VR/MR Systems Integrated with Heat Transfer Simulation for Training of Thermoforming: A Multicriteria Decision-Making User Study

Iman Jalilvand, Jiyoung Jang, Bhushan Gopaluni*, Abbas S. Milani*

Materials and Manufacturing Research Institute, University of British Columbia, Canada

Correspondence: abbas.milani@ubc.ca & bhushan.gopaluni@ubc.ca

Abstract

The emerging Industry 5.0 paradigm emphasizes the importance of collaboration between machines and human operators in smart factories, thereby preventing full automation and elimination of the skilled workforce. One technology that has been proposed recently as a means of achieving this collaboration, is the Extended Reality (XR), particularly for operators training and upskilling. However, current XR applications in advanced manufacturing sector are rather limited, mainly focusing on maintenance or sequential training. Additionally, most of earlier studies in the field have examined the use of one type of XR environment solely, such as Virtual Reality (VR), Augmented Reality (AR), or Mixed Reality (MR) in a respective application. The present study aims to develop and compare the performance of both MR and VR environments for training operators during the monitor and control of a given thermoforming manufacturing process. To achieve this, Meta Quest 2 (VR headset) and Microsoft HoloLens 2 (MR headset) were utilized in combination with heat transfer numerical simulations of the process. A video training tool was also created and employed as a reference/conventional training method, for comparisons. A user study involving 26 participants (with a mix of prior experience) was conducted, assessing different performance criteria including the usefulness, trust, reliability, user satisfaction, training effectiveness, and overall user preference over each training tool. Results, using statistical analysis along with five Multiple Criteria Decision Making (MCDM) algorithms, collectively indicated that the users preferred MR over VR and video training in the present case study; particularly as MR provided greater presence experience and satisfaction, higher user-friendliness and reliability. A demonstration of the developed applications can be found in the **Supplementary Material**.

Keywords: Extended Reality; Integrated Heat Transfer Simulation; Immersive Training; User Study; Multiple Criteria Decision-Making.

1. Introduction

Thermoplastic materials and their composites have been increasingly the focus of aerospace, marine, renewables, and automotive industries, to name a few, due to their superior thermo-mechanical properties and lightweight. Nevertheless, several barriers, such as the high cost of manufacturing trials and errors, time-consuming design tools, and low product quality repeatability rates, still restrict some industries from mass-producing advanced thermoplastic structural parts. To overcome such barriers, since the emergence of the 4th, and more recently the 5th industrial revolution (I4.0 and I5.0), various digital transformations of the underlying manufacturing processes have been adapted to increase the overall efficiency and speed of productions and reduce cost [1]. The enabling technologies, such as IoT, Simulation, Machine Learning, and Immersive Technology, are being researched to build Digital Twins (DT) that would allow operators to be efficiently trained, monitor and control the processes in close collaboration with automated/semi-automated machines (which is the primary motivation of I5.0 as compared to I4.0; i.e. bringing back empowered humans to the shop floor). In the present study, the primary focus is on the Immersive Technology component of I5.0 for workforce training, as follows.

1.1. Immersive platforms for manufacturing and workforce training

Over the past decade, the use of Extended Reality (XR) systems has increased in a range of applications, including training, education, and safety [2]. The XR platforms often include augmented reality (AR) as a perspective of the physical world with added computer-generated elements, virtual reality (VR) as a dynamic, computer-generated component within an entirely virtual environment, and mixed reality (MR) as a combination of the real-world and virtual components in an interactive representation. The manufacturing sector is not an exception to this overarching pattern. In particular, unprecedented interests are emerging for the applications of immersive technologies for complex machine operators/workforce training and upskilling.

Among different XR platforms, the use of VR in industrial trainings has been most popular to date [3], [4]. However, it remains as a continuing scientific endeavour to link this platform to, e.g., numerical simulation tools. Prattico et al. [5] argued that using VR training simulators as self-learning tools is still up for debate, even though they have been advantageous as alternatives to the traditional learning materials used by trainers. They performed a user study to determine the effectiveness of VR assistance systems in training KUKA industrial robot operators for routine maintenance duties. The study's generalizability was constrained by a modest sample size. It highlighted the need for further research to enhance skill transfer efficacy in Virtual Reality Training Systems (VRTSs), particularly for interactions with small and delicate objects. Leder et al. [6] examined VR and PowerPoint as means for providing safety training, using variables such as risk perception, learning, and dangerous decisions. Their research discovered an influence of training strategy on risk perception in terms of probability judgements and hazardous choices, but not on learning. The study was conducted among a sample of high-school students, where the participants were restricted to the role of observers, even in the immersive VR mode. Ragan et al. [7] evaluated an XR system that was created to teach individuals a visual scanning technique for identifying targets (armed individuals) in a virtual environment of an urban setting. In that complex scenario, those trained in realistic environments performed the technique more effectively. The study, however, did not quantify the impact of the VR training system fidelity on the training effectiveness itself. In the study conducted by Sportillo et al. [8], it was observed that using Head-

Mounted Displays (HMDs) improved performance of a virtual assembly task. Although the research showed promising preliminary results, no user study was conducted to objectively evaluate the efficacy of training system. López et al. [9] performed multiple case studies to determine the efficiency of VR systems in various training contexts. The study employed a relatively small sample size. Additionally, the complexity of the training scenarios resulted in a lower-than-expected retention performance two weeks after training. In another investigation on assembly tasks conducted in an artificially created environment, Vélaz et al. [10] studied the usefulness of VR as a teaching aid. Their evaluation relied solely on training time as a sole measure of XR efficiency assessment. Pan et al. [11] concentrated on a different aspect of VR, attempting to understand the impact of various forms of avatars on the learning and memory of the users/participants. However, the learning effect was inevitable through the sequence of experiments. Salah et al. [12] introduced a technique that utilized the predominant VR-based visualization method in product manufacturing and addressed the lack of innovation in training to keep pace with industry progress. The study was restricted to the task completion time as the primary evaluation criterion.

Next to the VR-based user studies above, several AR applications have also been developed and tested for training, maintenance, and monitoring purposes. Lai et al. [13] integrated multi-modal AR instructions with a deep learning framework for tool detection, improving assembly quality and efficiency. The system provided on-site instructions using visual renderings generated by a fine-tuned neural network. This integrated system enhanced the operator's performance by addressing assembly requirements. Although their research proposed a novel approach for integrating Deep Learning and AR, it did not provide detailed suggestions for generalizing the method in a commercialized setup. Moreover, Westerfield et al. [14] aimed at an AR with intelligent tutoring system (ITS), to enhance the training of manual assembly processes. Compared to conventional AR training, their strategy, which comprised a modular software architecture and prototype, increased task performance by 30% and test scores by 25%. It was concluded that an intelligent AR instructor considerably enhances learning. The study employed a small participant pool, necessitating repetitions with a larger student population in future work. Some participants had limited experience with motherboard assembly, potentially impacting the results. Several other studies have explored the application of AR in assembly processes. For instance, Chu et al. [15] examined AR user interface design for human-robot collaborative assembly. Visual and haptic cues were tested to convey robot intent, which reduced visual focus on the robot, with proximity cues being the most effective. The results had restrictions in assessing the influence of human trust on robot performance, as their correlation was weak in the experiment. Additionally, users found that recognizing complex information through haptic system vibration on different parts of the human hand is non-intuitive. Chu et al. [16] in another study developed AR functions to assist manual assembly in visually occluded scenarios. They evaluated the effectiveness of various AR assistive information and found that displaying operator hand movement in AR can reduce assembly time. Limitations encompassed the AR hand model's precision and a focus on occluded components. Further validating the effectiveness of additional AR information, like contextual clues, may be necessary. Mura et al. [17] addressed the challenge of achieving precise alignment between car panels in automotive assembly. They proposed an AR prototype system to guide operators during panel fitting operations, specifically for correcting alignment errors in terms of gap and flushness. A real case study was presented, focusing on the fine alignment of car body panels with the front light projector, to demonstrate the effectiveness of the AR system. Their study highlighted the potential of AR as a highly promising tool for supporting personnel in production processes. However, the assembly model developed was not evaluated through a user study, which is a critical step for XR implementations.

Webel et al. [18] developed a multimodal Augmented Reality-based framework for teaching maintenance and assembly skills, including subskill training and system assessment. In their evaluation, they determined that although the training period using the AR model was longer than the standard approach, the technicians' performance in a real-world maintenance scenario was quicker; after testing the system with two users and assessing their performance. The study employed a small sample size. Schwald et al. [19] proposed an augmented reality (AR) system for training and assisting in industrial equipment maintenance. The system comprised crucial hardware components such as an optical see-through Head Mounted Display, a tracking system, and a stand for installation. The research explored using an infrared optical tracking technology and associated calibration methods for ideal virtual overlay outcomes, as well as a scenario-based approach for guiding users through training or maintenance activities. Limitations included the need to reduce equipment weight, improve tracking system accuracy and latency, and enhance information access and internal communication for improved usability. Werrlich et al. [20] investigated the usefulness of AR-enabled head-mounted displays (HMDs) for training transfer tasks for assembly activities. They demonstrated that the AR approach outperformed conventional training approaches, such as video training and learning through printed text guidelines, in terms of immediate and long-term memory after training. Heinz et al. [21] created an AR system to aid users in comprehending complicated automated systems, despite their rising complexity and diminishing availability of experts. The study solely developed an AR platform and did not present a user study.

Within the domain of industrial applications leveraging MR, a select number of scholarly endeavors have been undertaken. Hoover et al. [22] utilized the Microsoft HoloLens HMD to instruct employees on how to assemble a wing. The study indicated that apart from assembly duties, MR may also be a valuable aid in maintenance chores. Some users noted issues with the HoloLens device's comfort, persistent graphics interference, and tracking errors as the limitations of the study. Muñoz et al. [23] introduced a novel MR-based user interface for quality control inspection, with the aim of improving ergonomics, reducing work-related stress, and enhancing productivity. They presented an experimental prototype and compared it with existing factory interfaces, like those at Mercedes-Benz, through usability tests. The study did not utilize a standard questionnaire like NASA Task Load to assess the effect of workload on each user. Baroroh et al. [24] proposed the use of MR simulation for the design of production systems with manual operations. They presented a MR-based planning tool that enabled collaboration between a manager and an operator wearing HoloLens. The planning results demonstrated the superiority of the MR tool over traditional simulation software in terms of planning quality and flexibility. The work validated the feasibility of MR as an effective approach for designing human-centric production systems. Murauer et al. [25] created language-independent attribute selection instructions using MR. The inconvenience induced by the HoloLens was highlighted as a system shortcoming.

In terms of training considerations, Sirakaya et al. [26] used AR to boost efficacy and enthusiasm in vocational education and training. From a strategic point of view, according to the critical study by Kaplan et al. [27] on multiple XR platforms, one of the principal advantages of immersive tools in training is that they can be incorporated in environments/applications that are *extremely difficult, time-consuming, or dangerous* for traditional methods of teaching. For instance, Farra et al. [28] devised a new reality simulation for disaster training to bolster the preceding point. XR has also proven advantageous in other application areas such as safety, education and military training, rehabilitation, and medical training. To exemplify, Pedram et al. [29] employed XR to assist mine rescuers in safety. Gonzalez et al. [30] outlined the procedures involved in developing an immersive simulation to hone naval abilities.

1.2. Physics-based process simulations within XR

Jalilvand et al. [31] conducted a case study on MR effectiveness for training thermoforming process operators, with a particular focus on User Interface (UI), User Experience (UX), and usability. They introduced a spatial UI for interacting with virtual objects and incorporated enhanced UX via real-time heat transfer simulations. The usability of their approach was validated only using Microsoft HoloLens 2. However, the study was limited to a small group of eight users. The present study herein builds on the above earlier work. Hernández et al. [32] integrated classical simulations with real-time physics engines using deep learning, resulting in a real-time AR simulator for interacting with virtual deformable solids. They reduced online computation time for XR HMDs, but faced challenges in scaling for larger meshes and required extensive offline training for their deep learning model. In a different study, Badías et al. [33] used a model order reduction to create a real-time AR interaction method, blending machine learning, computer vision, and computer graphics. They addressed occlusion with pixel-depth analysis and GLSL shaders; yet limitations in the depth camera hindered distant object detection.

Weistroffer et al. [34] developed the SEEROB framework for assessing safety and ergonomics in robotic workstations. It allowed direct communication with robot controllers, employed a precise physics engine for evaluations and workstation adjustments, used XR for user comprehension, and included motion capture for ergonomics assessments within the DT. Limitations included the need for hard-to-obtain DT parameters, custom adaptations for real-time robot data, and a lack of physical workstation feedback control. Li et al. [35] introduced an integrated AR framework that combined real-time multi-material simulation with efficient hand gesture interaction. They used an RGBD camera for 3D data acquisition, enabling real-time gestures in simulations and expanding AR applications. However, the user study was restricted to the PC use due to its reliance on the RGBD camera and Leap Motion for gesture recognition and scene acquisition. Koduri et al. [36] introduced AUREATE, an Augmented Reality Test Environment that enhanced automated driving testing by augmenting real-world sensor data with simulated data. It was effective for object detection algorithm validation. However, AUREATE faced limitations due to hardware constraints and resource-intensive LiDAR augmentation. Achieving sensor data realism for perception algorithm training presented challenges that required complex dynamics models beyond basic physics.

1.3. Summary of current gaps and objective of the present study

In the domain of VR, AR, and MR training systems for industrial applications, most of the above reviewed studies exhibited a few common limitations. Notably, most of the user studies suffered from small sample sizes, diminishing the extent to which the findings may be generalized. Additionally, certain studies limited participants selection to specific user groups, thereby potentially restricting the broader applicability of the training system for others. Furthermore, a prevalent limitation was deemed to be the adoption of a unidimensional (single criteria) evaluation approach, primarily centered on task completion time. Notably, user studies were omitted at times, consequently impeding the ability to evaluate practical usability of the developed systems. Moreover, recurrently unaddressed are technical concerns linked to system setup, equipment weight, tracking precision, and user comfort, which can potentially impact both the user experience and the interpretability of outcomes.

Finally, to the best of our knowledge, no study directly compared different XR training platforms within the context of workforce training in a same manufacturing operation. Typically, the studies have focused on singular XR platforms for adaptation and evaluation. Among immersive platforms, the utilization of AR applications for industrial training purposes may be deemed less interactive, given that AR applications primarily serve to augment virtual objects within the user's field of vision, with limited user interaction with the physical world. In contrast, MR systems emerge as a more viable platform facilitating practical interactions between users and augmented virtual objects. In this context, Iop et al. [37], demonstrated the interplay among different immersive XR platforms, explaining their potential use cases and areas of convergence.

In addition to the immersive platforms, it is deemed essential to consider and compare a traditional base training approach (e.g., 2D videos) as commonly employed in manufacturing settings. Finally, most of the established XR applications for training purposes, have generally followed sequential steps without involving physics-based process simulations (e.g., finite element models), which are often the backbone of design and decision-making in advanced manufacturing engineering.

In view of the above gaps, the main objective of this case study is to develop, implement, and compare two immersive training tools in a same complex manufacturing process. Namely, we employ VR as well as MR 3D platforms, in addition to a conventional 2D (video) training as the base method, and compare their performances from users' perspective in a thermoforming case study. To enhance the study's generalizability, a relatively large user study (comprising 26 participants) was conducted. The comparative analysis encompassed a diverse array of evaluation criteria, including usability, user-friendliness, trust, reliability, and training effectiveness, as well as the task completion time. Next to comparing these immersive training platforms, the novelty of the present research may lie in the integration of a numerical heat transfer simulation model of the process into the developed VR and MR platforms, as well as formal Multi-Criteria Decision Making (MCDM) techniques to rank the XR training method alternatives.

2. Methodology

As a case study, a thermoforming manufacturing process involving multiple human operations is considered, where heat transfer phenomena have a significant effect in the quality of the final formed product. As illustrated in Figure 1, the first stage of the proposed methodology is creating a DT of the process, where the heat transfer phenomenon between different components of the manufacturing system can be accounted for. Following the modelling of the heat transfer simulation tool (here in C#), and the development of its visual components using 3D model software, object behaviours are assigned using Unity to the components inside the XR (here VR and MR) applications. In the next steps, the user input methods and profiles for each VR/AR environment are established, based on the respective devices that are utilized (here Meta Quest 2 (VR headset) and Microsoft HoloLens 2 (MR headset)). Interactive elements of the programmes are adapted to the user's input system, such as joystick controls for Quest 2 and hand gestures for HoloLens 2. Finally, a user survey is conducted to gather data on how well the developed VR and MR applications would fit specific performance criteria. The obtained user study data is analyzed statistically and used in a MCDM model to rank and determine which platform is most suitable for training the operators.

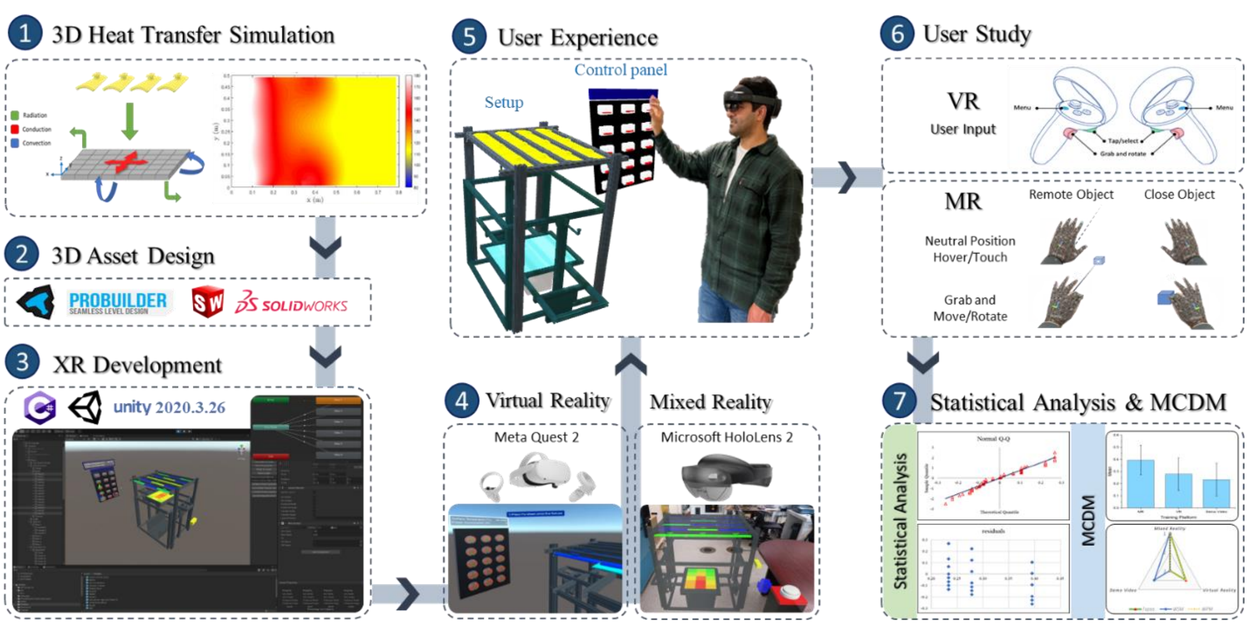

Figure 1 – Overview of the methods employed in the present study.

2.1. Thermoforming set-up

In a typical thermoforming process, thermoplastic sheet is heated up and placed into a mould by vacuum pressure. The aim is to heat the sheet to reach the desired temperature distribution before it forms. During this heating stage, maintaining the temperature at a certain level in each region of the sheet is crucial (to ensure the subsequent mechanical forming effectiveness and yield a defect-free final product with a uniform thickness). Figure 2 illustrates a typical procedure followed during the thermoforming process and the resulting sample regions with a typical thickness variation deficiency [38]. Unfortunately, the current industrial thermoforming methods have not all achieved the required efficiency to avoid such manufacturing defects through automation of the process. As a result, some of the final products are being discarded (the production failure rate has been seen to be high as 30% if the operator does not carefully optimize the process).

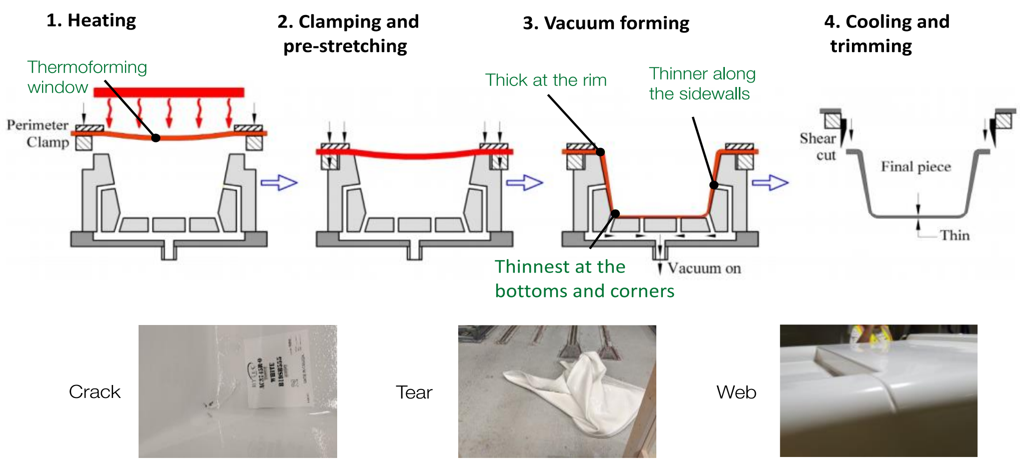

Figure 2 – A typical thermoforming process and example of potential defects in the final part [38].

2.2. Transient (time dependent) heat transfer modeling of the process

The numerical model consisted of fifteen heating elements above the thermoplastic sheet. It incorporates all the three conductive, radiative, and convective heat transfer modes, as illustrated in Figure 3. Figure 4 illustrates the thermoplastic sheet's constituent parts. Given the small sheet thickness in relation to its length and width, only one layer of elements is considered. The four sides (±X, ±Y) of each element are considered for heat conduction between elements. Radiation heat transmission from the heating coils occurs through the sheet's top face (+Z). Every surface of the element in contact with its neighbors and the environment is subject to heat conduction. q_{z+} and q_{z-} represent heat transmission in the vertical direction, while (q_{x+},q_{x-}) and (q_{y+},q_{y-}) represent heat transfer in the longitudinal and transverse directions, respectively. Equation (1) represents the heat transfer rate by conduction in a given direction [39].

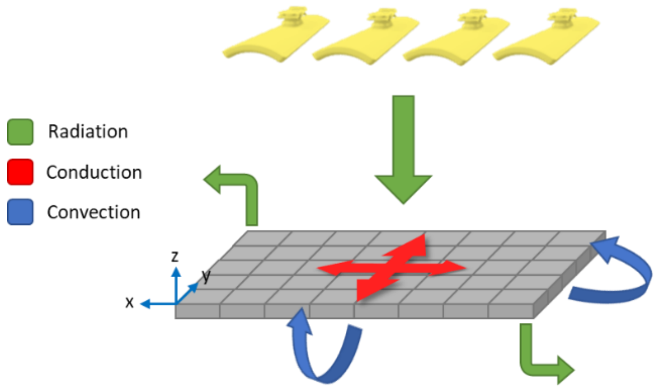

Figure 3 - Heat transfer between thermoplastic sheets, heaters, and the environment.

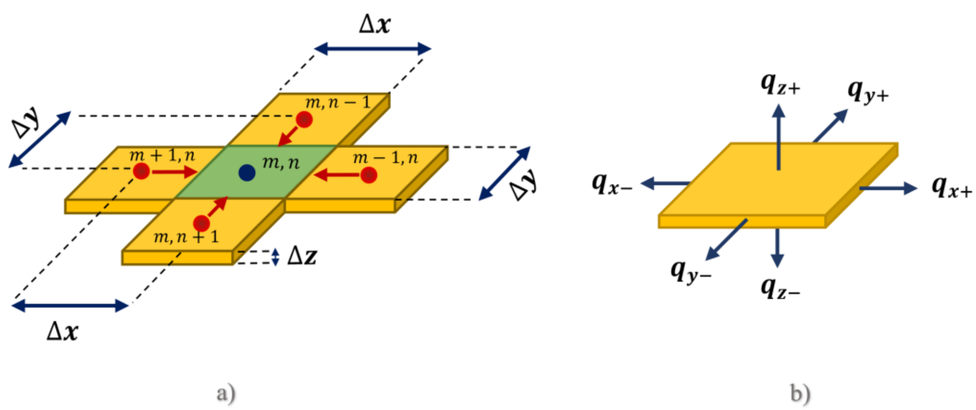

Figure 4 – Finite difference element modelling: (a) conduction heat transfer, and (b) including all modes of heat transfer.

$$Q_i^{cond.} = -kA\frac{\Delta T}{L} \tag{1}$$

Where $k\left[\frac{W}{m.K}\right]$ is the material's conductivity, $A\ [m^2]$ is the cross-sectional surface area, $\Delta T\ [K]$ is the temperature difference between the ends, and $L\ [m]$ is the distance between the ends. To calculate the conduction heat transfer, only $\pm x$, $\pm y$ directions are considered. Therefore, the corresponding in-plane conduction heat transfer equations are depicted as:

$$Q_x^{cond.} = k\Delta y\Delta z\frac{T_{m+1,n} - 2T_{m,n} + T_{m-1,n}}{\Delta x} \qquad , in \pm x\ direction \tag{2}$$

$$Q_y^{cond.} = k\Delta x\Delta z\frac{T_{m,n+1} - 2T_{m,n} + T_{m,n-1}}{\Delta y} \qquad , in \pm y\ direction \tag{3}$$

At the air-contacting surfaces of the element, convective heat transfer occurs between the element and its surroundings. This includes the $\pm z$ sides of the thermoplastic sheet. Equation (4) demonstrates the rate of heat convection from the i_{th} node to the surrounding environment [39].

$$Q_i^{conv.} = hA_i(T_i - T_{ambient}) \tag{4}$$

Where $h\left[\frac{W}{m^2.K}\right]$ is the convective heat transfer coefficient, A_i $[m^2]$ is the surface area of i_{th} node where the heat transfer takes place, T_i $[K]$ is the temperature of the i_{th} node of sheet and $T_{ambient}$ $[K]$ is the temperature of the ambient. When a cold environment is formed around a hot, horizontal sheet, free convection occurs. However, the coefficient of heat transfer will be variable between the plate's top and bottom surfaces as h_1 and h_2.

Equations (5) and (6) explain the free convection equation for elements positioned on various sides of the thermoplastic sheet.

$$Q_{z^+}^{conv.} = h_1 \Delta x \Delta y (T_{m,n} - T_{ambient}) \qquad , in + z\ direction \tag{5}$$

$$Q_{z^-}^{conv.} = h_2 \Delta x \Delta y (T_{m,n} - T_{ambient}) \qquad , in - z\ direction \tag{6}$$

Finally, as the radiation heaters are situated above the sheet, the input energy to elements can also enter via face +Z through radiation mode. Similarly, the heat may be lost from thermoplastic sheet by radiation from both $\pm Z$ sides to the environment. Equation (7) expresses the net rate of radiant heat transmission from the heaters to the i_{th} element on the plastic sheet [40].

$$Q_i^{rad.} = A_j \varepsilon_e \sigma \left[\sum_{j=1}^{M} F_{j,i}(T_j^4 - T_i^4) \right] \tag{7}$$

Where A_j $[m^2]$ is the surface area of emitter, ε_e is the effective emissivity, σ is the Stefan-Boltzmann constant: $5.67 \times 10^{-8} \left[\frac{W}{m^2K^4}\right]$, $F_{i,j}$ is the view factor between the j_{th} heater and the i_{th} node, T_j $[K]$ is the absolute temperature of the j_{th} emitter, T_i $[K]$ is the absolute temperature of the i_{th} node on the sheet and M is the total number of the heaters.

The net heat flow, which is equal to the change in internal energy, may be computed using the heat flow in all directions outward from the element, given in Equation (8). The temperature change may be calculated from the corresponding change in internal energy, where q_i represents the heat transfer from all surfaces to each element on the thermoplastic sheet. By combining the temperature shift caused by the change in internal energy and the temperature from the preceding time step, the new temperature (T_t) is obtained in equation (9). Where t is the time step, ρ is the sheet density, V is the element volume and C is the sheet material's heat capacity [41].

$$dU = \sum_{i=1}^{6} q_i = Q_i^{rad} + Q_i^{conv} + Q_i^{cond} \tag{8}$$

$$T_t = T_{t-\Delta t} + \left(\frac{\Delta t}{\rho V C}\right) dU \tag{9}$$

The view factor is the ratio of energy emitted by the first surface to that absorbed by the second surface. Mathematical calculations for view factors are known for the relative locations of the surfaces [42]. Based on these equations for view factor, models of the heating phase of the thermoforming process have been developed. In this instance, it is assumed that the heating source is a flat, temperature-maintaining plate.

Equation (10) depicts the time-dependent temperature variations of the sheet's elements due to radiation heat transfer using a single heater. The view factor between the components of the sheet and the heater may be calculated at the outset of the heating process using Equation (11) (given the thermoplastic's low thermal conductivity, which is due to its molecular structure and relatively weak intermolecular bonding).

$$\frac{dT_i}{dt} = \left(\frac{1}{\rho V C}\right) A_h \varepsilon_e \sigma F_{h,i} (T_h^4 - T_i^4) \tag{10}$$

$$F_{h,i} = \frac{\rho V C}{A_h \varepsilon_e \sigma (T_h^4 - T_i^4)} \times \frac{dT_i}{dt} \tag{11}$$

Following the above equations, the temperature variation can be obtained on each element of the thermoplastic sheet based on the applied power on each heater. Once combined for fifteen heaters, it is possible to simulate the whole process in real-time and measure the temperature distribution on the whole surface of the sheet. It is worth noting that the validation of this finite difference transient heat transfer simulation was performed through an established Finite Element (FE) model developed in [43].

2.3. XR Development

To create interactive DTs of the process (i.e. VR/AR tools integrated with the heat transfer simulations; Figure 1) for the operator training applications, 3D components of the system were designed in Unity® ProBuilder and SolidWorks®. Once all the components were designed, their STL files were converted into OBJ formats to be recognized as prefabs inside the Unity environment. These were used for object programming in C# language through Unity Engine. In the ongoing work of the authors, an actual thermoforming setup is also being designed and assembled in the lab environment. A demo video was recorded using this physical set-up and used as a baseline training method for the operators. The video training will be compared with the XR application tools developed in Section 3. The XR design and the actual lab-scale setup are demonstrated in Figure 5.

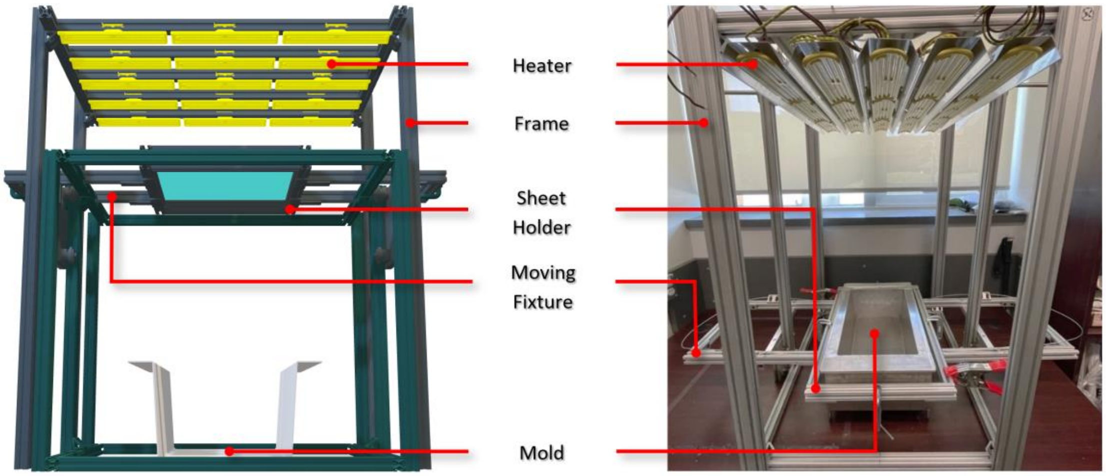

Figure 5 - XR design (left) versus the actual setup (right).

2.3.1. VR application

The VR application incorporates two distinct scenes: the "**Setup Outline**" and the "**Tutorial**." The "Setup Outline" serves as a central repository, where all components of the DT within the XR design are consolidated, as depicted in Figure 5. Its primary function is to provide users with a comprehensive perspective of the entire process, allowing them to acquaint themselves with the setup components prior to undergoing formal training. Conversely, the "Tutorial" scene is exclusively dedicated to instructional purposes. It encompasses all training modules designed for amateur users who require practice with the DT before engaging with the actual setup. The "Tutorial" scene primarily functions as the training environment for users to acquire proficiency in the thermoforming process.

Essentially, tutorial scene uses the Unity animation state machine in which each step is triggered once a specific set of actions (such as pressing a button or displacing an object) are completed, as demonstrated in Figure 6. This was done by incrementing the step numbers in the animator interface as soon as the condition for the completion of the previous tasks is met. These two scenes are interchangeable, and the application restarts as soon as switched.

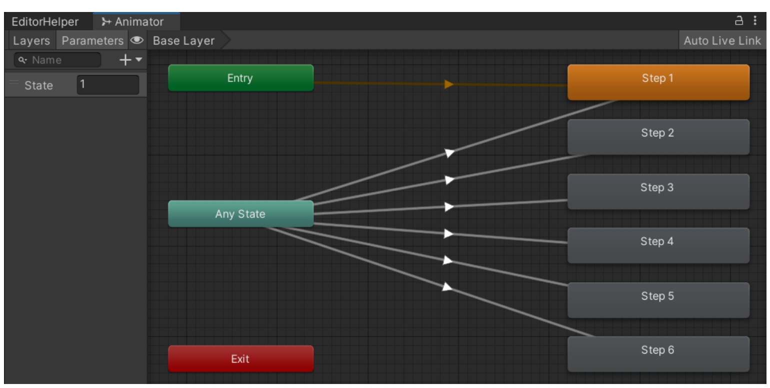

Figure 6 – The Unity state machine and animator employed.

Meta Quest 2 was selected as the target device for developing the VR application. To define the input system, three buttons were assigned to the controllers to allow the user to interact with the DT. As illustrated in Figure 7, the "Axis1D.PrimaryIndexTrigger" buttons were defined as the "Tap/Select" option and "Axis1D.PrimaryHandTrigger" buttons were defined as the "Grab and rotate" option. Moreover, the "Start" buttons were selected as the "Menu" option.

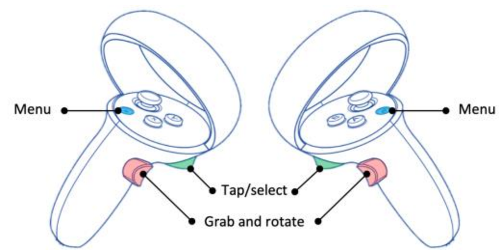

Figure 7 - VR input system definition.

2.3.2. MR application

The MR application concept development was very similar to what was implemented in the VR application (Section 2.3.1). However, the target device, user input system definition, profiles, and some of the interactive component designs were different. Namely, the training DT in MR case was developed using Microsoft's HoloLens 2, based on the input system provided by the Mixed Reality Toolkit (MRTK), to allow the user to interact with the environment. Similarly, the User Interface (UI) consisted of two scenes, the Setup outline and the Tutorial. The Setup outline scene includes all the components of the DT without any assistance presented to the operator. In contrast, in the tutorial scene, the user should follow six steps based on text instructions and visual guides, which are provided at each step.

To define the input system in the MR application, MRTK profiles were utilized to define various controls for the operator. Since HoloLens 2 does not come with separate controllers similar to Quest 2, hand gestures are specifically designed to define various controls. As illustrated in Figure 8, two modes of input systems were defined for close and remote interactions. In Figure 8a neutral position, hovering, and touching an object is demonstrated, where the palm is completely open, and for the remote selection, the ray cast was used, which is shown as a dashed line. Figure 8b exhibits grabbing, moving, and rotating the objects by pinching (closing the thumb and index fingers). The activated remote ray cast is shown as a solid line.

Figure 8 – The MR input system and hand gestures for near and remote interactions. a) Neutral position, Hover, and touch. b) Grab, move, and rotate.

2.4. User study design

One of the main features that was implemented in the XR applications above is the heat transfer model to represent physical behaviour among heaters, thermoplastic sheet, and the environment (Section 2.2). This was carried out in C# scripts and made the XR applications more intuitive and interactive. Overall, three modes of training scenarios were defined for the user study:

1. Virtual Reality Training (VRT)
2. Mixed Reality Training (MRT)
3. Demo Video Training (DVT) (as the base training method)

The objective of the first two training methods in practice would be to offer an immersive training tool for the novice operator to learn the essential steps of the thermoforming process considerably more quickly than the conventional video training method. Additionally, the operator can be provided with some valuable insights and real-time visual information about the physical status of the process (here the temperature distribution on the heaters and sheet).

The immersive training environment designed for each user was based on six sequential tasks:

1) Placing the thermoplastic sheet onto the fixture.
2) Turning on a specific set of heaters.
3) Adjusting each heater's power using the button/slider.
4) Checking the surface temperature on various heaters; once all the heaters reach their desired surface temperature, turning off all the heaters.
5) Releasing the sheet frame and moving it down onto the mould.
6) Turning on the vacuum pump.

All the above steps, except Step 4, incorporate human direct actions that can be broken down into moving, grabbing, hovering and pushing/pulling objects, which are all specified based on the colliding boundaries and physics-based relations defined for the target objects. Step 4, fueled by the numerical simulation model in real-time, requires a pre-condition to be satisfied, which is to check the sheet surface temperature on all the heaters by selecting them separately. Each VR and MR environment was designed based on their input systems to allow the operator to choose each heater as they touch/hover over them. The screen information on the control panel is updated instantly based on the governing heat transfer laws.

A video demonstration of the developed application can be found at the **Supplementary Material**.

2.4.1. Hypotheses

The overarching hypothesis considered in the user study was: *MR would be the most suitable platform for training purposes, compared to VR and video training.* To better outline the features of this hypothesis, it was divided into sub-hypotheses as follows:

Sub-H1: Interacting with the designed DT in MR environment would be more beneficial for training purposes compared to learning via a VR environment and training video.

Sub-H2: Operators would feel more *convenient* being trained by an MR environment compared to VR and the training video.

Sub-H3: Among all the provided training options, operators are *satisfied* with their performance during and their training via MR environment.

Sub-H4: MR would be the most *trustable* environment among the three platforms.

Sub-H5: MR would be the most *reliable* source for the trainees to function properly compared to the VR environment and training video.

2.4.2. Experimental procedure and questionnaires

Prior to the tests commencement, participants were assigned to workstations with access to either the DVT device (laptop) or MRT device (HoloLens 2) or the VRT device (Quest 2), following a Latin square sequence. This approach determined the order of experimental groups and mitigate the impact of the learning effect during the evaluation of the training methods.

Moreover, participants signed a permission form, answered a few demographic questions, and described their VR/AR experience before beginning the experiment. Participants were organized into three distinct groups, each assigned to a different session of the user study, as illustrated in Figure 9. In a systematic fashion, each group initiated their experimental procedure with a distinct training method, following the predetermined Latin square sequence. Specifically, Group One's procedure sequence was as follows: 1. DVT, 2. VRT, and 3. MRT. For Group Two, it proceeded as: 1. VRT, 2. MRT, and 3. DVT. Group Three adhered to: 1. MRT, 2. DVT, and 3. VRT as their sequence.

In the VRT phase of the study, the Meta Quest 2 was utilized as the VR device, offering users a fully immersive experience in a virtual environment. Participants were asked to follow all six steps (tasks) mentioned in Section 2.4. For safety reasons, the application marked a designated working area, enabling users to concentrate exclusively on their training while alleviating any concerns about potential collisions with physical objects in their surroundings.

For the DVT segment of the study, participants were introduced to the physical lab-scale test setup and the steps to finish the process. Additionally, they were provided with a concise demonstration video showcasing the experimental tasks on the actual setup. Participants followed the video instructions (similar to the other two training modes) to navigate through the thermoforming process.

During the MRT section, participants put on the HoloLens 2 headset, initiating the training experiment. They utilized both visual cues and audio instructions provided at each stage. As the HoloLens 2 is a MR device, users had the ability to maintain awareness of their physical surroundings through the goggles, while virtual interactive DT components augmented as they progressed through each step of the experiment.

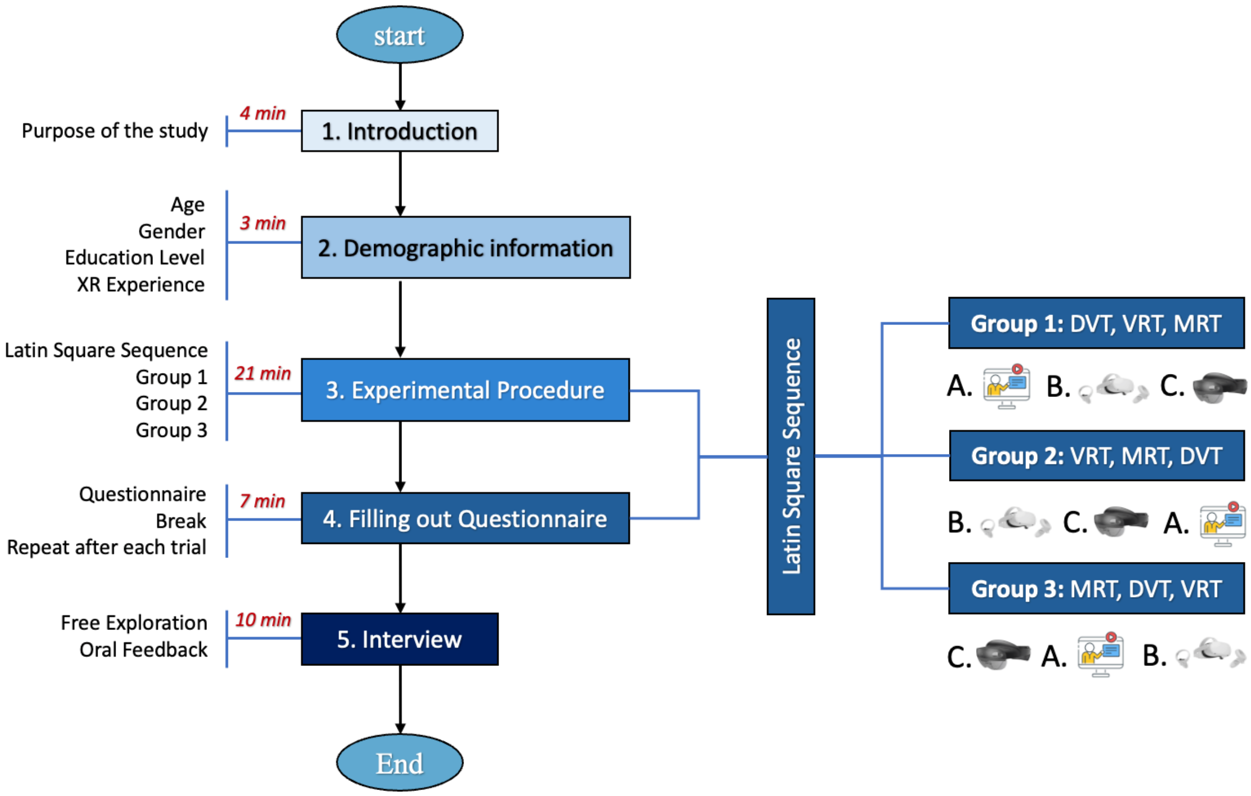

Figure 9 - Flowchart of the experimental procedure in the user study.

The user study employed a “within-subject” design, whereby the same group of participants experienced all conditions. In other words, each user was exposed to every training method. This approach enabled a comprehensive assessment of how individual responses varied across different training methods. After finishing all the three trials, they assessed their experience and offered qualitative feedback about the effectiveness of the immersive training methods in comparison to conventional video training and the general design of the immersive applications. To avoid tiredness and make the whole process consistent, all three sections of the user study were timely and kept short. On average, each participant required forty-five minutes to complete the three tests.

2.4.3. User survey

The authors previously conducted a study on thermoforming operator training exclusively using the MR system, while examining User Interface (UI), User Experience (UX), and usability of Microsoft HoloLens 2 [31]. The latter study involved eight participants and utilized the System

Usability Scale (SUS) questionnaire [44], resulting in an approximate usability score of 85/100. Building upon those findings, the current work expanded the user study to encompass a larger sample size (26 users) along with comparing both a VR and a modified MR application, in addition to the traditional video-based training. Also to address the SUS survey limitations in the earlier study [31], a revised questionnaire was devised based on users feedback from the earlier study. Namely, the new questionnaire was tailored to specifically assess criteria related to XR design and operator training, areas not traditionally addressed by SUS or other conventional UX questionnaires. Unlike the SUS questionnaire, which primarily measures usability and ease of use, the new survey evaluated the trust, reliability, user satisfaction, training effectiveness, as well as usability and ease of use, and lastly overall preference, as follows.

- ***Trust**:* Each training technique's trust level was evaluated by participants using a 5-point Likert scale (1 – lowest score, 5 – highest score). Additionally, an open-ended question invited participants to provide insights on how to enhance user trust in each training mode.

Q 1. I trust the system completely.
Q 2. The system is completely predictable.
Q 3. The system is completely safe. (Feels safe while working with)

- ***Reliability**:* On a 5-point Likert scale (1 – Strongly disagree, 5 – Strongly agree), participants were asked to evaluate their reliance on each of the training techniques. Moreover, one open-ended question was posed to gather user's comments on potential improvements.

Q 4. The system provides the type of help I require to make my decision.
Q 5. The system performs reliably.
Q 6. I can rely on the system to function properly (While using).

- ***User satisfaction:*** Participants were asked to rate their satisfaction with the training platforms on a 5-point Likert scale (1 – lowest score, 5 – highest score), mirroring the format of previous criteria questions. Also, an open-ended question requested ways to improve user satisfaction in each training mode.

Q 7. I feel comfortable while/after using the system.
Q 8. I would recommend others to use the system.
Q 9. Feels natural to use the system.

- ***Training effectiveness:*** Similarly, another short survey comprising three questions asked the user to assess the training efficacy based on their experience with the platforms. Questions were designed on a 5-point scale (1 – Strongly disagree, 5 – Strongly agree).

Q 10. I understand high-level details of the training with the system.
Q 11. The system is suitable for complex training purposes.
Q 12. The system is practical to follow step by step.

- ***Usability and ease of use:*** Participants were queried to assess the usability and ease of use for each training method with responses on a scale of 1 to 5 (1 - strongly disagree, 5 - strongly agree). A few questions focused on the learnability, followed by a similar approach that was taken by Lewis et al. [45], suggesting that factor analysis of two independent SUS data sets included two inherent factors: "Usable" and "Learnable". Additionally, participants offered qualitative input through two open-ended questions in a post-experiment questionnaire, suggesting potential improvements for each training mode in terms of usability and ease of use.

Q 13. The system would help me to be more productive.
Q 14. The system would be useful over learning manufacturing-related topics.
Q 15. The system would give me more control for learning manufacturing-related topics.
Q 16. The system is easy to use.
Q 17. The system is user friendly.
Q 18. The system is simple to use.

- ***Preference:*** This part of survey consisted of two questions, one with a maximum of five responses (from "Do not prefer" to "Prefer very much"), to evaluate the overall preference of the user over the training methods, and one open-ended question for user's comments.

Q 19. My level of preference towards using the system.

In order to statistically evaluate the reliability of the designed questionnaire, Cronbach's alpha [46] was used, which is a widely known measure for assessing the internal consistency of a given questionnaire or survey. In other words, it is a tool that helps to determine whether the survey questions are measuring the same underlying concept. The possible values for Cronbach's alpha range from 0 to 1, where a score of 0 implies no consistency, whereas a score of 1 indicates perfect consistency. Typically, a threshold value of 0.7 or higher is considered acceptable for scientific studies [36].

In addition to the users' responses to the designed questionnaire (subjective feedback), the completion time for each of the training methods was also employed as an objective metric to assess the training effectiveness of DVT, VRT, and MRT.

2.4.4. Participants diversity

26 participants (13 males and 13 females) were recruited from the local university community, as shown in Figure 10a. The educational levels of the participants ranged from high school diploma to doctorate, details are shown in Figure 10b (53.85% had a master's degree, 19.23% had a bachelor's degree, 15.38% had a doctorate, 7.69% had a high school diploma, and 3.85% held other credentials). Half of the participants had prior experience with immersive applications in VR and MR environments, while the other half had no prior experience (Figure 10c). This was primarily due to the recruitment of a large group of participants for a more comprehensive assessment of the training methods. Additionally, limited access to experienced XR users resulted

in half of the participants being novices. Therefore, the presence or absence of XR technology experience was treated as an independent variable in the user study, enabling the analysis of differences between XR novices and experts for all the criteria introduced in the survey.

The distribution of all participants with prior (relative) XR experience could be classified into three groups: "little experience" with two participants (less than one month), "moderate experience" with four participants (one to six months), and "extensive experience" with seven participants (six months and more). The majority of experiences were gained via gaming (13 participants). Moreover, participants' ages ranged from 18 to 36, with a mean of 27.54 and a standard deviation of 4.23, as indicated in Figure 11, with normal distribution. Lastly, everyone in the study had normal or corrected-to-normal eyesight.

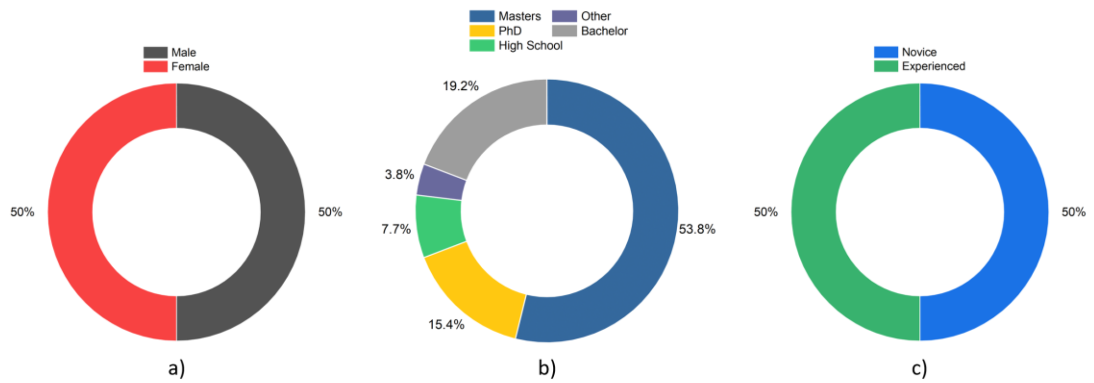

Figure 10 - Participants' demographic information. a) gender, b) educational level and c) experience level.

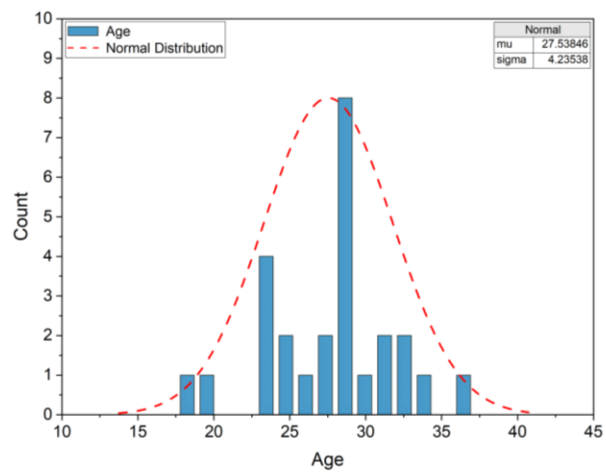

Figure 11 - Participants' age distribution.

2.5. Statistical analysis with blocks

Although there are several methods for testing user study hypotheses, the analysis of variance (ANOVA) is perhaps the most widely used method in similar experimentations [47], [48], [49]. ANOVA is essentially used to determine whether there are significant differences between the means of three or more distinct (unrelated) study groups (often with a significance level of 5%;

[25], [26]). When implementing ANOVA, data blocks can also be employed to control potential sources of data variation other than the treatments or primary independent variables themselves. In other words, blocks are used to eliminate known causes of variance in the interpretation of final results. In this research, because the participants were completing the same questionnaire against the three training methods (VRT, MRT, DVT) and their individual perspective was common between the questionnaires, each user was considered as a block during ANOVA. Post-ANOVA analysis methods such as Fisher LSD, Tukey, and Bonferroni can also be employed to determine the statistically significant difference among pairs of the study groups, and in this study Bonferroni correction is employed. To utilise the standard ANOVA for experimental data analysis, it is necessary to verify (meet) the following assumptions [49]:

- The populations from which the samples are taken, have a normal distribution.
- The samples/data points are independent.
- Data populations have the same variance; i.e., the error distribution variances are homogeneous across populations.

According to the study by Mircioiu et al. [50], comparing parametric and non-parametric methods for analyzing Likert scale data, both methods consistently yielded identical results in inter-subgroup comparisons, regardless of whether they indicated significance or non-significance. This consistency held true when working with "large" (>15) response datasets and similar, though not necessarily normal, distributions among various subgroups. Given the large number of participants (26 persons) in the current study, ANOVA method was used. Regardless, the underlying assumptions of ANOVA were also checked (more details to follow in Section 3.4.1).

To analyze the results obtained from the questionnaire, the Likert scales were quantified by assigning values ranging from 1 (strongly disagree) to 5 (strongly agree), denoted as $i = \{1,2,3,4,5\}$, with i representing the scales index. Subsequently, the scale was normalized to a range between 0 and 1, represented as weights (W_i), as demonstrated in Eq. (12). This transformation merely ensured a normalized representation of the Likert scales for subsequent analyses.

$$W_i = \frac{x_i}{\sum_i x_i} \tag{12}$$

Given that the survey comprised multiple questions denoted by $j = \{1,2,\dots,N\}$, where N represents the maximum number of questions per criterion (see also Section 2.4.3), the weighted response (Z_{ijk}) for each question was computed via:

$$W_i \times Y_{jk} = Z_{ijk} \tag{13}$$

In this context, Y_{jk}, which has a binary value of {0 or 1}; 1 for the selected response and 0 for the rest of the weights, signifies the responses provided by a user for each question (j) and each training method ($k = \{1,2,3\}$). Subsequently, the normalized weighted response (S_{jk}) was derived following equation, with C_j being defined as $\sum_i \sum_k Z_{ijk}$.

$$S_{jk} = \frac{\sum_i Z_{ijk}}{C_j} \tag{14}$$

Finally, to compute the score of each training method relative to each criterion, the average of the user responses to all the questions was calculated as follows.

$$M_k = \frac{\sum_j S_{jk}}{N} \tag{15}$$

2.6. Multi-Criterial Decision Making (MCDM)

The study employed five common MCDM techniques for final ranking of the three alterative training methods under the multiple performance criteria listed in Section 2.4.2. These methods included TOPSIS, Weighted Sum Method (WSM), Weighted Product Method (WPM), Multi MOORA, and VIKOR, as briefly reviewed below.

The technique for order preference by similarity to the ideal solution (TOPSIS) was introduced in [51]. The TOPSIS MCDM technique ranks alternatives by emphasizing those that are "closest" to the positive ideal solution (PIS) and "furthest" from the Negative Ideal Solution (NIS) [52]. The NIS is used to maximize cost-type criteria and minimize benefit-type criteria, whereas the PIS is used to accomplish the reverse. Due to its theoretical precision, TOPSIS has been widely used in different MCDM applications [53].

In WSM, objectives are scaled (aggregated) into a single objective by multiplying each criterion by a decision maker-defined weight. Typically, the weight allocated to a criterion is proportionate to the criterion's importance relevance in the given application. The weights are normalized to guarantee that none of the criteria are over-scaled. On the other hand, the WPM is closely linked to WSM, with the key distinction being that WPM involves a product of criterial values, rather than a sum [54].

MULTIMOORA, similar to other popular MCDM methods, employs vector normalization to provide uniform ratings and uses three supplementary ranking techniques: The Ratio System, the Reference Point Approach, and the Full Multiplicative Form. Given that each of the three ranking approaches has advantages and disadvantages, MULTIMOORA employs a hybrid version of these strategies [55].

As another viable ranking technique for MCDM, VIKOR provides a compromise ranking strategy [56], [57]. The VIKOR technique was developed to handle discrete choice problems and to improve complicated systems with inconsistent and non-comparable characteristics on a multi-criteria basis [56], [57]. When the decision-maker cannot convey their choice to find compromise solutions, VIKOR becomes particularly a helpful method. The reached compromise solutions minimize individual regret while maximizing community benefit, which is why the decision-makers can accept them. This method focuses on ranking and selecting among possibilities when there are conflicting requirements. To aid decision-makers in minimising the trade-offs necessary to attain the best choice, ranking by VIKOR may be undertaken with varied values of criteria weights. This analysis examines the impact of criterion weights on the recommended compromise solution [56], [57]. According to [58], fuzzy VIKOR and VIKOR are currently among the most practical methods for ranking and selection problems in large-scale settings.

3. Results and Discussion

3.1. VR application

As described in Section 2.3, a practical six-step guidance was presented to each operator/user to follow and finish the manufacturing procedure via the developed VR application. Figure 12 illustrates a typical scene recorded from one of the user's tests on each of these steps. The first step requires the operator to lay the thermoplastic sheet on the moving fixture. To simplify the visualization and minimize confusion, the remaining components of the setup are not superimposed on the fixture in this stage (Figure 12a). Next in Figure 12b, the operator is instructed to complete the next three steps in a single scene in which all the setup components, including the main frame, heaters, adjustable frame, sheet, and mould, are presented. Step 2 requires the operator to switch on the heaters by pressing the button on the left side of the setup. During the third step of the procedure, the user may modify the power for each heater by choosing the desired heater and then rotating the nob. In the fourth step, after the heating procedure is finished, the operator may switch off all heaters simultaneously by pressing the prompted button on the right. After completing the preceding stages, the user should release the moving frame from the side hangers and lower it to the mould, which is indicated as step 5 in Figure 12c. After placing the sheet on top of the mould and securing the moving frame, it is time to begin the vacuum process by activating the vacuum pump. This is demonstrated in Figure 12d as the last step of the thermoforming process.

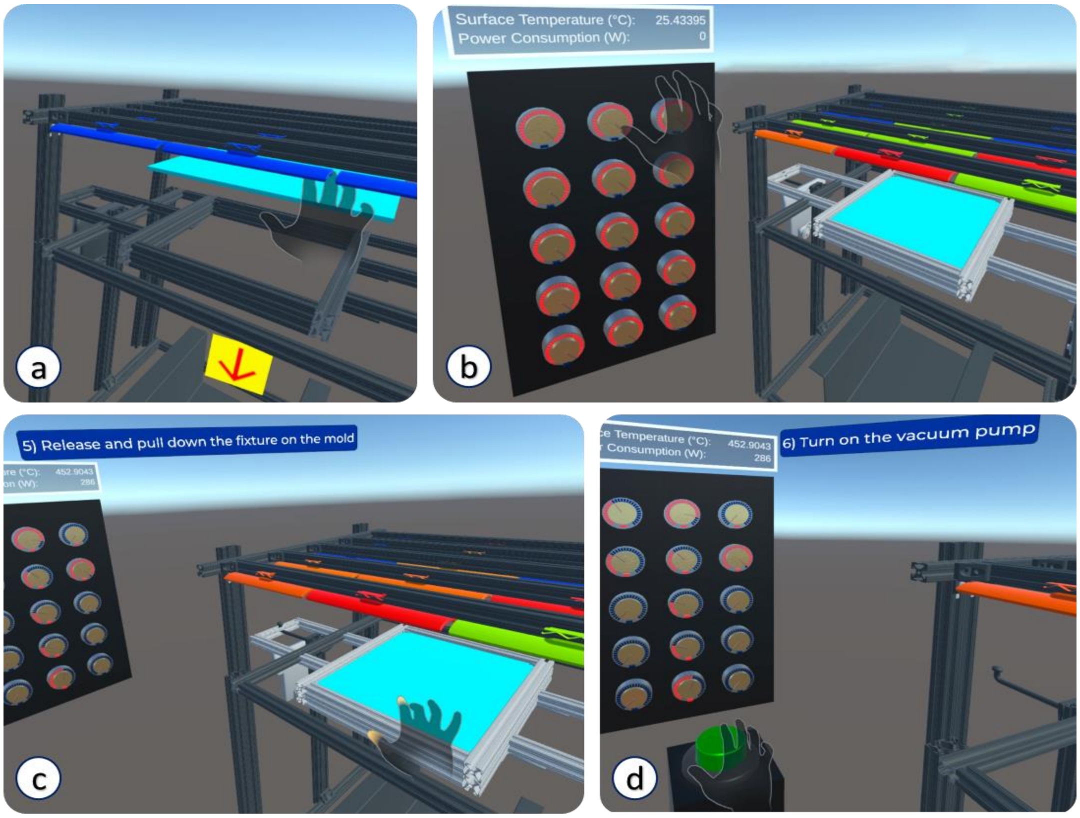

Figure 12 – Illustration of the VR app steps followed by a user in the experiments.

3.2. MR application

Figure 13 shows typical scenes recorded from one of the study user’s tests using MR application, to complete the same six tasks identified in Section 2.4. In the first step as depicted in Figure 13a, the user is asked to grab the thermoplastic sheet and position it on top of the moving fixture. Then, to assist the user in navigating between the panels if they are not at the proper height or if the items are far away, the operator can select and rearrange the objects. Only the relevant interactive elements are shown in each step to make the MRT more intuitive. In the second step on visualization, a control panel is added to the screen to activate the heating elements (Figure 13b). There are 15 buttons on the panel, one for each heating element. Each button consists of a red push button for switching the heaters on and off and a slider for adjusting the power usage for each heater (Figure 13c and Figure 13d). In addition, a screen above the control panel presents real-time information regarding each heater's surface temperature and power consumption, which was initially accessible in VRT mode. As soon as the heating process is finished, a push button is activated to allow user to deactivate all the heaters at once (Figure 13e). Also, once the operator's palm hovers above a heating element, the screen's information is updated immediately. Figure 13f illustrates step 5, in which the user will move down the sheet holder and place it on the mould. Lastly, by activating the vacuum pump thermoplastic sheet is formed as indicated in Figure 13g. Figure 14 demonstrates the MR application’s features, which are accessible to the operator who is working beside the actual developed thermoforming setup in the laboratory environment.

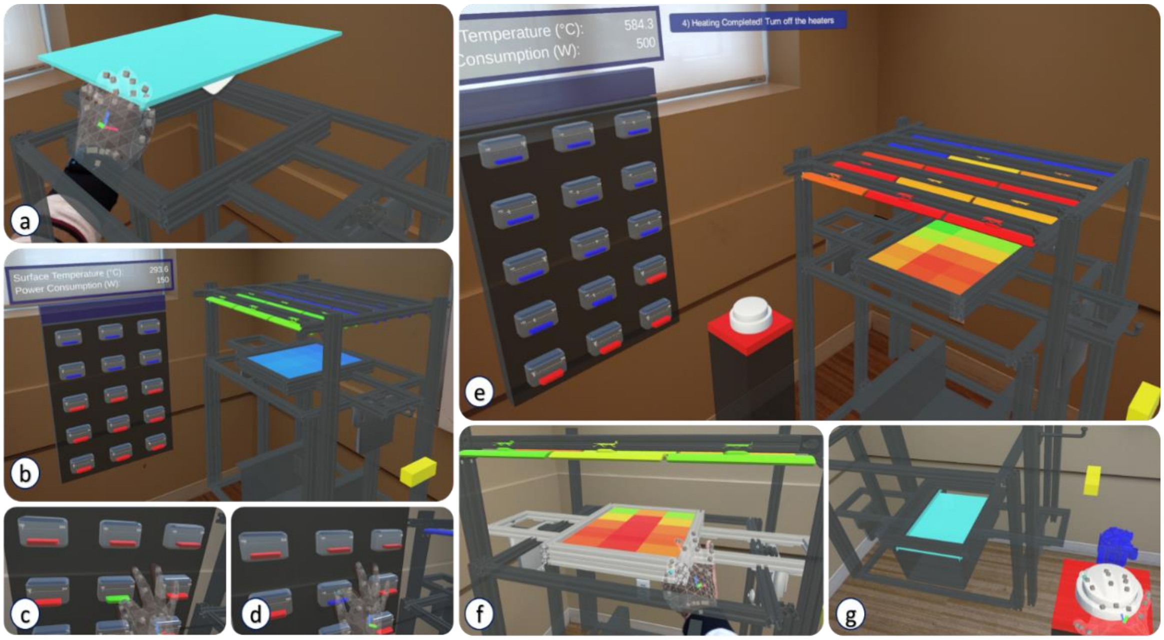

Figure 13 – Steps of the developed MR application for training operators. a) Placing the thermoplastic sheet on the sheet holder. b) Prompted control panel to allow user interactions. c) Activating the heaters using the buttons on control panel. d) Adjusting the heater's power range using the slider. e) Deactivating all the heaters at once using the push button. f) Placing the sheet holder on the mould. g) Activating the vacuum pump, using the push button.

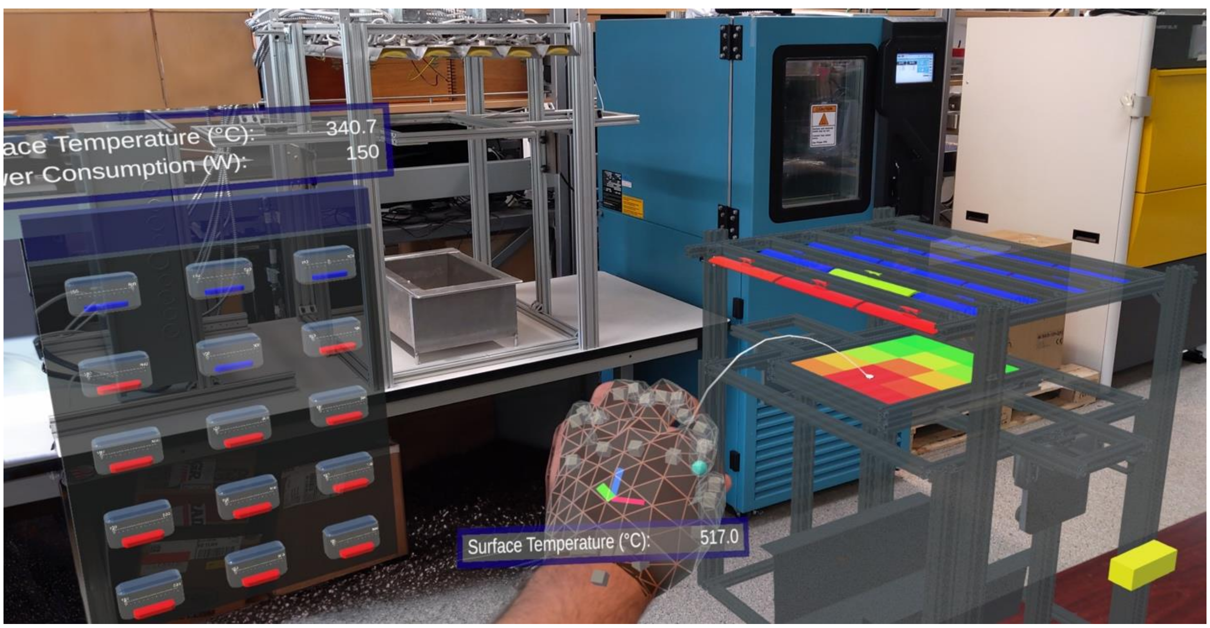

Figure 14- Features of the developed MR application: control panel on the left, the virtual setup on the right and the actual setup between them. In the shown example, the first row of heaters (in red) in the virtual set-up are working at full power (here 500W), and a heater in the middle of the second row (in green) is working at 150W power (the screen on top of the control panel depicts the info about the latter heater). Additionally, the sheet is heated, and the colour gradient in various regions reflects its surface temperature, as it is also shown on the operator's hand once remotely selected via hand raycast.

3.3. Heat transfer model visualization

Within both VR and MR apps as soon as the heating process begins in step 3, the colour gradient (temperature of the heaters) is updated graphically in real-time depending on the power consumption of each heater; with dark blue representing the lowest temperature and red representing the highest.

3.4. User study

After going through all the training methods (DVT, MRT and VRT) in their Latin square orders, the participants were asked to take a survey and provide feedback on the three training platforms experienced, considering usefulness, ease-of-use, satisfaction, trust, reliability, training effectiveness and overall preference (as outlined in section 2.4.2). In order to compare all the obtained data from the user's feedback, the procedure defined in Section 2.5 was employed. Figure 15 illustrates the normalized mean and standard deviation for each training method, considering each performance criterion. As seen, in most criteria, MRT has had the highest mean compared to VRT and DVT, except for the 'ease of use and trust' criterion. Moreover, VRT was never ranked as the top priority in any of the mean measurements, while being considered the least trustable, convenient, and reliable training platform.

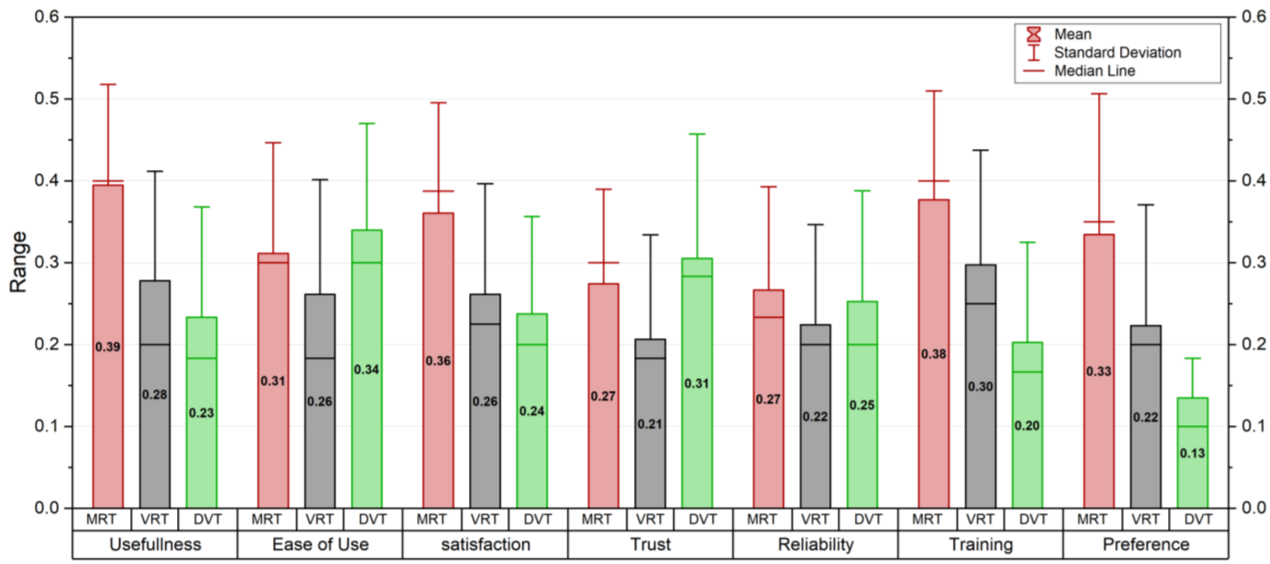

Figure 15 – The distribution of data obtained for each training method with respect to the performance criteria.

Additionally, a comparison of cumulative means among the three training methods is shown in Figure 16. The latter also clearly indicates that MRT has the highest cumulative mean with respect to all the criteria combined. Finally, to compare the responses between novice and experienced XR users, the means of all the criteria for each training method were compared, as illustrated in Figure 17 (discussion of results to follow in subsequent sections).

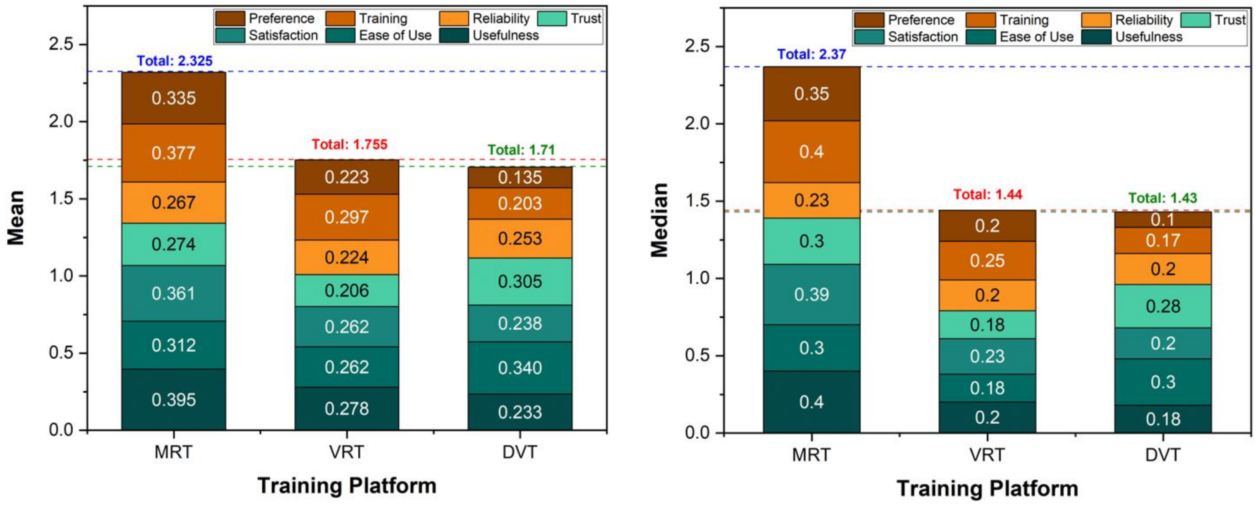

Figure 16 – Stacked mean (left) and median (right) performance scores, among different training platforms.

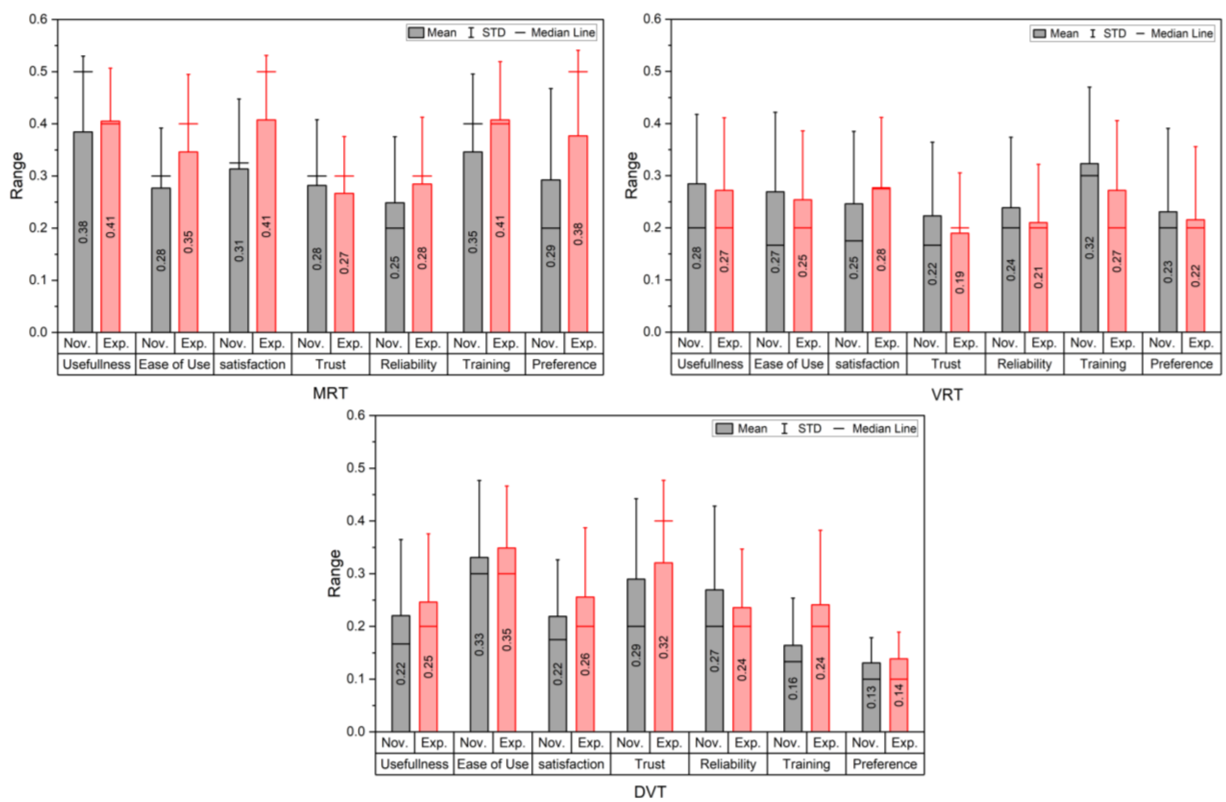

Figure 17 - Comparison of XR novices and XR experienced with respect to the measured criteria, for each of the three training methods.

To assess the reliability of the designed questionnaire, the Cronbach's alpha coefficient was calculated as a measure of overall internal consistency of the survey. The obtained mean value of Cronbach's alpha was 0.922, indicating a very high level of consistency between the designed survey questions and responses.

3.4.1. ANOVA with blocks and post-hoc tests

One-way ANOVA with blocks was implemented on the collected data to specify whether the described criteria means in Figure 15 are statistically different under each criterion. To further analyze the mean differences, post hoc tests were conducted, to specifically identify which groups are different. Bonferroni correction was used as the post hoc method to conclude the results. Before conducting the ANOVA test on each criterion, the normality and the variance of the data were also checked. Figure 18a demonstrates a sample normality check as a Q-Q plot with a 95% confidence interval for the usefulness criterion. Moreover, the data population variance for the same criterion is illustrated in Figure 18b. The same procedure was employed for all the other criteria to ensure the data is suitable for conducting ANOVA tests (see the **Appendix)**. The ANOVA with block measurements ($F(2, 77) = 15.086$, p-value = 7.48E-06) determined a statistically significant difference in **usefulness** criterion across the training platforms, as indicated in Table 1 (see the **Appendix** for results of other criteria).

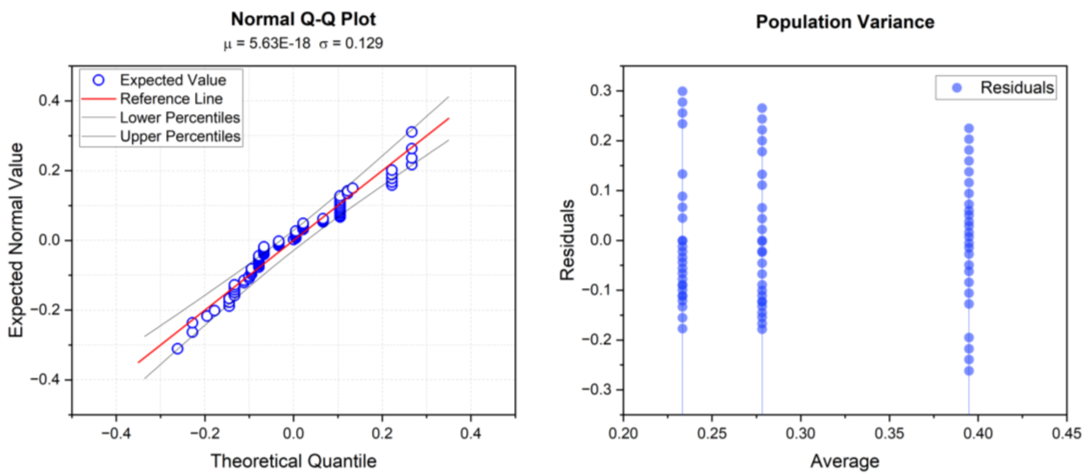

Figure 18 – a) Normality check via Q-Q plot of the usefulness criterion data. b) Constant variance check for the criterion.

Table 1 - ANOVA with blocks results for the usefulness criterion.

Source of Variation	*Sum of Squares*	*df*	*Mean Square*	*F-value*	*P-value*	*F critical*
Training Platforms	0.361	2	0.180	15.086	7.48E-06	3.182
Users	0.682	25	0.027	2.277	0.006	1.727
Error	0.599	50	0.012			
Total	1.643	77				

The post-hoc analysis using the Bonferroni correction revealed that the VRT - MRT as well as DVT – MRT pairs have statistically significant different means, and in both cases, MRT is more useful compared to the others (MRT ($M = 0.394$, $SD = 0.123$); VRT ($M = 0.278$, $SD = 0.133$); DVT ($M = 0.233$, $SD = 0.134$)). There was no significant difference in usefulness between DVT – VRT pair. Same procedure for the rest of the criteria was followed to identify the significant factors affecting the training methods. Participants' answers for the **ease-of-use** criterion (F(2, 77)

= 1.903, p-value = 0.159) reflected that in terms of convenience, there is no statistically significant difference between the alternatives and therefore, no pairwise comparison is necessary. Moreover, the outcome of the one-way ANOVA with block test for the **satisfaction** (F(2, 77) = 9.247, p-value = 3.82E-4), suggested that MRT is statistically more significantly satisfactory compared to VRT and DVT. Additionally, a pairwise Bonferroni correction test indicated that there is statistically significant difference between VRT (M = 0.261, SD = 0.134) and MRT (M = 0.36, SD = 0.135); as well as DVT (M = 0.237, SD = 0.119) and MRT (M = 0.36, SD = 0.135). There was no significant difference found in usefulness between DVT – VRT pair.

In terms of user's **trust** on the training platforms, participants provided a statistically significantly higher rating for the DVT (M = 0.305, SD = 0.152) over the MRT (M = 0.274, SD = 0.115) and VRT (M = 0.206, SD = 0.127) methods, where one-way ANOVA with block test indicated F(2, 77) = 6.101, p-value = 0.004. A post hoc pairwise test via Bonferroni method showed that DVT – VRT is the only pair with statistically significant mean difference, where DVT is preferred over VRT in terms of user's trust. On the other hand, the obtained results of one-way ANOVA with block for **reliability** measurements, proved that there is no statistically significant difference between the training platforms (F(2, 77) = 0.796, p-value = 0.456).

Statistical analysis demonstrated that the **training** (F(2, 77) = 12.52, p-value = 3.90E-5) with MRT (M = 0.376, SD = 0.132) is the most efficient method in comparison to VRT (M = 0.297, SD = 0.140) and DVT (M = 0.202, SD = 0.122). Bonferroni pairwise comparison suggested that the DVT – MRT and DVT – VRT pairs are the ones with statistically significant mean differences, which is not the case for the VRT – MRT pair. Finally, the outcome of the one-way ANOVA with block test for the **Preference** criterion (F(2, 77) = 13.35, p-value = 2.25E-5), suggested that MRT is statistically significantly preferred the most, compared to VRT and DVT. Furthermore, the pairwise Bonferroni test showed that there is statistically significant difference between the pair of VRT (M = 0.223, SD = 0.147) and MRT (M = 0.334, SD = 0.171); as well as the pair of DVT (M = 0.134, SD = 0.048) and MRT (M = 0.334, SD = 0.171). There was no significant difference found in usefulness between DVT – VRT pair.

Moreover, an evaluation was conducted to assess the impact of participants' prior exposure to XR technology, as depicted in Figure 17. The results of the statistical analysis indicated that no statistically significant differences between the two groups are observed with respect to any of the criteria, including **Usefulness** (F(1, 77) = 0.136, p-value = 0.713), **Ease of use** (F(1, 77) = 0.608, p-value = 0.438), **User Satisfaction** (F(1, 77) = 3.459, p-value = 0.067), **Trust** (F(1, 77) = 0.039, p-value = 0.844), **Reliability** (F(1, 77) = 0.086, p-value = 0.770), **Training** (F(1, 77) = 0.943, p-value = 0.335), and **Overall preference** (F(1, 77) = 0.713, p-value = 0.401).

In terms of task completion time, the initiation and conclusion of participants' training sessions for each training method were recorded and subjected to analysis. A statistically significant difference in **Completion Time** across the training platforms was determined using ANOVA with block measurements (F(2, 77) = 4.302, p-value = 0.017). Subsequent post-hoc analysis, employing the Bonferroni correction, revealed that the DVT and VRT exhibited statistically significant differences in means. Specifically, VRT (M = 220.1, SD = 28.62) emerged as the fastest training method, while DVT (M = 244.1, SD = 33.1) was the slowest. In contrast, MRT (M = 236.2, SD = 28.2) demonstrated intermediate performance.

3.4.2. MCDM

As outlined in Section 2.6, five conventional approaches of MCDM were implemented to find the best alternative among the three training methods. Based on the above results from the statistical analysis and the ANOVA tests, out of all the seven subjective criteria defined to assess the training efficiency of the platforms, Ease-of-use and Reliability were eliminated as there was no statistically significant mean differences, and hence they cannot be part of final decision-making in this case study. Furthermore, Completion Time was found to be statistically significant as an objective metric. This left the MCDM problem with six criteria, which are shown in Table 2; including the criterion types indicated by min or max (cost-type or benefit-type) and their weights, as well as the calculated means of the alternatives. As can be seen, the best mean values (shown in bold) under different criteria pertain to different training methods, and therefore to decide the best alternative, MCDM techniques are needed.

Given the study's focus on comparing three distinct training methods, the highest weight was assigned to training effectiveness, reflecting its paramount importance. Additionally, usability holds significant weight due to its crucial role. User satisfaction and trust, while substantial, occupy an intermediate level of importance compared to the preceding criteria. Finally, overall preference and completion time bear the lowest weights, as factors such as preference and training speed were comparatively less critical when compared to training effectiveness and usability. Lastly, in preparation for the MCDM analysis, normalization of the mean measurements for each criterion was undertaken (not included in Table 2).

Table 2 - Means of the three training methods (decision alternatives) per each criterion along with the criteria weights assigned.

CRIT. / ***ALT.***	***Usefulness (max) W. 0.25***	***Satisfaction (max) W. 0.2***	***Trust (max) W. 0.15***	***Training (max) W. 0.3***	***Preference (max) W. 0.05***	***Time (min) W.0.05***
MRT	**0.3949**	**0.3606**	0.2744	**0.3769**	**0.3346**	236.2
VRT	0.2782	0.2615	0.2064	0.2974	0.2231	**220.1**
DVT	0.2333	0.2375	**0.3051**	0.2026	0.1346	244.1

Figure 19 demonstrates the final rankings of alternatives based on the five MCDM approaches of TOPSIS, WSM, WPM, Multi MOORA, and VIKOR. MRT is considered the first rank based on the results obtained from all the MCDM methods, whereas VRT and DVT share the second and third ranks.

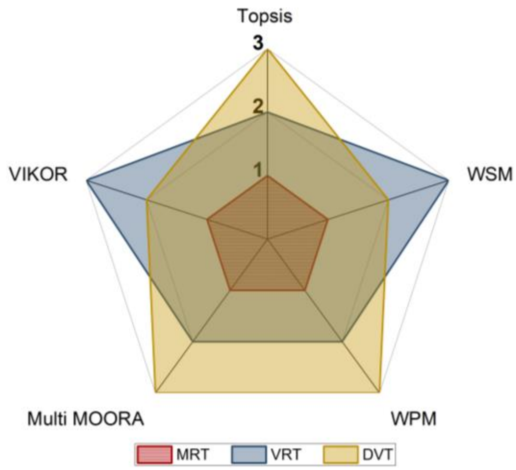

Figure 19 - Comparative rankings of the three training methods based on five MCDM approaches; numbers within the Spider plot indicate ranks.

Finally, to assess the effect of user/expert opinion criteria weights on the MCDM outcome, a sensitivity analysis was conducted, considering seven different weighting scenarios/cases. Case 1 provides equal weights among all the criteria, whereas the rest of the cases considered five criteria with the minimum weights, to emphasize on the sixth criterion with the highest weight (Table 3). The TOPSIS method was then implemented for each case. Based on Table 3, the rankings of the alternatives based on various criteria weighting cases consistently indicated that MRT is the preferred alternative as compared to the other training methods. In particular, notice that the score of MRT (out of the maximum 1.0) is consistently higher than 0.8, except for cases 2 and 5.

Table 3 – TOPSIS Rankings of the training methods (alternatives) based on various sets of criteria weights.

Criteria	**Base case**	**Case 1**	**Case 2**	**Case 3**	**Case 4**	**Case 5**	**Case 6**	**Case 7**
Usefulness	0.25	0.166	0.1	0.1	0.1	0.1	0.1	0.5
Satisfaction	0.2	0.166	0.1	0.1	0.1	0.1	0.5	0.1
Trust	0.15	0.166	0.1	0.1	0.1	0.5	0.1	0.1
Training Effectiveness	0.3	0.166	0.1	0.1	0.5	0.1	0.1	0.1
Overall Preference	0.05	0.166	0.1	0.5	0.1	0.1	0.1	0.1
Completion Time	0.05	0.166	0.5	0.1	0.1	0.1	0.1	0.1
Alternatives				***Ranking (scores)***				
MRT	**1 (*0.917*)**	**1(*0.887*)**	**1 (*0.771*)**	**1 (*0.965*)**	**1 (*0.951*)**	**1 (*0.677*)**	**1 (*0.935*)**	**1 (*0.948*)**
VRT	2 (*0.416*)	2 (*0.399*)	2 (*0.477*)	2 (*0.438)*	2 (*0.529*)	3 (*0.237*)	2 (*0.273*)	2 (*0.296*)
DVT	3 (*0.165*)	3 (*0.201*)	3 (*0.189*)	3 (*0.065*)	3 (*0.091*)	2 (*0.558*)	3 (*0.121*)	3 (*0.098*)

4. A Qualitative Interpretation of User Study Outcomes

The quantitative results derived from the above user study collectively revealed a significant enhancement in users' training experiences upon utilizing the MRT system. Remarkably, even among participants who had no prior exposure to either the MR system or the thermoforming process, the consensus favored the MRT interface over both the DVT and VRT alternatives. It was interesting to note that participants' experience levels had no discernible impact on the subjective criteria, which could be attributed to the intuitive UI design of the present immersive applications.

Despite MRT consistently ranking the highest according to the MCDM, trust emerged as a pivotal factor influencing users' preferences; as some users favored DVT over VRT. Particularly, users highly valued MRT in terms of training effectiveness and overall preference, while DVT garnered the least favor in both criteria. Users' post-experimental qualitative feedback revealed a desire to maintain awareness of their surroundings during training, offering a rationale for the preference for MRT. Additionally, VRT received the highest score in terms of training effectiveness, with users indicating a preference for the input method within the VRT system over MRT. The joystick input, as opposed to the hand gestures utilized in the MR HMD, was cited as being more user-friendly. However, despite this input method preference, the results decisively established MRT as the superior environment when compared to VRT, affirming its suitability for training purposes in the present manufacturing application. In essence, while the input method within VRT may have been preferred due to its familiarity, the broader context of user experience, effectiveness, and situational awareness within MRT ultimately established it as the superior choice for training purposes. The users' preference for input method alone did not overshadow the holistic advantages offered by the MRT system.

Based on the initial sub-hypotheses and the quantitative evaluations conducted through the user study and MCDM, it became evident that overall, MRT is the most suitable platform for training purposes in the current application. Sub-H1 posited that interacting with the designed DT in the MRT environment would be more *beneficial* for training, and the statistical analysis supported this assertion. Sub-H2, which had suggested that operators would feel more *comfortable* being trained in an MRT environment compared to VRT and DVT, was similarly substantiated by the statistical results. Moreover, Sub-H3, which examined *user satisfaction* with their performance during and after training, favored MRT over VRT and DVT. Sub-H4 proposed that MRT would be the most *trustable* environment, and Sub-H5 suggested it would be the most *reliable* source for trainees to function properly, both of which were confirmed by the user study findings.

Limitations: Although the study made efforts to include a diverse range of target users, it still faced limitations in terms of accessing participants with specific manufacturing skills, particularly thermoforming. The cognitive skills, notably procedural memory, as perceived by thermoforming experts, may have had an impact on the study's outcomes. Furthermore, it is worth noting that Head-Mounted Displays (HMDs) have limitations in computational capacity, leading to the necessity of reducing simulation resolution to enable real-time interaction for operators. There is also room for improvement in the user interface (UI) design to enhance the overall user experience, acknowledging that UI design enhancements can continually refine the user's interaction with the system. Another constraint of the study was its reliance solely on completion time as the objective metric. Additional factors, such as the frequency of errors during task completion for each training method, or an examination of training transfer encompassing both immediate and long-term recall, can possibly provide a more comprehensive perspective. Future MCDM evaluation framework may also include complex human factors such as emotions or e.g. cognitive resistance to change [59] toward a non-conventional training method, etc.

5. Conclusion

With recent breakthroughs in immersive technologies and the emergence of head-mounted displays (HMD), new prospects await industrial processes that rely on smart interactions between machines and human operators. Advanced polymers and composites manufacturing sector is a seamless beneficiary of such technologies, with a promising future in joining I5.0, as it involves many manual or minimally automated procedures during different material processing techniques. Manual methods for staff training in such manufacturing processes using experimental trials can be very costly and time-consuming, which brings about a critical need for the adaptation of immersive technologies.

This case study developed a novel adaptive and interactive DT interface in both VR and MR environments to demonstrate how a complex thermoforming process used in thermoplastic manufacturing can be simulated in real-time and used for training novice operators under immersive experience. Meta Quest 2 and Microsoft HoloLens 2 were employed as the target devices to validate the study hypothesis. Aligned with the earlier research, several evaluation criteria for Virtual Reality Tool (VRT) and Mixed Reality Tool (MRT) were employed, including Usefulness, Ease of use, Satisfaction, Trust, Reliability, Training effectiveness, and overall user Preference in addition to completion time. To provide a comprehensive evaluation of the immersive platforms, a demo video tool (DVT) was also employed as the benchmark training method. Regarding the aforementioned criteria and training methods, a user study (n=26) was carried out and the acquired data was analyzed using one-way ANOVA with blocks. Only five of the seven criteria, including Usefulness, Satisfaction, Training Efficacy, Trust, and Overall User Preference, exhibited statistically significant mean differences across the training platforms. Accordingly, Multiple Criteria Decision Making (MCDM) techniques were employed to rank the training platforms. Except for the 'Ease of use' and 'Trust' criterion, the MCDM findings and sensitivity analysis of criteria weights demonstrated a strong preference of the users towards MRT (score range: 0.677 to 0.965) compared to VRT and DVT. Although it was the fastest training method, the VRT consistently scored second, with ratings ranging from 0.237 to 0.529, despite being the least reliable, convenient, and trustable training platform. Except for one scenario in the sensitivity analysis (case 5), DVT was overall the least favored training technique. In the base case with weights determined by expert opinion, the MCDM results revealed that the preference score for MRT was 2.2 times higher than that of VRT; and 5.56 times higher than that of DVT.

According to the sensitivity analysis, overall, 'Trust' was the most influential criterion affecting the rating order of the training platforms.

Acknowledgment

The authors wish to acknowledge the financial support from the New Frontiers in Research Fund-Exploration program in Canada (NFRFE-2019-01440). In addition, the authors are grateful to their colleagues at the Composites Research Network and Mr. Hadi Hosseinionari at UBC for stimulating discussions.

Declaration of Competing Interest

The authors declare that they have no conflicts of interest.

References

[1] A. Ustundag and E. Cevikcan, *Industry 4.0: Managing The Digital Transformation*, no. January. 2018.

[2] S. Doolani *et al.*, “A Review of Extended Reality (XR) Technologies for Manufacturing Training,” *Technologies*, vol. 8, no. 4, p. 77, 2020, doi: 10.3390/technologies8040077.

[3] J. E. Naranjo, D. G. Sanchez, A. Robalino-Lopez, P. Robalino-Lopez, A. Alarcon-Ortiz, and M. V. Garcia, “A scoping review on virtual reality-based industrial training,” *Appl. Sci.*, vol. 10, no. 22, pp. 1–31, 2020, doi: 10.3390/app10228224.

[4] A. de Giorgio, M. Romero, M. Onori, and L. Wang, “Human-machine Collaboration in Virtual Reality for Adaptive Production Engineering,” *Procedia Manuf.*, vol. 11, no. June, pp. 1279–1287, 2017, doi: 10.1016/j.promfg.2017.07.255.

[5] F. G. Pratticò and F. Lamberti, “Towards the adoption of virtual reality training systems for the self-tuition of industrial robot operators: A case study at KUKA,” *Comput. Ind.*, vol. 129, 2021, doi: 10.1016/j.compind.2021.103446.

[6] J. Leder, T. Horlitz, P. Puschmann, V. Wittstock, and A. Schütz, “Comparing immersive virtual reality and powerpoint as methods for delivering safety training: Impacts on risk perception, learning, and decision making,” *Saf. Sci.*, vol. 111, no. May 2018, pp. 271–286, 2019, doi: 10.1016/j.ssci.2018.07.021.

[7] E. D. Ragan, D. A. Bowman, R. Kopper, C. Stinson, S. Scerbo, and R. P. McMahan, “Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task,” *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 7, pp. 794–807, 2015, doi: 10.1109/TVCG.2015.2403312.

[8] D. Sportillo, G. Avveduto, F. Tecchia, and M. Carrozzino, “Training in VR: A Preliminary Study on Learning Assembly/Disassembly Sequences,” 2015, pp. 332–343.

[9] M. Murcia López, “The effectiveness of training in virtual environments,” *Dr. thesis, UCL (University Coll. London).*, p. 157, 2018, [Online]. Available: https://discovery.ucl.ac.uk/id/eprint/10063411/.

[10] Y. Vélaz, J. R. Arce, T. Gutiérrez, A. Lozano-Rodero, and A. Suescun, “The influence of interaction technology on the learning of assembly tasks using virtual reality,” *J. Comput. Inf. Sci. Eng.*, vol. 14, no. 4, pp. 1–9, 2014, doi: 10.1115/1.4028588.

[11] Y. Pan and A. Steed, “Avatar type affects performance of cognitive tasks in virtual reality,” *Proc. ACM Symp. Virtual Real. Softw. Technol. VRST*, 2019, doi: 10.1145/3359996.3364270.

[12] B. Salah, M. H. Abidi, S. H. Mian, M. Krid, H. Alkhalefah, and A. Abdo, “Virtual reality-based engineering education to enhance manufacturing sustainability in industry 4.0,” *Sustain.*, vol. 11, no. 5, pp. 1–19, 2019, doi: 10.3390/su11051477.

[13] Z. H. Lai, W. Tao, M. C. Leu, and Z. Yin, “Smart augmented reality instructional system for mechanical assembly towards worker-centered intelligent manufacturing,” *J. Manuf. Syst.*, vol. 55, no. July 2019, pp. 69–81, 2020, doi: 10.1016/j.jmsy.2020.02.010.

[14] G. Westerfield, A. Mitrovic, and M. Billinghurst, “Intelligent augmented reality training for

motherboard assembly," *Int. J. Artif. Intell. Educ.*, vol. 25, no. 1, pp. 157–172, 2015, doi: 10.1007/s40593-014-0032-x.

[15] C. H. Chu and Y. L. Liu, "Augmented reality user interface design and experimental evaluation for human-robot collaborative assembly," *J. Manuf. Syst.*, vol. 68, no. April, pp. 313–324, 2023, doi: 10.1016/j.jmsy.2023.04.007.

[16] C. H. Chu and C. H. Ko, "An experimental study on augmented reality assisted manual assembly with occluded components," *J. Manuf. Syst.*, vol. 61, no. April, pp. 685–695, 2021, doi: 10.1016/j.jmsy.2021.04.003.

[17] M. Dalle Mura and G. Dini, "An augmented reality approach for supporting panel alignment in car body assembly," *J. Manuf. Syst.*, vol. 59, no. February, pp. 251–260, 2021, doi: 10.1016/j.jmsy.2021.03.004.

[18] S. Webel, U. Bockholt, T. Engelke, N. Gavish, M. Olbrich, and C. Preusche, "An augmented reality training platform for assembly and maintenance skills," *Rob. Auton. Syst.*, vol. 61, no. 4, pp. 398–403, 2013, doi: 10.1016/j.robot.2012.09.013.

[19] B. Schwald and B. DeLaval, "An Augmented Reality System for Training and Assistance to Maintenance in the Industrial Context," *11th Int. Conf. Cent. Eur. Comput. Graph. Vis. Comput. Vis.*, pp. 425–432, 2003.

[20] S. Werrlich, E. Eichstetter, K. Nitsche, and G. Notni, "An Overview of Evaluations Using Augmented Reality for Assembly Training Tasks," *Int. J. Comput. Inf. Eng.*, vol. 11, no. 10, pp. 1129–1135, 2017, [Online]. Available: https://www.researchgate.net/profile/Stefan_Werrlich/publication/324438098_An_Overview_of_Evaluations_Using_Augmented_Reality_for_Assembly_Training_Tasks/data/5acdb18e4585154f3f40e904/DOK-20170812-An-Overview-of-Evaluations-Using-Augmented-Reality-for-Ass.

[21] M. Heinz, S. Büttner, and C. Röcker, "Exploring training modes for industrial augmented reality learning," *ACM Int. Conf. Proceeding Ser.*, pp. 398–401, 2019, doi: 10.1145/3316782.3322753.

[22] M. Hoover, "An evaluation of the Microsoft HoloLens for a manufacturing-guided assembly task," Iowa State University, Digital Repository, Ames, 2018.

[23] A. Muñoz, X. Mahiques, J. E. Solanes, A. Martí, L. Gracia, and J. Tornero, "Mixed reality-based user interface for quality control inspection of car body surfaces," *J. Manuf. Syst.*, vol. 53, no. August, pp. 75–92, 2019, doi: 10.1016/j.jmsy.2019.08.004.

[24] D. K. Baroroh and C. H. Chu, "Human-centric production system simulation in mixed reality: An exemplary case of logistic facility design," *J. Manuf. Syst.*, vol. 65, no. March, pp. 146–157, 2022, doi: 10.1016/j.jmsy.2022.09.005.

[25] N. Murauer, D. Schön, F. Müller, N. Pflanz, S. Günther, and M. Funk, "An analysis of language impact on augmented reality order picking training," *ACM Int. Conf. Proceeding Ser.*, pp. 351–357, 2018, doi: 10.1145/3197768.3201570.

[26] M. Sirakaya and E. K. Cakmak, "Effects of augmented reality on student achievement and self-efficacy in vocational education and training," *Int. J. Res. Vocat. Educ. Train.*, vol. 5, no. 1, pp. 1–18, 2018, doi: 10.13152/IJRVET.5.1.1.

[27] A. D. Kaplan, J. Cruit, M. Endsley, S. M. Beers, B. D. Sawyer, and P. A. Hancock, "The Effects of Virtual Reality, Augmented Reality, and Mixed Reality as Training Enhancement Methods: A Meta-Analysis," *Hum. Factors*, vol. 63, no. 4, pp. 706–726, 2021, doi: 10.1177/0018720820904229.

[28] S. Farra, E. Miller, N. Timm, and J. Schafer, "Improved Training for Disasters Using 3-D Virtual Reality Simulation," *West. J. Nurs. Res.*, vol. 35, no. 5, pp. 655–671, 2013, doi: 10.1177/0193945912471735.

[29] S. Pedram, S. Palmisano, R. Skarbez, P. Perez, and M. Farrelly, "Investigating the process of mine rescuers' safety training with immersive virtual reality: A structural equation modelling approach," *Comput. Educ.*, vol. 153, no. March, p. 103891, 2020, doi: 10.1016/j.compedu.2020.103891.

[30] M. Gonzalez-Franco *et al.*, "Immersive mixed reality for manufacturing training," *Front. Robot. AI*, vol. 4, no. FEB, pp. 1–8, 2017, doi: 10.3389/frobt.2017.00003.

[31] I. Jalilvand, J. Jiyoung, H. Hosseinionari, R. Seethaler, B. Gopaluni, and A. S. Milani, "An Interactive Digital Twin of a Composite Manufacturing Process for Training Operators via Immersive Technology," *Commun. Comput. Inf. Sci.*, vol. 1836 CCIS, pp. 207–214, 2023, doi: 10.1007/978-3-031-36004-6_28.

[32] Q. Hernández, A. Badías, F. Chinesta, and E. Cueto, "Thermodynamics-informed neural networks for physically realistic mixed reality," *Comput. Methods Appl. Mech. Eng.*, vol. 407, p. 115912, 2023, doi: 10.1016/j.cma.2023.115912.

[33] A. Badías, D. González, I. Alfaro, F. Chinesta, and E. Cueto, "Real-time interaction of virtual and physical objects in mixed reality applications," *Int. J. Numer. Methods Eng.*, vol. 121, no. 17, pp. 3849–3868, 2020, doi: 10.1002/nme.6385.

[34] V. Weistroffer, F. Keith, A. Bisiaux, C. Andriot, and A. Lasnier, "Using Physics-Based Digital Twins and Extended Reality for the Safety and Ergonomics Evaluation of Cobotic Workstations," *Front. Virtual Real.*, vol. 3, no. February, pp. 1–18, 2022, doi: 10.3389/frvir.2022.781830.

[35] J. Li, H. Deng, Y. Gao, A. Chen, Z. Song, and A. Hao, "Real-time Physics-based Interaction in Augmented Reality," *Proc. - 2023 IEEE Conf. Virtual Real. 3D User Interfaces Abstr. Work. VRW 2023*, pp. 573–574, 2023, doi: 10.1109/VRW58643.2023.00129.

[36] T. Koduri, D. Bogdoll, S. Paudel, and G. Sholingar, "AUREATE: An Augmented Reality Test Environment for Realistic Simulations," *SAE Tech. Pap.*, vol. 2018-April, pp. 1–9, 2018, doi: 10.4271/2018-01-1080.

[37] A. Iop *et al.*, "Extended Reality in Neurosurgical Education: A Systematic Review," *Sensors*, vol. 22, no. 16, 2022, doi: 10.3390/s22166067.

[38] W. de O. Leite, J. C. C. Rubio, F. M. Cabrera, A. Carrasco, and I. Hanafi, "Vacuum thermoforming process: An approach to modeling and optimization using artificial neural networks," *Polymers (Basel).*, vol. 10, no. 2, 2018, doi: 10.3390/polym10020143.

[39] J. P. Holman, *Heat transfer, New York, 10th edition.* McGraw-Hill, 2009.

[40] M. I. Chy and P. O. November, "Estimation and Control of Plastic Temperature in Heating Phase of Thermoforming Process," 2013.

[41] M. M. I. Chy and B. Boulet, “A new method for estimation and control of temperature profile over a sheet in thermoforming process,” *Conf. Rec. - IAS Annu. Meet. (IEEE Ind. Appl. Soc.*, pp. 5–12, 2010, doi: 10.1109/IAS.2010.5615714.

[42] F. Erchiqui and G. D. Ngoma, “Analyse comparative des méthodes de calcul des facteurs de formes pour des surfaces à contours rectilignes,” *Int. J. Therm. Sci.*, vol. 46, no. 3, pp. 284–293, 2007, doi: 10.1016/j.ijthermalsci.2006.06.001.

[43] H. Hosseinionari, M. Ramezankhani, R. Seethaler, and A. S. Milani, “Development of a Computationally Efficient Model of the Heating Phase in Thermoforming Process Based on the Experimental Radiation Pattern of Heaters,” 2023.

[44] J. Brooke, “SUS: A ‘Quick and Dirty’ Usability Scale,” *Usability Eval. Ind.*, no. July, pp. 207–212, 1995, doi: 10.1201/9781498710411-35.

[45] J. R. Lewis and J. Sauro, “The factor structure of the system usability scale,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5619 LNCS, pp. 94–103, 2009, doi: 10.1007/978-3-642-02806-9_12.

[46] L. J. Cronbach, “Coefficient alpha and the internal structure of tests,” *Psychometrika*, vol. 16, no. 3, pp. 297–334, 1951, doi: 10.1007/BF02310555.

[47] K. Bilisik, “Pull-out properties of polyester woven fabrics: Effects of softening agent and interlacement on single and multiple yarn pull-out forces and analysis by statistical model,” *Fibers Polym.*, vol. 12, no. 8, pp. 1106–1118, 2011, doi: 10.1007/s12221-011-1106-0.

[48] B. IşIk and E. Ekici, “Experimental investigations of damage analysis in drilling of woven glass fiber-reinforced plastic composites,” *Int. J. Adv. Manuf. Technol.*, vol. 49, no. 9–12, pp. 861–869, 2010, doi: 10.1007/s00170-009-2440-x.

[49] D. C. A. S. U. Montgomery, *D esign and Analysis of Experiments Ninth Edition*. 2017.

[50] C. Mircioiu and J. Atkinson, “A Comparison of Parametric and Non-Parametric Methods Applied to a Likert Scale,” *Pharmacy*, vol. 5, no. 4, p. 26, 2017, doi: 10.3390/pharmacy5020026.

[51] J. J. Tzeng, G. H., & Huang, *Multiple attribute decision making: methods and applications.* CRC press., 2011.

[52] M. Rabbani, N. A. Saravi, H. Farrokhi-Asl, S. F. W. T. Lim, and Z. Tahaei, “Developing a sustainable supply chain optimization model for switchgrass-based bioenergy production: A case study,” *J. Clean. Prod.*, vol. 200, pp. 827–843, Nov. 2018, doi: 10.1016/j.jclepro.2018.07.226.

[53] J.-J. H. Gwo-Hshiung Tzeng, *Multiple Attribute Decision Making: Methods and applications*, vol. 4, no. 1. 2011.

[54] O. Onajite and S. A. Oke, “The Application of WSM, WPM and WASPAS Multicriteria Methods for Optimum Operating Conditions Selection in Machining Operations,” *J. Rekayasa Sist. Ind.*, vol. 10, no. 1, pp. 1–14, 2021, doi: 10.26593/jrsi.v10i1.4271.1-14.

[55] E. K. Zavadskas, J. Antucheviciene, J. Saparauskas, and Z. Turskis, “MCDM methods WASPAS and MULTIMOORA: Verification of robustness of methods when assessing alternative solutions,” *Econ. Comput. Econ. Cybern. Stud. Res.*, vol. 47, no. 2, 2013.

[56] S. Opricovic and G. H. Tzeng, "Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS," *Eur. J. Oper. Res.*, vol. 156, no. 2, pp. 445–455, 2004, doi: 10.1016/S0377-2217(03)00020-1.

[57] S. Opricovic and G. H. Tzeng, "Extended VIKOR method in comparison with outranking methods," *Eur. J. Oper. Res.*, vol. 178, no. 2, pp. 514–529, 2007, doi: 10.1016/j.ejor.2006.01.020.

[58] P. Tučník and V. Bureš, "Experimental Evaluation of Suitability of Selected Multi-Criteria Decision-Making Methods for Large-Scale Agent-Based Simulations," *PLoS One*, vol. 11, no. 11, p. e0165171, Nov. 2016, doi: 10.1371/journal.pone.0165171.

[59] A. S. Milani, A. Shanian, and C. El-Lahham, "A decision-based approach for measuring human behavioral resistance to organizational change in strategic planning," *Math. Comput. Model.*, vol. 48, no. 11–12, pp. 1765–1774, 2008, doi: 10.1016/j.mcm.2008.06.018.

Appendix A: Additional ANOVA Data and Checking its Assumptions

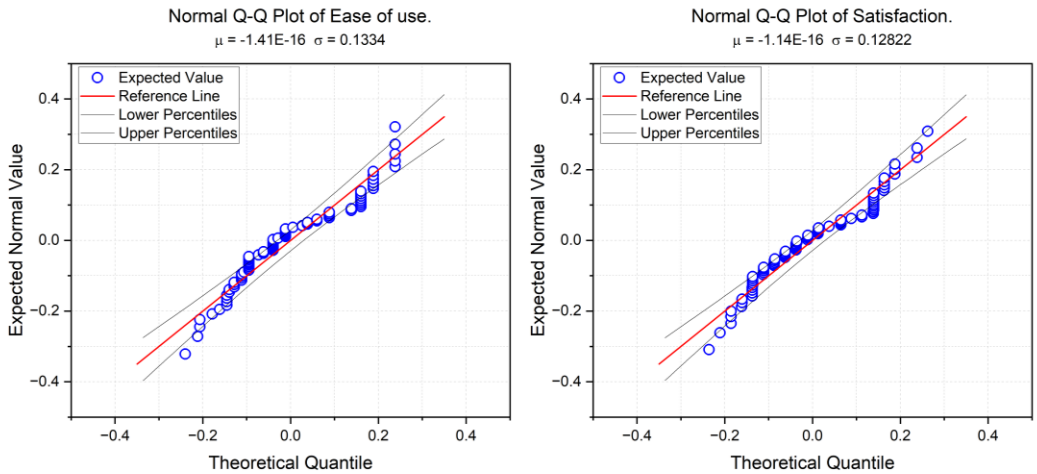

Figure A1 – Normality check via Q-Q plots: ease of use (left) and satisfaction (right).

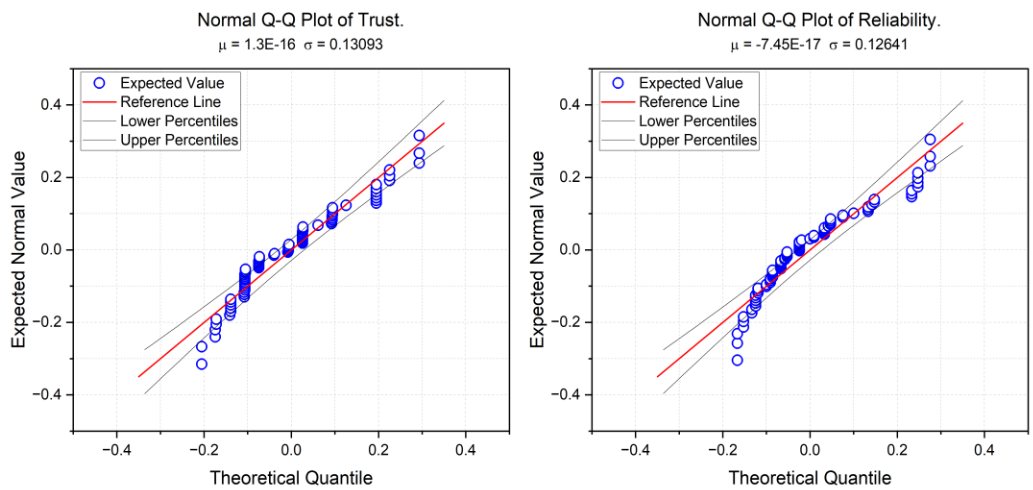

Figure A2 - Normality check via Q-Q plots: trust (left) and reliability (right).

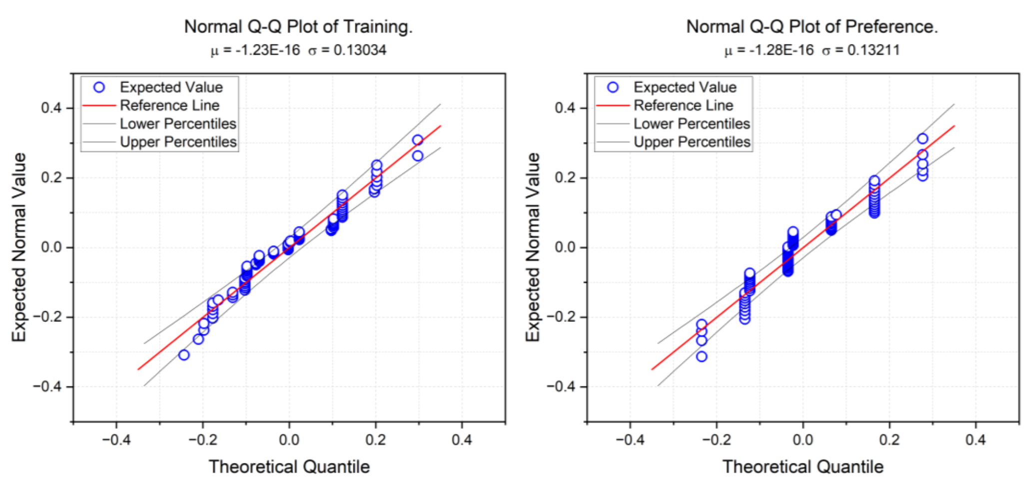

Figure A3 - Normality check via Q-Q plots: training (left) and preference (right).

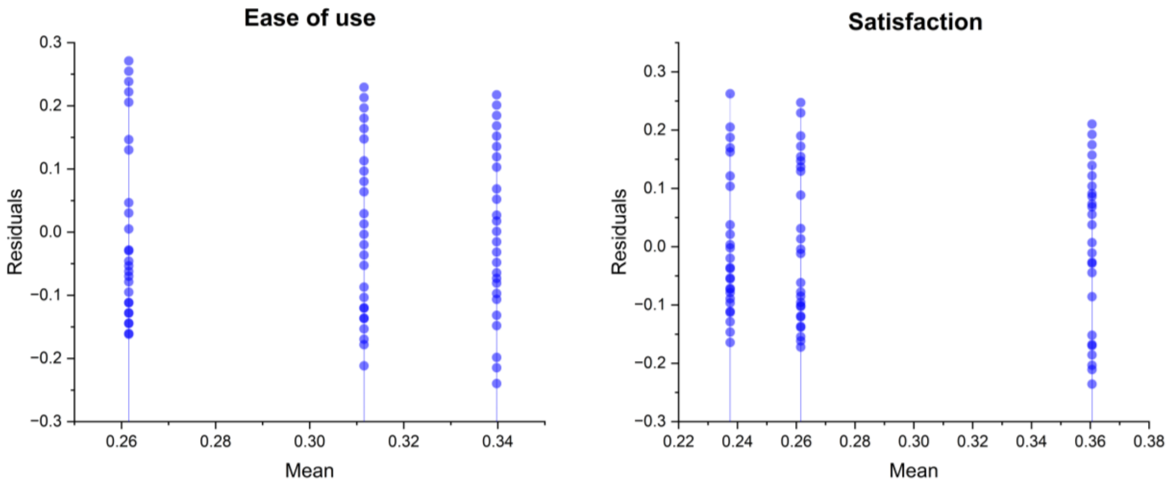

Figure A4 - Constant variance check: ease of use (left) and satisfaction (right).

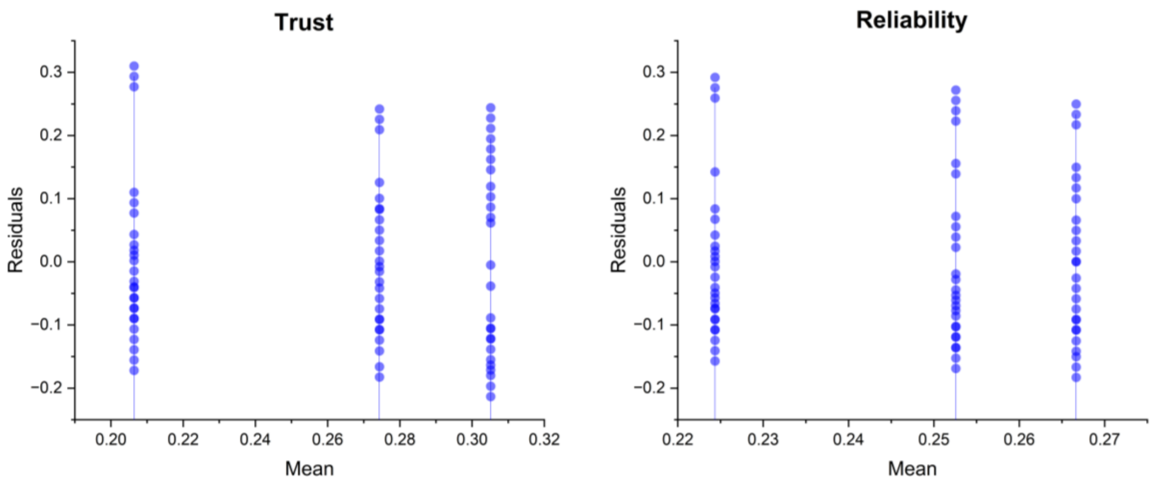

Figure A5 - Constant variance check: trust (left) and reliability (right).

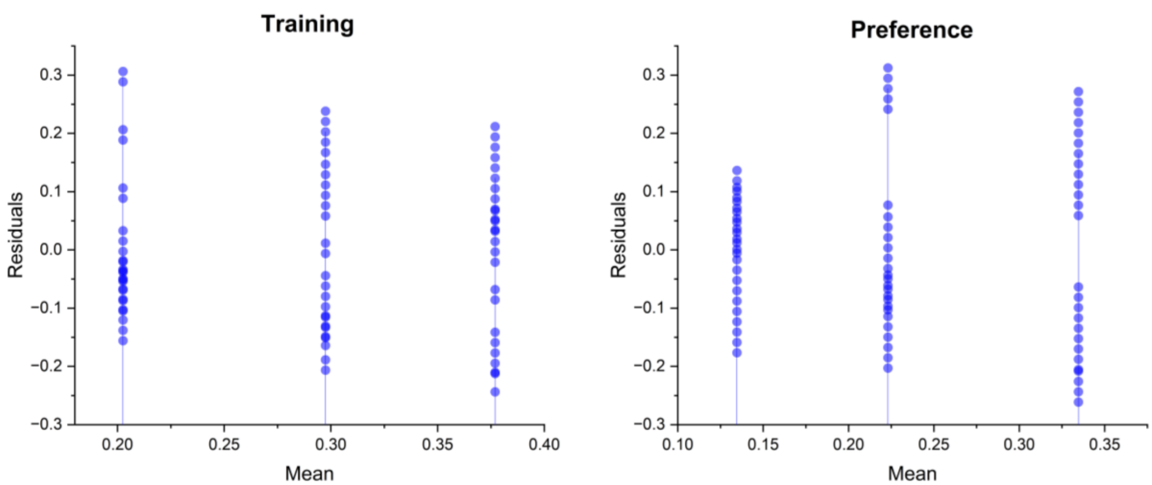

Figure A6 - Constant variance check: training (left) and preference (right).

Table A1 - ANOVA with blocks results for ease of use.

Source of Variation	***Sum of Squares***	***df***	***Mean Square***	***F-value***	***P-value***	***F critical***
Training Platform	0.082	2	0.041	1.904	0.160	3.183
Users	0.299	25	0.012	0.559	0.942	1.727
Error	1.071	50	0.021			
Total	1.452	77				

Table A2 - ANOVA with blocks results for satisfaction.

Source of Variation	***Sum of Squares***	***df***	***Mean Square***	***F-value***	***P-value***	***F critical***
Rows	0.221	2	0.111	9.247	0.0004	3.183
Columns	0.668	25	0.027	2.231	0.0079	1.727
Error	0.598	50	0.012			
Total	1.487	77				

Table A3 - ANOVA with blocks results for trust.

Source of Variation	***Sum of Squares***	***df***	***Mean Square***	***F-value***	***P-value***	***F critical***
Rows	0.133	2	0.066	6.102	0.004	3.183
Columns	0.776	25	0.031	2.857	0.001	1.727
Error	0.544	50	0.011			
Total	1.453	77				

Table A4 - ANOVA with blocks results for reliability.

Source of Variation	***Sum of Squares***	***df***	***Mean Square***	***F-value***	***P-value***	***F critical***
Rows	0.024	2	0.012	0.797	0.457	3.183
Columns	0.473	25	0.019	1.250	0.247	1.727
Error	0.757	50	0.015			
Total	1.255	77				

Table A5 - ANOVA with blocks results for training.

Source of Variation	***Sum of Squares***	***df***	***Mean Square***	***F-value***	***P-value***	***F critical***
Rows	0.396	2	0.198	12.521	0.00004	3.183
Columns	0.517	25	0.021	1.307	0.207	1.727
Error	0.791	50	0.016			
Total	1.704	77				

Table A6 - ANOVA with blocks results for preference.

Source of Variation	***Sum of Squares***	***df***	***Mean Square***	***F-value***	***P-value***	***F critical***
Rows	0.522	2	0.261	13.356	0.000	3.183
Columns	0.366	25	0.015	0.749	0.781	1.727
Error	0.978	50	0.020			
Total	1.866	77				