Humanoid Robot.* **2008**, *5*, 47–66. [CrossRef]
21. Lasota, P.A.; Rossano, G.F.; Shah, J.A. Toward safe close-proximity human-robot interaction with standard industrial robots. In Proceedings of the IEEE International Conference on Automation Science and Engineering, New Taipei, Taiwan, 18–22 August 2014; pp. 339–344. [CrossRef]
22. Delgado JM, D.; Oyedele, L.; Ajayi, A.; Akanbi, L.; Akinade, O.; Bilal, M.; Owolabi, H. Robotics and automated systems in construction: Understanding industry-specific challenges for adoption. *J. Build. Eng.* **2019**, *26*, 100868. [CrossRef]
23. Tashtoush, T.; Garcia, L.; Landa, G.; Amor, F.; Laborde, A.N.; Oliva, D.E.; Safar, F.G.E. Human-Robot Interaction and Collaboration (HRI-C) Utilizing Top-View RGB-D Camera System. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 10–17. [CrossRef]
24. Lasota, P.A.; Fong, T.; Shah, J.A. A Survey of Methods for Safe Human-Robot Interaction. *Found. Trends® Robot.* **2017**, *5*, 261–349. [CrossRef]
25. Zacharaki, A.; Kostavelis, I.; Gasteratos, A.; Dokas, I. *Safety Bounds in Human Robot Interaction: A Survey*; Elsevier B.V.: Amsterdam, The Netherlands, 2020. [CrossRef]
26. Akalin, N.; Kiselev, A.; Kristoffersson, A.; Loutfi, A. A Taxonomy of Factors Influencing Perceived Safety in Human–Robot Interaction. *Int. J. Soc. Robot.* **2023**, *15*, 1993–2004. [CrossRef]
27. Lichte, D.; Wolf, K.-D. Use case-based consideration of safety and security in cyber physical production system. In *Safety and Reliability-Safe Societies in A Changing World*; Taylor & Francis: Abingdon, UK, 2018; pp. 1395–1401.
28. Hoffman, G. Evaluating Fluency in Human-Robot Collaboration. *IEEE Trans. Hum. Mach. Syst.* **2019**, *49*, 209–218. [CrossRef]
29. Wang, X.; Shen, L.; Lee, L.-H. Towards Massive Interaction with Generalist Robotics: A Systematic Review of XR-enabled Remote Human-Robot Interaction Systems. *arXiv* **2024**, arXiv:2403.11384. [CrossRef]
30. Karpichev, Y.; Charter, T.; Hong, J.; Enayati AM, S.; Honari, H.; Tamizi, M.G.; Najjaran, H. Extended Reality for Enhanced Human-Robot Collaboration: A Human-in-the-Loop Approach. In Proceedings of the 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN), Pasadena, CA, USA, 26–30 August 2024; pp. 1991–1998. [CrossRef]
31. Brooke, J. SUS: A Quick and Dirty Usability Scale. 1995. Available online: https://www.researchgate.net/publication/228593520 (accessed on 5 December 2024).
32. Kim, Y.M.; Rhiu, I.; Yun, M.H. A Systematic Review of a Virtual Reality System from the Perspective of User Experience. *Int. J. Hum. Comput. Interact.* **2020**, *36*, 893–910. [CrossRef]
33. Hart, S.; Stavenland, L. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Adv. Psychol.* **1988**, *52*, 139–183.
34. Dünser, A.; Grasset, R.; Billinghurst, M. survey of evaluation techniques used in augmented studies. In Proceedings of the ACM SIGGRAPH ASIA 2008 Courses, SIGGRAPH Asia'08, Singapore, 10–13 December 2008. [CrossRef]
35. Csikszentmihalyi, M. Flow: The Psychology of Optimal Experience. 1990. Available online: https://www.researchgate.net/publication/224927532 (accessed on 4 December 2024).
36. Nakamura, J.; Csikszentmihalyi, M. Flow Theory and Research. In *The Oxford Handbook of Positive Psychology*, 2nd ed.; Oxford University Press: Oxford, UK, 2012. [CrossRef]
37. Michailidis, L.; Balaguer-Ballester, E.; He, X. Flow and immersion in video games: The aftermath of a conceptual challenge. *Front. Psychol.* **2018**, *9*, 1682. [CrossRef]

38. Hassan, L.; Jylhä, H.; Sjöblom, M.; Hamari, J. Flow in VR: A Study on the Relationships Between Preconditions, Experience and Continued Use. In Proceedings of the Annual Hawaii International Conference on System Sciences, Maui, HI, USA, 7–10 January 2020; pp. 1196–1205. [CrossRef]
39. Moore, J.W.; Middleton, D.; Haggard, P.; Fletcher, P.C. Exploring implicit and explicit aspects of sense of agency. *Conscious. Cogn.* **2012**, *21*, 1748–1753. [CrossRef] [PubMed]
40. Caspar, E.A.; Desantis, A.; Dienes, Z.; Cleeremans, A.; Haggard, P. The sense of agency as tracking control. *PLoS ONE* **2016**, *11*, e0163892. [CrossRef]
41. Beck, B.; Di Costa, S.; Haggard, P. Having control over the external world increases the implicit sense of agency. *Cognition* **2017**, *162*, 54–60. [CrossRef] [PubMed]
42. Tapal, A.; Oren, E.; Dar, R.; Eitam, B. The sense of agency scale: A measure of consciously perceived control over one's mind, body, and the immediate environment. *Front. Psychol.* **2017**, *8*, 1552. [CrossRef]
43. Meta, "Meta Quest 3", Meta Platforms, Menlo Park, CA, USA. 2023. Available online: https://www.meta.com/quest/quest-3/ (accessed on 1 December 2024).
44. Dogangun, F.; Bahar, S.; Yildirim, Y.; Temir, B.T.; Ugur, E.; Dogan, M.D. RAMPA: Robotic Augmented Reality for Machine Programming and Automation. *arXiv* **2024**, arXiv:2410.13412. [CrossRef]
45. Pearse, N. Deciding on the scale granularity of response categories of likert type scales: The case of a 21-point scale. *Electron. J. Bus. Res. Methods* **2011**, *9*, 159–171. Available online: www.ejbrm.com (accessed on 6 December 2024).
46. NASA Task Load Index. Available online: https://digital.ahrq.gov/health-it-tools-and-resources/evaluation-resources/workflow-assessment-health-it-toolkit/all-workflow-tools/nasa-task-load-index (accessed on 30 November 2023).
47. Jackson, S.; Marsh, H.W. Development and Validation of a Scale to Measure Optimal Experience: The Flow State Scale. *J. Sport Exerc. Psychol.* **1996**, *18*, 17–35. [CrossRef]
48. Charalambous, G.; Fletcher, S.; Webb, P. The Development of a Scale to Evaluate Trust in Industrial Human-robot Collaboration. *Int. J. Soc. Robot.* **2016**, *8*, 193–209. [CrossRef]
49. Evans, R.H. An Analysis of Criterion Variable Reliability in Conjoint Analysis. *Percept. Mot. Ski.* **1996**, *82*, 988–990. [CrossRef]
50. Toothaker, L.E. Book Review: Nonparametric Statistics for the Behavioral Sciences (Second Edition): Sidney Siegel and N. John Castellan, Jr. New York: McGraw-Hill, 1988, 399 pp., approx. $47.95. *Appl. Psychol. Meas.* **1989**, *13*, 217–219. [CrossRef]