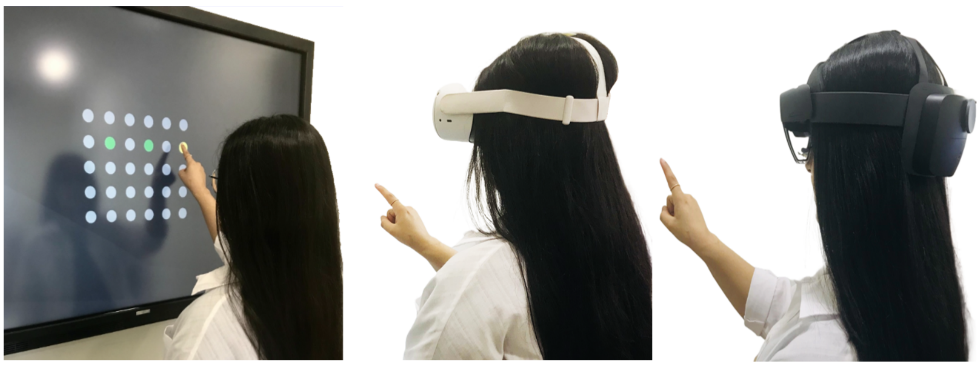
Eye-Hand Coordination Training: A Systematic Comparison of 2D, VR, and AR Display Technologies and Task Instructions

Aliza Aliza* Kadir Has University
Irene Zaugg† Colorado State University
Elif Çelik‡ Kadir Has University
Wolfgang Stuerzlinger§ Simon Fraser University
Francisco Raul Ortega¶ Colorado State University
Anil Ufuk Batmaz ‖ Concordia University
Mine Sarac ** Kadir Has University



(a) (b) (c)

Figure 1: Eye-Hand Coordination Training system with (a) 2-Dimensional Touchscreen Display, (b) Virtual Reality Head-Mounted Display (VR HMD), and (c) Augmented Reality (AR) HMD.

Abstract

Previous studies on Eye-Hand Coordination Training (EHCT) focused on the comparison of user motor performance across different hardware with cross-sectional studies. In this paper, we compare user motor performance with an EHCT setup in Augmented Reality (AR), Virtual Reality (VR), and on a 2D touchscreen display in a longitudinal study. Through a ten-day user study, we thoroughly analyzed the motor performance of twenty participants with five task instructions focusing on speed, error rate, accuracy, precision, and none. As a novel evaluation criterion, we also analyzed the participants' performance in terms of effective throughput. The results showed that each task instruction has a different effect on one or more psychomotor characteristics of the trainee, which highlights the importance of personalized training programs. Regarding different display technologies, the majority of participants could see more improvement in VR than in 2D or AR. We also identified that effective throughput is a good candidate for monitoring overall motor performance progress in EHCT systems.

Index Terms: Human-centered computing—Human Computer Interaction (HCI); Human-centered computing—Virtual Reality; Human-centered computing—Pointing; Human-centered computing—Touch screens

*e-mail: aliza.rind@stu.khas.edu.tr
†e-mail: irene.zaugg@colostate.edu
‡e-mail: elifcelik@stu.khas.edu.tr
§e-mail: w.s@sfu.ca
¶e-mail: fortega@colostate.edu
‖e-mail: ufuk.batmaz@concordia.ca
**e-mail: mine.sarac@khas.edu.tr

1 Introduction

In recent years, the affordability and accessibility of current Virtual Reality (VR) systems has revolutionized training methods in many different fields. These VR systems create immersion into lifelike virtual environments that offer a promising avenue for efficient skill development and improvement. Typically, these experiences are facilitated through Head-Mounted Displays (HMDs), providing users with an immersive, interactive environment and clear benefits over conventional techniques in various applications [65]. In addition, Augmented Reality (AR) blends the tangible and virtual components with the real environment to enrich users' immediate surroundings with contextually pertinent information.

Virtual training environments offer today a more effective approach to learning and skill development, as VR HMDs enable trainees to track their progress and pinpoint areas that require improvement [36] while repeatedly engaging with specific, controlled scenarios [33, 47]. Such repetitive training promotes deliberate practice and skill refinement [31]. These virtual systems closely replicate real-world situations, providing trainees with a sense of physical presence and allowing them to engage in training scenarios [2]. This enhanced immersion increases the engagement of the trainees while enhancing information retention [30]. Finally, virtual training is cost-efficient in the long run, as it reduces the need for physical equipment, materials, maintenance, and travel expenses associated with traditional training methods [59]. This approach can also accommodate larger numbers of trainees simultaneously, ensuring scalability and cost savings.

More specifically, in sports training, the use of VR and AR HMDs has had a considerable influence on performance evaluation throughout the years [1, 17, 29, 34, 39]. Eye-hand coordination training systems (EHCTSs) are one of the influential approaches for improving a user's reaction time. Beyond systems mounted on real 2-dimensional (2D) surfaces (e.g., a wall) [22, 55], EHCTSs based on 2D [28, 54] or 3D displays [13, 15, 16] have expanded the horizons of athlete training across various sports in terms of sensory-motor abilities and perception. They encourage athletes to respond to dynamically displayed visual stimuli with precise hand gestures that closely mimic real-game conditions. This approach has proven highly effective in enhancing athletic performance, providing athletes with a valuable tool for skill improvement [28,54]. The commercial NIKE's SPARQ sensory station [51] is the most common EHCTS based on a touchscreen display where trainees are instructed to touch a sequence of randomly activated targets as fast and accurately as possible — ultimately improving their reaction time and accuracy [61].

VR HMDs have been shown to be a potential alternative to conventional EHCTSs using physical objects and/or interactive 2D displays [13, 15]. Although there is a single longitudinal study on EHCTs [48], it used only 3 participants with a VR headset and did not comprehensively investigate motor learning. Moreover, previous EHCTS research with AR HMDs did not demonstrate good performance – and only evaluated the experience for a single day [13]. This could be due to the fact that AR technology is still unfamiliar to most participants, and that its true potential might thus not emerge after using it for only one day. Consequently, a longitudinal AR HMD EHCT study is needed to demonstrate whether AR HMDs can be used in EHCT systems and if they demonstrate benefits. Finally, instructing participants to focus on different strategies while performing EHCT yields different performance and motor learning outcomes [48, 52], but a systematic comparison is still missing. Our novel comparative analysis between a standard 2D touchscreen display, and a VR and AR HMD also adds valuable insights to the existing body of knowledge in the field.

This paper presents an extensive and systematic evaluation of the impact of different display technologies and task instructions for EHCTs based on a longitudinal user study. By tracking user performance over time we identify long-term trends in the user experience, learning curves, and adaption patterns. This design aligns with real-world usage scenarios, which makes our findings generally applicable. First, examining different task instructions allows trainers to identify which specific training methods or techniques are most effective in improving eye-hand coordination. This information can help refine training programs, making them more efficient and enabling targeting to the current needs of trainees. We focus on five specific task instructions – completing tasks with the fewest errors, the shortest time, the highest accuracy, the highest precision, or no specific strategy to focus on. Second of all, the level of immersion and/or familiarity with the display technology might change the task performance or the motor learning curve for the trainees. Therefore, we focus on three display technologies: 2D touchscreen displays, VR HMDs, and AR HMDs. Finally, over time, individuals can adapt to training strategies or develop new strategies on their own. Longitudinal studies can capture these adaptations and help researchers understand how individuals maintain and adapt their eye-hand coordination skills over the long term. In this paper, we designed an EHCT user study with a training process that consists of 10 days of sessions together with a pre-training and post-training assessment. This training approach can be used in different sports that require eye-hand coordination, such as tennis, baseball, basketball, cricket, and American football.

2 Previous Work

2.1 EHCTSs with 2D Touchscreen, VR, and AR Displays

EHCTSs challenge trainees' perceptual abilities and necessitate rapid and precise responses to specific stimuli. This form of training has the potential to improve cognitive processing speed, visual perception, and motor coordination. Through repeated participation in such tasks, trainees can improve functioning with their minds and bodies to respond with greater efficiency and precision. EHCTSs can be extended to sports training [20, 28, 31, 35, 50, 54], military training [18], and post-injury rehabilitation [25].

Moreover, EHCTSs help trainees react faster to visual stimuli, directly affecting the performance of trainees during sports and game outcomes [28]. Several companies, such as Batak [55], developed an eye-hand coordination training system with wall-mounted hardware. EHCT can also be done on a 2D touchscreen display [45], which is favored by trainers due to their ability to gather precise motor performance data and provide immediate feedback. This facilitates real-time adjustments in learning techniques and reduces the trainer's time commitment. Previous studies have also investigated the skill transfer and its efficacy as a response to external factors such as daily variations (e.g., sleepiness [61]) or technology (e.g., 2D displays [28, 35, 54, 61]).

VR HMDs have been shown to improve the user performance and decision-making of participants during EHCT tasks [40]. For example, a VR-based EHCTS used a grid layout of potential targets in a virtual environment, with only one target being highlighted at any time [15], similar to the Nike SPARQ sensory station [51]. Participants selected targets with a virtual cursor attached to their palms while trying to rapidly and accurately select each sequentially highlighted target by hitting their palms on a wall. Their results illustrated that VR HMD-based EHCTSs possess the ability to replace existing 2D setups. Later on, the authors followed up by extending their work to AR HMDs and by instructing participants to select targets using a virtual cursor attached to their index fingertip [13]. In terms of time and error rate, VR HMDs and 2D touchscreen displays revealed similar results, while AR HMDs fell significantly short in terms of user performance. Finally, the optimal performance in VR-based EHCT was obtained when users interacted with mid-air objects with increased target size, utilized their dominant hand, and engaged in a vertically oriented task plane [16].

Another research direction regarding EHCTSs is based on the dimension of the grid layout in which the potential targets are presented to the user. In earlier VR-based EHCTSs, a 6x6 grid layout had been adopted to replicate previous 2D touchscreen display setups [13, 15, 48]. However, investigating the gaze behavior during EHCT tasks revealed that many users incurred significant neck strain to see the overall targets in the 6x6 layout while using VR HMDs [49], due to the limited field-of-view of such devices. This result emphasized the need for EHCTS designs that maximize the effectiveness of the gaze movements.

2.2 Performance assessment in EHCTSs

In the context of training, a central research question is how to evaluate the performance of trainees and provide feedback [56]. In EHCT, and to improve feedback quality and refine training strategies, the trainer should take different assessment criteria simultaneously into account (such as combinations of accuracy, precision, error rate, or time) [7, 11, 17, 21, 62]. Previously, precision has been highlighted as a primary assessment criterion for tasks involving eye-hand coordination [9]: participants exhibit swifter task execution and enhanced stimulus perception when emphasizing precision rather than speed or concentration on feedback. The authors claimed that adopting a precision-oriented approach can accelerate sensorimotor learning and optimize skill acquisition, facilitating seamless task adaptation [10].

On the other hand, participants might get more benefit from the training process if they focus on certain assessment criteria to enhance the corresponding performance [48]. Particularly, focusing on task speed could help trainees improve their task completion time, albeit at the potential cost of (at least temporarily) increasing their error rate. This method not only results in notable improvements, but also recognizes the essential role of a broader development of psychomotor skills for eye-hand coordination tasks [20, 31].

2.3 Fitts' Law

Fitts' Law [24] encapsulates the dynamics of the complete human receptor-neural-effector system during pointing tasks. Derived from information theory, MacKenzie's Shannon formula (in Equation 1) [41] stands as a widely employed approach within human-computer interaction research. It embodies one of the most frequently utilized methods in this field.

$$MovementTime = a + b * \log_2\left(\frac{A}{W} + 1\right) = a + b * ID \qquad (1)$$

This model employs movement amplitude A and target size W, where constants a and b are derived through linear regression. The *Index of Difficulty ID*, determined by the logarithmic term, signifies task complexity. Later on, a prominent rendition of this index was put forth by MacKenzie [41], drawing inspiration from the Shannon capacity theorem. This formulation has evolved into what is now recognized as throughput based on effective measures, commonly known as effective throughput. As per ISO 9241-411:2012, throughput denotes the "rate of information transfer when a user is operating an input device to control a pointer on a display" [32]. In alignment with this, we computed throughput accordingly.

$$Throughput = \left(\frac{ID_e}{MovementTime}\right) \qquad (2)$$

Equation 2 delineates movement time as the interval spanning movement initiation and target selection. The *effective* Index of Difficulty ID_e is characterized as the "measure of user precision attained in task accomplishment" [32] as follows:

$$ID_e = \log_2\left(\frac{A_e}{W_e} + 1\right) \qquad (3)$$

Equation 3 introduces the effective distance A_e as the actual movement distance to the target while accounting for participants' task performance accuracy. We denote the effective target width, derived from the distribution of selection coordinates – calculated by $W_e = 4.133 \times SD_x$, where SD_x represents the standard deviation of selection coordinates along the task axis — as the task performance precision [32,42,43]. Fitts' Law in (Equation 1) stands as the foremost model for characterizing a user's pointing performance. It establishes a link between the user's movement time and factors like target size and distance between targets.

The term "throughput" plays an essential role as an assessment criterion in VR systems and training research to combine task execution time, accuracy, and precision into a single measure [58]. This comprehensive measure supports a quick assessment of trainee development. Previous studies on VR-based EHCTSs [13,15,16] have shown the value of throughput in identifying trends like decreased precision among participants due to difficulties with stereo display technologies [5,6,8]. Although MacKenzie and Isokoski [42] argued for throughput's speed-accuracy invariance, current research indicates that task execution strategies may actually change it [48,52].

Previous work also investigated the users' effective throughput performance with longitudinal studies. Boritz et al. conducted a user study that spanned four sessions [19]. Mughrabi et al. conducted a ten-day training session but did not explore different assessment criteria [48]. After all, using *only* effective throughput might decrease the ability to monitor motor abilities. Consequently, to best support trainees, it is crucial to keep track of each different performance criterion, such as time, accuracy, or precision, and draw conclusions only based on a more thorough analysis.

Hence our study employed effective throughput to evaluate user performance. Additionally, ISO 9241-411 defines task precision as the "measure of the user precision achieved in accomplishing a task" [32], denoted as *ID*. In this context, the term "precision" refers to the relative closeness of selected points to each other. In our study, we use ISO 9241-411's equations to assess participants' accuracy SD_x and precision ID_e.

3 Motivation and Research Questions

While the previous work on EHCTSs and its performance assessment techniques indicate its potential benefits for future use, there are many other gaps that need to be addressed. In this paper, we are motivated to address these gaps partially by investigating the impact of the training period and the task execution instructions and display technologies used in EHCTSs:

Training Period: Previous EHCTSs have mostly been tested through preliminary studies to investigate how human performance and decision-making are affected by different display or task execution strategies [13, 15, 16]. There is one study focusing on the impact of EHCTSs on motor learning in a systematic, longitudinal study; but, instead of an extended user study, they present only the results of a preliminary study with 3 participants with only VR headset [48]. Similarly, the previous studies on effective throughput go up to 4 days of longitudinal studies [19], and we did not find a comprehensive study for longitudinal effective throughput. Thus, the literature lacks a longitudinal study that analyzes the long-term performance impact of HMD-based VR EHCTSs for training and effective throughput. *In this paper,* we use a 10-day experimental training protocol, where the participants' performance was evaluated with pre-training and post-training sessions on day-0 and day-11. Such a longitudinal user study allows us to observe changes in behavior and skill acquisition over different phases of the learning process. It also enables a more comprehensive understanding of how performance evolves. Thus, and following a previous VR training study [46], we set the longitudinal study duration to 10 days, which also increases the comparability of our results.

Task Execution Instructions: Previously, a VR-based EHCTS has been tested on participants focusing on speed, accuracy, or without a specific task instruction [48]. However, this study did not investigate some other important task instructions (i.e., error rate and precision), which might be essential to improve the overall progress and task performance of trainees in future uses of EHCTSs. This is particularly important since *precision* had been highlighted as the primary assessment criterion for tasks involving eye-hand coordination in other work [9]. *In this paper,* we instruct the participants to perform the EHCT tasks focusing on one of the five different task strategies: precision and error in addition to the previously used strategies of speed, accuracy, or without a specific task instruction.

Display Technologies: EHCTSs have been previously implemented on 2D touchscreen displays, in VR HMDs, and AR HMDs to investigate whether using different display technologies would reveal differences in user performance. 2D displays and VR HMDs have been reported to yield similar performance, while AR HMDs were observed to be significantly worse than the first two [13]. Yet, how these displays affect motor learning at the end of a longitudinal training process is still unknown. *In this paper,* each participant was asked to complete the training tasks while they were displayed on a 2D touchscreen display or in VR and AR HMDs.

In summary, our main motivation is to investigate the impact of different display technologies and different task execution instructions for participants on different evaluation metrics as an indication of user performance at the end of a 10-day training process. At the end of our study, we aim to answer the following research questions:

- **RQ1:** Do different task execution instructions yield different trends in different aspects of the user performance metrics? If so, does "precision" offer the most optimal training procedure, as previously hypothesized [9]?

- **RQ2:** Is "effective throughput" a reliable assessment criterion to observe the overall performance of trainees with different task instruction strategies over a 10-day training process? If so, does it contribute to the speed/accuracy trade-off discussions?

- **RQ3:** Can VR HMDs be used as an effective display technology for virtual EHCTSs?

4 EYE-HAND COORDINATION TRAINING EXPERIMENT

We created an EHCTS in Unity, shown in Fig. 2, following a previously presented design [48]. Participants see a 5x6 grid of gray spherical virtual buttons with 6 cm distance between their centers, which allows them to see and reach every target while reducing the risk of neck strain [49]. Buttons were identical in terms of color, size, and shape across all display technologies. At any given time, one of these buttons is chosen as a target sphere and highlighted in yellow. We instructed participants to select this target sphere by "clicking" on it with their index finger. To enhance the virtual interaction and overcome the lack of haptics, our study strategically incorporated different visual cues for the participants during different stages of the selection, following previous work by Batmaz et al. [15]. This previous work highlighted the issues related to haptic feedback, as that addition increased the execution time and lowered throughput value when participants engaged with solid objects during virtual target selections.

(i) When the virtual cursor associated with the finger is inside a sphere, it is highlighted in blue. Moving the finger closer to the target center after seeing the blue cue results in the sphere selection. *(ii)* If the participant chooses the target correctly, the selection is labeled as a "hit", and the selected sphere becomes green. *(iii)* If the participant chooses a different sphere that is not the target, the selection is labeled as a "miss", the selected sphere becomes red, and the system generates a beep sound. The green and red coloring of the spheres is maintained during the next two selections, but then revert back to gray.

Once a selection has been made, the next target is automatically assigned – following a straight path with one of two step sizes: North (N2 and N4), West (W2 and W4), South (S2 and S4), and East (E2 and E4), or a diagonal path with a single step size: North-west (NW3), South-west (SW3), South-east (SE3), and North-east (NE3). The straight directions correspond to target distances of 12 cm and 24 cm, while the diagonals correspond to 25.45 cm. The system randomly selects the next target among these 12 options under the constraint that the target has not previously been used in the current round of trials. Once there are no more unused targets that can be reached within the constraints, the round of trials stops. On average, each round of trials thus involved the selection of 20 targets. Each participant took up to 10 minutes to complete the experiment in each display technology, and each participant completed a total of 180 trials for each display technology. As a result, each participant contributed a total of 540 trials in a single day across all three display technologies and overall 6480 trials over the 12 days. The order of the displays was counterbalanced across participants. Assuming a large effect size (η^2=0.14, f(U)=0.4035), we performed an a-priori power analysis to make sure that our sample size was appropriate. With a selected sample size of 20, the analysis revealed a high level of statistical power (0.977). This comprehensive validation supports the robustness of our experiment design.

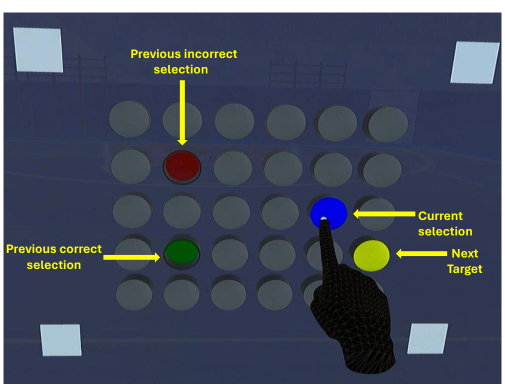


Figure 2: An image illustrating the virtual task. Participants interacted with a 5x6 grid of target spheres, initially all grey. The current target is shown in yellow. When the index finger was in contact with or inside a sphere, it was highlighted in blue – indicating a possible future selection. The previous two selections were also shown in green if successful and in red if missed.

While a Fitts' Law task itself may be considered straightforward, our study systematically explores different task instructions and different visualization environments (2D touch display, VR HMD, AR HMD). Also, longitudinal user studies benefit from collecting much more data from the same participants over a longer time, which means that statistically valid outcomes can be expected from a sample size smaller than with typical short-term studies.

4.1 Display Technologies

During the experiment, we showed the EHCT environment to the participants through various platforms, including a regular 2D touchscreen display, VR HMDs (e.g., Oculus Quest), and AR HMDs (e.g., Microsoft HoloLens), as a within-subjects factor (i.e., all participants experience all display technologies). We detailed the technical aspects of each display system in the following.

4.1.1 2D Touchscreen Display

We used an Android 2D display with x64-based processor and 8GB RAM, running at 3840x2160 resolution. Participants interacted with the 2D display only using their dominant index finger, as shown in Fig. 1 (a). During the experiment, the experimenter observed how they interacted with the screen and gave immediate feedback if they deviated from these instructions. In the software, we recorded the exact point they touch on the 2D display in real time. Participants adjusted the height of the target grid to their eye level on the first day, and its height was kept constant for the same participant throughout the rest of the training.

4.1.2 VR HMD

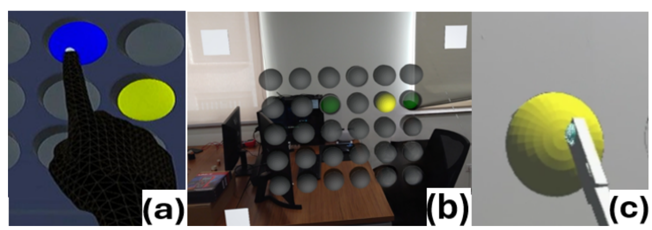


Figure 3: (a)EHCTS with a VR HMD: The cursor attached to the index finger of their avatar (hand) allows participants to select the spheres. (b) EHCTS with an AR HMD: The augmented scene allows participants to visualize the virtual content displayed overlaid over the real world. During the experiment, participants faced an empty wall to improve the quality of hand tracking and their attention span. (c) The cursor attached to the index finger of their avatar allows participants to select the spheres.

We chose an Oculus Quest 2 with a resolution of 1920 x 1832 pixels and a diagonal field-of-view of 110°. Participants interacted with the grid using only their dominant index finger, as shown in Fig. 1 (b). The hand tracking software integrated in the headset allowed participants' hand movements to be tracked and visualized in the virtual environment in real time, so that they could complete the eye-hand coordination task presented in Fig. 2. A white cursor attached to their index fingertip allows them to interact with the virtual targets (Fig. 3 (c)). We adopted a similar GUI to adapt the virtual environment to the participants' eye level and to set the experiment parameters before the experiment.

4.1.3 AR HMD

We used a HoloLens 2, with a resolution of 2K and a field-of-view of 52°. Participants interacted with the grid using only their dominant index finger, as shown in Fig. 1 (c). The HoloLens 2 features a transparent display system that allows users to see the real world while overlaying digital content over it, while also providing an integrated hand tracking system. Fig. 3 (b) shows how the EHCT grid appeared within the participant's view of the real world. However, to achieve the highest performance with hand tracking and to increase the user's ability to focus on the experiment task, we asked participants to face an empty wall during the experiment.

During the EHCT task, participants can see a series of virtual links representing the structure of the hand overlaid on top of their actual hands. At the index fingertip of their virtual dominant hand, we show a white cursor to allow participants to interact with the virtual spheres as in Fig. 3 (c). HoloLens 2 automatically renders the visual display at their eye level, so there is no need for manual adjustments. For this system, we recorded information such as experimental conditions and participant ID directly in Unity.

4.2 Task Instructions

We chose the task instructions (i.e., execution strategies) to be a between-subject factor, where each participant was randomly assigned to a group receiving a single, specific task instruction:

Speed-focus participants were instructed to select the targets "as fast as possible" with a reasonable error rate. They were reminded to prioritize their movement speed over other performance metrics.

Error-focus participants were instructed to select the targets "with the fewest errors possible" at a reasonable speed. They were reminded to prioritize the number of errors they made over other performance metrics.

Accuracy-focus participants were instructed to select the targets "as close to the center as possible" (i.e., their midpoint) at a reasonable speed. They were reminded to prioritize choosing the targets' center over other performance metrics.

Precision-focus participants were instructed to select each target "consistently at the same point relative to its center" at a reasonable speed. The exact location of the selected point did not matter. They were reminded to prioritize choosing targets at the same point consistently over other performance metrics.

No-focus participants were only instructed to complete the tasks with no specifics, except that they should use a reasonable speed and not make too many errors.

The experimenter observed the participants to remind them about the assigned task instructions as needed, except for the no-focus participants. To ensure that the participants followed instructions, the authors monitored their results and provided daily feedback.

4.3 Participants

20 participants (8 females and 12 males) were recruited from the local university, with ages ranging between 19 and 33 years (M = 21, SD = 4.6). The local Institutional Review Board approved the experimental protocol, and all participants gave informed consent. By enrolling in the experiment, they agreed to participate in the experiment for 12 consecutive days (1 day for pre-training assessment, 10 days for training, and 1 day for post-training assessment). We assigned 4 participants to each task instruction randomly. Upon arrival for the first session, they completed a questionnaire (see supplementary material) about their demographic information and hand dominance. Sixteen participants were right-handed (four left-handed).We also collected information about their prior experience with *(i)* 2D touchscreen displays (four participants had used it 5+ times, ten 1-3 times, and six had never used it), *(ii)* VR HMDs (five participants had used it 5+ times, one 3-5 times, nine 1-3 times, and five had never used it), and *(iii)* AR HMDs (one participant had used it 5+ times, five 1-3 times, and 14 had never used it before).

5 Experiment Results

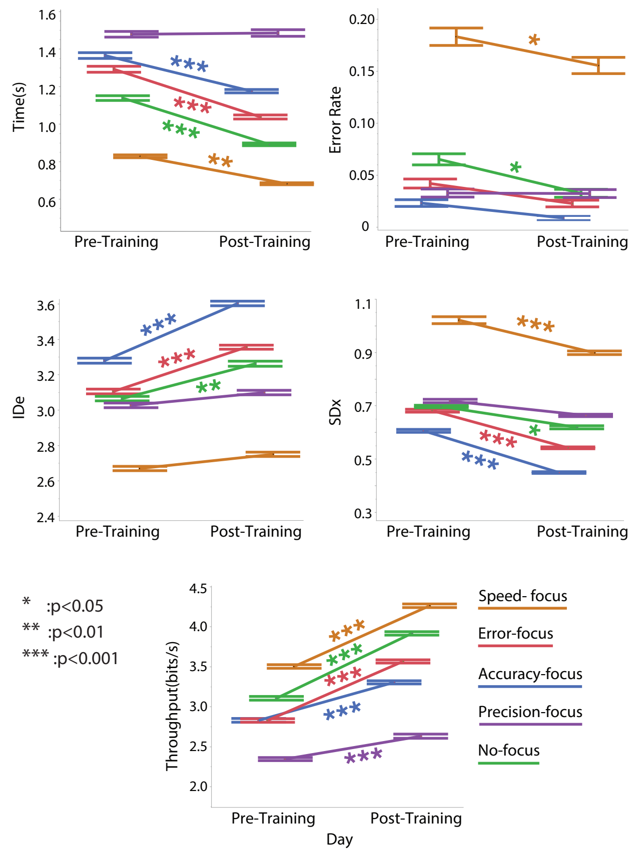


Figure 4: Pre- and Post-training results with different task instructions for execution time, error rate, IDe (precision), SDx (accuracy), and throughput. * indicates a significant difference with $p < 0.05$, ** with $p < 0.01$, and *** with $p < 0.001$.

We evaluated the motor performance of participants using execution time, error rate, SDx (i.e., accuracy of target selection), IDe (i.e., precision of target selection), and throughput and analyzed them using SPSS 24 and JMP. When the Skewness (S) and Kurtosis (K) of the data distribution were within ±1.5 [27, 44], we considered the data to have a normal distribution. Otherwise, we used log-transform. We performed a 3-factor ANOVA with two within-subject factors (displays and training days) and one between-subject factor (task instruction) for each assessment criterion.

5.1 Pre-Training & Post-Training Results

Table 1 and Fig. 4 provide a summary of the main factors and their interactions only for the pre-training and post-training evaluation. The supplementary material illustrates the data distributions.

The analysis revealed that all participant groups except the precision-focus group exhibited a significantly shorter execution time at post-training compared to the pre-training assessment. Only the speed-focus and no-focus participant groups exhibited significantly fewer errors at post-training compared to pre-training assessment. Accuracy-focus, error-focus, and no-focus participant groups exhibited a significantly higher *IDe* (better precision) at post-training compared to pre-training assessment. All participant groups except the precision-focus group exhibited a significantly higher *SDx* (better

Table 1: ANOVA results of the main experiment participants between pre- and post-training. Statistically significant results are in bold.

	Day	Display	Task instruction
Time	**F(1,15) = 6.299, $p < 0.05, \eta^2 = 0.296$**	**F(2,30) = 36.310, $p < 0.001, \eta^2 = 0.708$**	**F(4,15) = 5.698, $p < 0.05, \eta^2 = 0.603$**
Error Rate	$F(1,15) = 2.517$ $p > 0.05, \eta^2 = 0.144$	$F(2,30) = 1.935,$ $p > 0.05, \eta^2 = 0.114$	**F(4,15) = 7.298, $p < 0.05, \eta^2 = 0.661$**
IDe	**F(1,15) = 14.187, $p < 0.05, \eta^2 = 0.486$**	**F(2,30) = 71.933, $p < 0.001, \eta^2 = 0.827$**	**F(4,15) = 12.309, $p < 0.001, \eta^2 = 0.766$**
SDx	**F(1,15) = 8.150, $p < 0.05, \eta^2 = 0.351$**	**F(2,30) = 36.784, $p < 0.001, \eta^2 = 0.710$**	**F(4,15) = 14.835, $p < 0.001, \eta^2 = 0.798$**
Throughput	**F(1,15) = 26.670, $p < 0.001, \eta^2 = 0.640$**	**F(2,30) = 43.434, $p < 0.001, \eta^2 = 0.743$**	**F(4,15) = 6.542, $p < 0.05, \eta^2 = 0.636$**

accuracy) at post-training compared to pre-training assessment. All participant groups exhibited a significantly higher throughput (better overall task performance) at post-training compared to pre-training assessment.

In terms of different display technologies, Fig. 5 shows that all display technologies are statistically significantly different from each other for execution time and error rate – with the AR HMD being significantly worse than the other two for both metrics. Participants exhibited the shortest execution time using 2D displays and the lowest error rate using VR HMD. Participants performed the tasks with significantly higher precision (IDe) and accuracy (SDx) using 2D displays compared to the other two, but we observed no significance between VR HMD and AR HMD. Finally, all display technologies were found to be statistically significantly different from each other in terms of throughput – 2D displays being statistically the highest and AR HMDs the lowest.

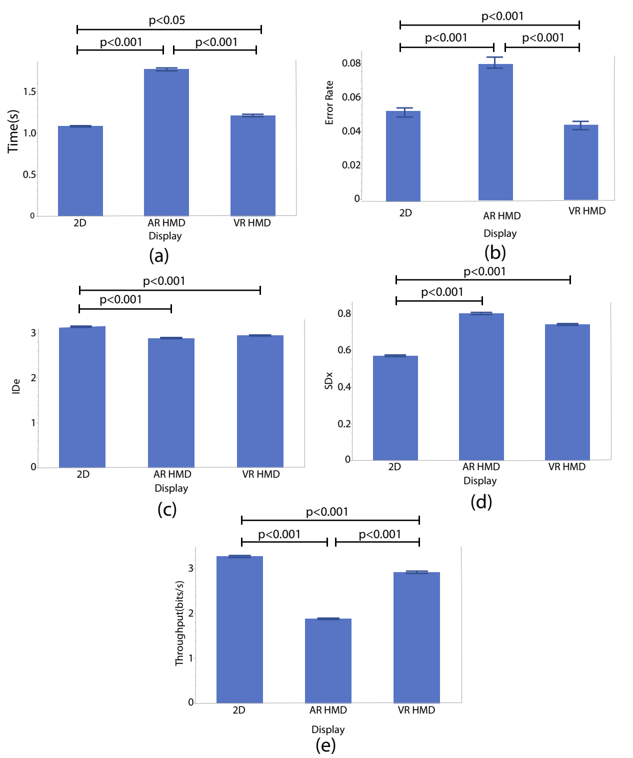


Figure 5: Statistically significant results of display condition for average of pre- and post-training data for (a) execution time, (b) error rate, (c) IDe, (d) SDx, and (e) throughput.

5.2 Longitudinal Training Results

Table 2 details the results of the main effect values of the longitudinal study while their interactions, the detailed post-hoc analyses, as well as the growth rate results, are provided in the supplementary material. Since Table 2 shows that participants' performances varied significantly across training days, we performed a post-hoc analysis to investigate their motor learning in more detail. Fig. 6 shows the longitudinal results to select each target for all target distances, target sizes, and display types for each task instruction. Data points on each day represent the mean, and the error bars represent the standard error of the mean. The supplementary material illustrates the data distributions.

Time: The post-hoc analysis of time across training days indicated that participants' target selection times improved significantly in the first 5 days, but starting with the 6^{th} day, their performance did not vary significantly, indicating a plateau effect. In addition, further post-hoc analysis indicates that participants selected the targets statistically significantly slower with the AR HMD than with both the 2D display and the VR HMD, but there was no significant difference between the 2D display and VR HMD. Finally, we observe a steeper learning curve with VR and AR HMDs compared to the 2D display.

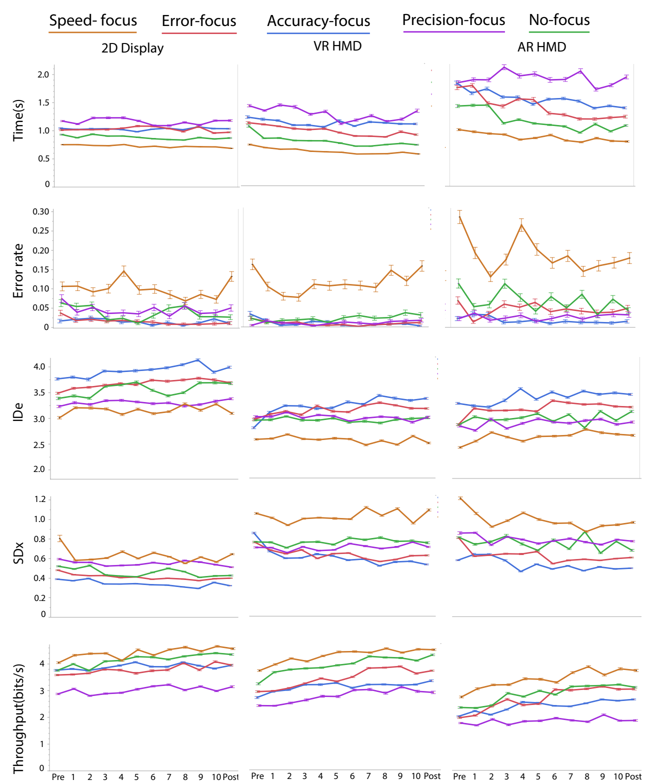


Figure 6: Longitudinal study results for 3 displays and 5 task instructions for execution time, error rate, IDe, SDx, and throughput.

Table 2: ANOVA results of longitudinal training days. Statistically significant results are in bold.

	Day	Display	Task instruction
Time	**F(11,165) = 4.795, $p < 0.05, \eta^2 = 0.242$**	**F(2,30) = 37.351, $p < 0.001, \eta^2 = 0.713$**	**F(4,15) = 6.057, $p < 0.05, \eta^2 = 0.618$**
Error Rate	**F(11,165) = 3.412, $p < 0.001, \eta^2 = 0.185$**	$F(2,30) = 3.241,$ $p = 0.053, \eta^2 = 0.178$	**F(4,15) = 6.112, $p < 0.05, \eta^2 = 0.620$**
IDe	**F(11,165) = 5.149, $p < 0.001, \eta^2 = 0.256$**	**F(2,30) = 101.877, $p < 0.001, \eta^2 = 0.872$**	**F(4,15) = 12.213, $p < 0.001, \eta^2 = 0.765$**
SDx	**F(11,165) = 5.116, $p < 0.001, \eta^2 = 0.254$**	**F(2,30) = 71.983, $p < 0.001, \eta^2 = 0.828$**	**F(4,15) = 12.802, $p < 0.001, \eta^2 = 0.773$**
Throughput	**F(11,165) = 16.202, $p < 0.001, \eta^2 = 0.519$**	**F(2,30) = 58.701, $p < 0.001, \eta^2 = 0.796$**	**F(4,15) = 7.572, $p < 0.05, \eta^2 = 0.669$**

Error Rate: Post-hoc analysis for error rate across training days indicated that participants' error rate seemed to vary systematically in the first and fourth day, but then remained statistically non-significant. Further post-hoc analysis indicates that participants made significantly more errors with the AR HMD than with both the 2D display and the VR HMD, and with the 2D display than with the VR HMD.

IDe: Post-hoc analysis for IDe across training days indicated that participants' *IDe* score (i.e., precision in selecting the targets) improved systematically for the first two days and then again on the 9^{th} day, while the other days showed no significant difference. Regardless, the change on day 9 seems not to have affected the ultimate performance of the users, so its longitudinal impact is somewhat questionable. Further post-hoc analysis indicates that participants selected the targets with significantly higher precision on a 2D display compared to the VR and AR HMDs, which were significantly different. Finally, we observe a more consistent learning curve for VR and AR HMDs compared to the 2D display.

SDx: Post-hoc analysis for SDx across training days indicated that participants' target selection seemed to vary systematically in the first 2 days, but then remained statistically non-significant. Further post-hoc analysis indicates that participants selected the targets significantly more accurately with a 2D display compared to VR and AR HMDs. There was a significant difference between 2D and VR and 2D and AR but no significant difference between VR and AR. Finally, we observe a steeper learning curve for VR and AR HMDs than 2D display for all groups except speed-focus.

Throughput: Post-hoc analysis for throughput across training days indicated that participants' target selection varied systematically in the first 6 days, but then their performance did not vary statistically anymore. Further post-hoc analysis indicates that participants selected the targets statistically significantly better in terms of throughput for all display types. Finally, we observe a steeper learning curve with the VR and AR HMDs than the 2D display.

5.3 Subjective Comments

After the experiment, we asked our participants to fill a questionnaire regarding their preferences on a 7-point Likert scale and user preference questions (see supplementary material). Their rankings on the user preference questions were analyzed using a Kruskal-Wallis test. The mean rank obtained from the test for the 2D display was 26.78, for the VR HMD was 36.83, and 27.90 for the AR HMD. Hence we observed that participants preferred using the VR HMD over the AR HMD and using the AR HMD over the 2D display. When asked about which display technology they preferred for EHCTS, twelve participants reported the VR HMD — claiming having perceived an increase in their task performance throughout the longitudinal training sessions. They further reported: *"it was easier to click on targets"*, *"it was more interactive"*, *"the virtual environment was good, and the virtual hands were visible clearly"*, and *"it was the smoothest experience"*. Four participants reported 2D touchscreen display – claiming *"touching the screen makes me more confident about whether I am pressing exactly the center of the circles or not"* and *"it was easy and clearer to choose the options"*. Four participants reported AR HMD – claiming *"I am not in a completely virtual environment, I can see my real hand's and fingers' movement"*, *"I feel extremely comfortable"*, and *"to perform the task was very smooth"*.

When asked about which display technology they disliked for EHCTS, seven participants reported 2D touchscreen display – claiming *"I cannot tap the exact location of the button"*, *"it was very sensitive and after pressing buttons in the air for a while, it felt odd pressing an actual surface"*, *"I got used to completing the task in AR and VR, and they became easier and enjoyable for me"*, and *"since there is no feeling of [visual] depth [with the] 2D touchscreen, I did not feel comfortable"*. Ten participants reported AR HMD – claiming *"it was more difficult to locate the yellow button because my viewport was limited to a few rows, and the contrast between the yellow button and background was not as visible as the rest of the environment"* and *"poor hand tracking made it harder to select the correct button accurately"*. Three participants reported VR HMD – claiming *"I like the view in VR, but touching the circles does not feel realistic"*, and *"there is not a feel of depth"*.

When asked about frustration, four participants reported the task as annoying, long, and repetitive; ten participants reported that it was fun; three participants reported the AR HMD as annoying, but the other two displays as fun; and three participants reported it as annoying due to small target sizes. Regarding the level of physical fatigue, ten participants reported the experiment to be physically tiring, five participants reported it not to be tiring, and four participants reported it to be tiring in the first 2-3 days only. Regarding the level of mental fatigue, six participants reported the experiment to be mentally tiring in terms of switching between different display technologies and focusing on the task instructions being (too) hard. Fourteen participants did not report suffering from mental fatigue.

5.4 Detailed Analysis for Precision-focus Task Execution

The supplementary material illustrates the data distributions. Our results indicated that the precision-focus participants made the highest number of errors and did not improve in terms of other evaluation matrices – contrary to previous work [9]. This conflict raises compelling questions about the relationship between precision-focus instructions and their task performance. Firstly, we investigated the user performance separated for each target size and target distance. Our results show that participants were not able to follow the precision-focus instructions regardless of the target size and distance (i.e., different levels of task difficulty) and still deviated from the first selection point of the targets (see supplementary material).

Since the participants could not follow the precision-focus task instructions, we assumed that language barriers could be a factor that might have played an important role. Both the experimenter and the participants of the experiment were non-native English speakers, yet the experiment instructions were presented in English. Yet, the concept of 'precision' might be more challenging to capture than other instructions. To investigate this even further, we established a collaboration with a university in an English-speaking country (Colorado State University, Colorado, USA) to conduct a replication study of the precision-focus group with native English-speaking participants with the same hardware and software setups of the experiment. This study was conducted with 3 native English speakers and the results were compared to 3 non-natives.

We conducted a comparative analysis across these two groups (Native and Non-native English speakers) to investigate the impact of language on their performance. Table 3 summarizes the results of RM ANOVA main effect values of pre-training and post-training. We found no significant differences in any of the performance measures between native and non-native speakers, but we found significant interaction results between display and main-language, as shown in Fig. 7. The results indicate that native English speakers had a higher ID_e performance compared to non-native English speakers. We present the statistics of the interactions and the longitudinal study results for this comparison in the supplementary material.

Table 3: ANOVA results for native and non-native English speakers over pre- and post-training. Statistically significant results are in bold.

	Day	Display	English Nativity
Time	$F(1,4) = 0.123$, $p > 0.05$, $\eta^2 = 0.030$	**F(2,8) = 25.856, p < 0.001, $\eta^2 = 0.866$**	$F(1,4) = 2.669$, $p > 0.05$, $\eta^2 = 0.400$
Error Rate	$F(1,4) = 2.715$, $p > 0.05$, $\eta^2 = 0.404$	$F(2,8) = 2.245$, $p > 0.05$, $\eta^2 = 0.360$	$F(1,4) = 6.454$, $p > 0.05$, $\eta^2 = 0.617$
IDe	$F(1,4) = 3.003$, $p > 0.05$, $\eta^2 = 0.429$	**F(2,8) = 8.191, p < 0.05, $\eta^2 = 0.672$**	$F(1,4) = 0.162$, $p > 0.05$, $\eta^2 = 0.039$
SDx	$F(1,4) = 4.418$, $p > 0.05$, $\eta^2 = 0.525$	**F(2,8) = 8.512, p < 0.05, $\eta^2 = 0.680$**	$F(1,4) = 0.099$, $p > 0.05$, $\eta^2 = 0.024$
Throughput	$F(1,4) = 6.166$, $p > 0.05$, $\eta^2 = 0.607$	**F(2,8) = 57.989, p < 0.001, $\eta^2 = 0.935$**	$F(1,4) = 1.656$, $p > 0.05$, $\eta^2 = 0.293$

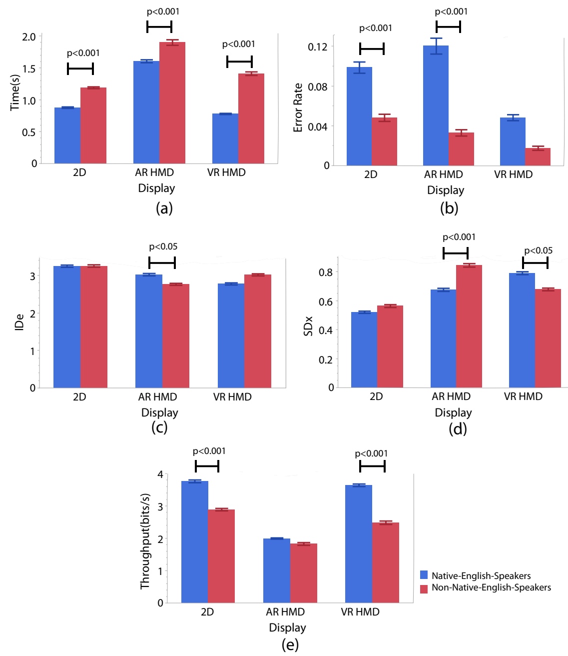


Figure 7: (a) Execution time, (b) error rate, (c) IDe, (d) SDx, and (e) throughput results for display condition across pre- and post-training data between native and non-native English speakers.

6 Discussion

In this paper, we studied *(i)* the skill transfer of EHCTSs using AR- and VR-based displays [28, 35, 54], *(ii)* participants' performance change over 10 days of training, and *(iii)* five task execution strategies instead of three with multiple participants for each strategy [48].

6.1 Task Execution Instructions

Our findings of the sub-analysis showed that after 10 days of training, each participant's motor skills improved overall, while the specific evaluation metric that was improved directly depended on the participants' task instructions, as hypothesized.

Speed-focus participants selected the targets in significantly less time than the other groups while their performance was not high in other performance metrics – as instructed. Matching the results of previous work [14], we believe that their ability to follow the instructions successfully resulted in the highest throughput results.

Error-focus participants completed the experiment with fewer errors – as instructed, but also more accurately and more precisely. Further, post-hoc analyses in error rate showed that participants performed the tasks with a very similar growth rate to accuracy-focus participants and (interestingly) to no-focus participants.We also observed no significant difference between error-focus and accuracy-focus participants in terms of precision (i.e., the consistency of choosing the same point while respecting the target limits, ID_e) and accuracy (i.e., the distance between the points selected by the participants and the center points of the targets being the smallest, SD_x) – implying that they tend to choose points around the targets' center, although not instructed so implicitly.

Accuracy-focus participants demonstrated the *highest* performance in terms of task accuracy – as instructed, but also task precision. They also showed a similar error-rate performance with error-focus participants. As a result, they also exhibited the second highest throughput measure. We speculate that accuracy-focus participants aimed to select the center of the targets as instructed, which intrinsically led to fewer errors and performing precise selections. However, executing the task so carefully resulted in a significantly higher completion time, as expected. This implies that focusing on performing the tasks accurately also yields high performance in terms of error rate and precision.

Precision-focus participants made relatively fewer errors, albeit at the cost of exhibiting the highest time. We observed that their level of choosing the targets accurately and precisely was also poor – but still slightly higher than speed-focus participants. As a result, they showed the lowest overall performance in the form of throughput from the start to the end.

No-focus participants performed the task in less time than precision-focus, accuracy-focus, or error-focus participants and with a similar number of errors as error-focus and accuracy-focus participants. Interestingly, their task precision and accuracy were in the middle when ranked among the other groups, as expected. Thus, it is not surprising that their effective throughput results indicate that no-focus participants performed the experiment *second highest*, following speed-focus participants.

To sum up, as shown by the results, all participant groups followed the assigned task instructions properly, except the precision-focus group. Thus, we can state that assigning different task instructions might, in fact, change the ultimate performance on the task: *(i)* if participants need to get faster, they should train with a speed-focus, *(ii)* if they need to achieve minimum error, they should train with an error-focus or accuracy-focus, *(iii)* if they need to get more accurate, they should be trained with an accuracy-focus, and *(iv)* Thus, these results answer our research question RQ1 as **different task execution instructions yield different trends in different aspects of the user performance metrics, and different instructions should be adopted for future training sessions if there are certain expectations regarding participants' growing goals**.

Since the precision-focus participants were expected to show the highest performance [9], these findings are particularly surprising and raise intriguing questions regarding the relationship between precision-focus instructions and task performance. To understand the factors affecting the performance of the precision-focus participants, we conducted a supplementary study to evaluate the potential effect of language barriers. The results indicate that there was no significant difference between native and non-native English speakers. However, the interaction results showed that native English speakers were more accurate and precise with AR HMDs and more accurate with VR HMDs. These results highlight that participants might require more localized instructions to perform EHCT in VR systems to optimize training results. Thus, the follow-up question for RQ1 can be concluded as **"precision" does not offer the most optimal training procedure, unlike hypothesized in previous work.**

6.2 Effective Throughput

Our results show that speed-focus, error-focus, accuracy-focus, and no-focus participants followed and complied with the assigned instructions. Interestingly, we observed that their throughput rankings comply with participants' ability to comply with the assigned instructions to them. Speed-focus participants got faster, error-rate participants made fewer errors, and accuracy-focus participants were more accurate. In each case, the effective throughput of the participants increased. Moreover, since the precision-focus participants got faster, their effective throughput also increased. In general, any performance improvement is visible in the throughput, even when trainees' performance might not improve in terms of other measures (e.g., time, accuracy, precision, and error rate). For example, if a trainee wants to improve their accuracy or precision, they should focus on accuracy, and this would also be visible in the effective throughput. Furthermore, if a trainee does not focus on a specific task execution strategy, their error rate, accuracy, and precision performance could still improve, which are again reflected in the effective throughput results. Thus, it is important to plan the task execution strategy ahead of the training sessions to improve user performance in specific ways in HMD-based VR EHCTSs. Thus, these results answer our research question RQ2 as **effective throughput can be used as an assessment criterion for EHCTSs in 2D touchscreen display, VR-, and AR-based HMD**.

The speed-accuracy trade-off in Fitts' Law and effective throughput is still inconclusive. While MacKenzie and Isokoski showed that effective throughput is speed-accuracy invariant [42], Olafsadir et al. [52] and Batmaz and Stuerzlinger [14] identified contradictory results. The performance growth observed in our longitudinal study also showed that effective throughput increases regardless of whether participants focus on speed or accuracy. Still, the highest throughput performance was observed with speed-focus participants. In a motor learning study, one can expect such an increase in the motor performance of the participants; however, the pre-experiment results on time, accuracy, and effective throughput still exhibit substantial throughput variance. These outcomes still contradict MacKenzie and Isokoski's work [42]. Yet, we did not use an ISO9241-411 setup in this study, which is a standard method for evaluating the usability of software user interfaces [32]. Thus, we invite researchers to conduct longitudinal motor learning studies to further investigate effective throughput invariance with ISO9241-411 setups.

The results of the longitudinal study showed that different headsets might require different task instructions to increase the rate of learning (detailed results are shown in supplementary material). For instance, we observed a higher improvement in time for no-focus participants in 2D and VR HMDs potentially because of the usability challenges they might have faced with AR, where virtual elements are shown overlaid over the real world. On the other hand, the most improvement in time for AR-based training was observed for the error-focus participants. This issue might require a detailed analysis of the used hardware system to reveal its advantages and disadvantages. Still, our outcomes might also be used to create better training programs for the participants using different headsets.

In general, participants were observed to be better in terms of time, error rate, accuracy, precision, and throughput with a 2D display compared to VR and AR HMDs. These outcomes match most of the participants' expressed preferences during the experiment. This result also coincides with previous results of studies comparing real and virtual worlds [6, 7, 13]. The impact of various depth cues or stereo deficiencies, such as the vergence and accommodation conflict, are reasonable justifications for the difference. However, this result contradicts previous work on VR-based EHCTS, where the authors did not find a significant difference between 2D display and VR-based EHCTSs [13]. We believe that this is due to the length of our study, where participants had to experience stereo deficiencies for a long time, increasing their detrimental effects on motor performance [8]. Further, participants performed better in terms of time, error, and throughput with VR HMD compared to AR HMD. This result also aligns with user motor performance results getting lower using AR-based EHCTSs [13]. We believe that current AR HMDs can be used as a training system, but the user motor skills might not reach the level of 2D and VR-based EHCTS. Thus, these results answer our research question RQ3 as **VR HMDs can be used as an effective display technology also during longitudinal training studies for virtual EHCTSs**.

6.3 Usability of VR-based EHCTS

Participants did not exhibit or report significant mental or physical fatigue after the experiments. We speculate that manually adjusting the height of the virtual targets decreased the potential fatigue of the task. In addition, since users did not move their heads much in virtual and physical space, they did not experience and report simulator or motion sickness.

Overall, the results of this work can be applied to various fields to create better training systems where motor learning plays an important role. VR and AR-based training systems can be adjusted for specific needs, creating scenarios and virtual environments that closely mimic training settings, ensuring relevance and effectiveness. The data collection capabilities that exist with modern VR and AR HMDs allow trainers to gain valuable insights into trainee performance, enabling them to assess the effectiveness of the training program, identify areas for improvement, and make data-driven decisions to optimize training outcomes. Our findings here can be applied to create efficient training programs to increase the motor performance of the trainees.

6.4 Limitations and Future Work

Previous studies on Fitts' Law and human motor performance have already investigated the effects of movement direction on user performance, including [4, 7, 12, 23, 26, 37, 38, 57]. As visible in the supplementary materials, the results of our study match the outcomes of this previous literature; subjects are slower when they move their hand upwards. Also, their movement time increases and their accuracy decreases when they reach for further away targets [24, 53, 60, 63, 64]. Therefore, we did not analyze movement direction further in our study.

In our detailed precision-focus analysis, we conducted an RM-ANOVA with six participants, where three participants were native English speakers, and three were not. Although we found significant results with a large effect size ($\eta^2 > 0.14$), the number of participants was limited. For future work, we recommend conducting the same experiment with a large pool of participants since personal differences might have an impact on the results.

Future research will involve gathering information from professional athletes who have participated in real-world training and examining the associated learning effects in VR and AR HMDs. Investigating skill transfer from VR/AR systems to real-world performance is another goal of ours. In addition, EHCTSs may be employed in rehabilitation and medical research [3]; we thus intend to broaden the scope of our study to investigate this prospective application area for VR and AR-based EHCT. To understand the cognitive aspects of the interaction better, we will also study error-related negativity, providing insights into users' psychological responses to error-related visual and auditory feedback.

7 Conclusion

This research embarked on a comprehensive exploration of the impact of various task instructions on eye-hand coordination training using different display technologies. The study included 20 participants, each 4 participants focusing on speed, accuracy, precision, error rate, or no specific task instruction over a 10-day training period. The outcomes shed light on the intricate relationship between task instruction assignment, participant adherence, and task performance. The results emphasized the various effects of different task instructions on participants' psychomotor measures, which can be used to create more effective training systems. We also identified that effective throughput is potentially the best option for tracking the overall motor performance in eye-hand coordination training systems. Still, the unexpected outcome of the precision-focus group emphasizes the complexity of the relationship between task instruction, participant behavior, and task performance. To understand the underlying causes of this discrepancy, more investigation is required.

Acknowledgments

This work was funded by 1001 - TUBITAK Scientific and Technological Research Projects Support Program: grant number 221M458 and in part by NSF 2327569, NSF 2238313, and ONR N00014-21-1-2949. We would like to thank Moaaz Hudhud Mughrabi and Furkan Kaya for their contributions.

REFERENCES

[1] K. Ahir, K. Govani, R. Gajera, and M. Shah. Application on virtual reality for enhanced education learning, military training and sports. *Augmented Human Research*, 5:1–9, 2020.

[2] A. Ahmed and M. J. Sutton. Gamification, serious games, simulations, and immersive learning environments in knowledge management initiatives. *World Journal of Science, Technology and Sustainable Development*, 14(2/3):78–83, 2017.

[3] B. M. Asken, J. P. Mihalik, J. D. Schmidt, A. C. Littleton, K. M. Guskiewicz, and J. B. Hopfinger. Visual performance measures and functional implications in healthy participants: A sports concussion perspective. *Athletic Training & Sports Health Care*, 8(4):145–153, 2016.

[4] D. A. Barany, V. Della-Maggiore, S. Viswanathan, M. Cieslak, and S. T. Grafton. Feature interactions enable decoding of sensorimotor transformations for goal-directed movement. *Journal of Neuroscience*, 34(20):6860–6873, 2014. doi: 10.1523/JNEUROSCI.5173-13.2014

[5] M. D. Barrera Machuca and W. Stuerzlinger. Do stereo display deficiencies affect 3d pointing? In *Extended abstracts of the 2018 CHI conference on human factors in computing systems*, pp. 1–6, 2018.

[6] M. D. Barrera Machuca and W. Stuerzlinger. The effect of stereo display deficiencies on virtual hand pointing. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pp. 1–14, 2019.

[7] A. U. Batmaz. *Speed, precision and grip force analysis of human manual operations with and without direct visual input*. PhD thesis, Strasbourg, 2018.

[8] A. U. Batmaz, M. D. Barrera Machuca, J. Sun, and W. Stuerzlinger. The effect of the vergence-accommodation conflict on virtual hand pointing in immersive displays. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–15, 2022.

[9] A. U. Batmaz, M. de Mathelin, and B. Dresp-Langley. Getting nowhere fast: Trade-off between speed and precision in training to execute image-guided hand-tool movements. *BMC psychology*, 4:1–19, 2016.

[10] A. U. Batmaz, M. de Mathelin, and B. Dresp-Langley. Seeing virtual while acting real: Visual display and strategy effects on the time and precision of eye-hand coordination. *PloS one*, 12(8):e0183789, 2017.

[11] A. U. Batmaz, M. de Mathelin, and B. Dresp-Langley. Effects of 2d and 3d image views on hand movement trajectories in the surgeon's peri-personal space in a computer controlled simulator environment. *Cogent Medicine*, 5(1):1426232, 2018.

[12] A. U. Batmaz, M. de Mathelin, and B. Dresp-Langley. Effects of image size and structural complexity on time and precision of hand movements in head mounted virtual reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 167–174, March 2018. doi: 10.1109/VR.2018.8446217

[13] A. U. Batmaz, A. K. Mutasim, M. Malekmakan, E. Sadr, and W. Stuerzlinger. Touch the wall: Comparison of virtual and augmented reality with conventional 2d screen eye-hand coordination training systems. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 184–193. IEEE, 2020.

[14] A. U. Batmaz and W. Stuerzlinger. Effective throughput analysis of different task execution strategies for mid-air fitts' tasks in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3939–3947, 2022.

[15] A. U. Batmaz, X. Sun, D. Taskiran, and W. Stuerzlinger. Hitting the wall: Mid-air interaction for eye-hand coordination. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–5, 2019.

[16] A. U. Batmaz, X. Sun, D. Taskiran, and W. Stuerzlinger. Eye-hand coordination training for sports with mid-air vr. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–10, 2020.

[17] B. Bideau, R. Kulpa, N. Vignais, S. Brault, F. Multon, and C. Craig. Using virtual reality to analyze sports performance. *IEEE Computer Graphics and Applications*, 30(2):14–21, 2009.

[18] A. T. Biggs, D. J. Geyer, V. M. Schroeder, F. E. Robinson, and J. L. Bradley. Adapting virtual reality and augmented reality systems for naval aviation training. *Naval medical research unit dayton wright-patterson AFB United States*, 2018.

[19] J. Boritz and K. S. Booth. A study of interactive 3d point location in a computer simulated virtual environment. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pp. 181–187, 1997.

[20] A. Covaci, C.-C. Postelnicu, A. N. Panfir, and D. Talaba. A virtual reality simulator for basketball free-throw skills development. In *Technological Innovation for Value Creation: Third IFIP WG 5.5/SOCOLNET Doctoral Conference on Computing, Electrical and Industrial Systems, DoCEIS 2012, Costa de Caparica, Portugal, February 27-29, 2012. Proceedings 3*, pp. 105–112. Springer, 2012.

[21] C. Craig. Understanding perception and action in sport: How can virtual reality technology help? *Sports Technology*, 6(4):161–169, 2013.

[22] Dynavision International, 2019.

[23] L. Fernandez and R. J. Bootsma. Effects of biomechanical and task constraints on the organization of movement in precision aiming. *Experimental Brain Research*, 159(4):458–466, Dec 2004. doi: 10.1007/s00221-004-1964-4

[24] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6):381, 1954.

[25] K. L. Gao, S. S. Ng, J. W. Kwok, R. T. Chow, and W. W. Tsang. Eye–hand coordination and its relationship with sensori-motor impairments in stroke survivors. *Journal of rehabilitation medicine*, 42(4):368–373, 2010.

[26] F. Grimm, G. Naros, and A. Gharabaghi. Closed-loop task difficulty adaptation during virtual reality reach-to-grasp training assisted with an exoskeleton for stroke rehabilitation. *Frontiers in Neuroscience*, 10:518, 2016. doi: 10.3389/fnins.2016.00518

[27] J. F. Hair Jr, W. C. Black, B. J. Babin, and R. E. Anderson. Multivariate data analysis, 2014.

[28] J. A. Harpham, J. P. Mihalik, A. C. Littleton, B. S. Frank, and K. M. Guskiewicz. The effect of visual and sensory performance on head impact biomechanics in college football players. *Annals of biomedical engineering*, 42:1–10, 2014.

[29] C. Huang, Y. Zhang, C. Zhu, C. Zhang, and H. Meng. Chinese sports basketball teaching tactics training system combined with multimedia interactive model and virtual reality technology. *Multimedia Tools and Applications*, 82, 02 2019. doi: 10.1007/s11042-019-7298-9

[30] H.-M. Huang, U. Rauch, and S.-S. Liaw. Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist approach. *Computers & Education*, 55(3):1171–1182, 2010.

[31] Y. Huang, L. Churches, and B. Reilly. A case study on virtual reality american football training. In *proceedings of the 2015 virtual reality international conference*, pp. 1–5, 2015.

[32] I. Iso. 9241–411 ergonomics of human-system interaction–part 411: Evaluation methods for the design of physical input devices. *International Organization for Standardization*, 2012.

[33] C. Jin, L. Dai, and T. Wang. The application of virtual reality in the training of laparoscopic surgery: A systematic review and meta-analysis. *International Journal of Surgery*, 87:105859, 2021.

[34] J. W. Keogh and P. W. Winwood. The epidemiology of injuries across the weight-training sports. *Sports medicine*, 47(3):479–501, 2017.

[35] K. Krasich, B. Ramger, L. Holton, L. Wang, S. R. Mitroff, and L. Gregory Appelbaum. Sensorimotor learning in a computerized athletic training battery. *Journal of motor behavior*, 48(5):401–412, 2016.

[36] D. G. Liebermann, L. Katz, M. D. Hughes, R. M. Bartlett, J. McClements, and I. M. Franks. Advances in the application of information technology to sport performance. *Journal of sports sciences*, 20(10):755–769, 2002.

[37] C. J. Lin, H. Chen, P. Cheng, and T. Sun. Effects of displays on visually controlled task performance in three-dimensional virtual reality environment. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 25(5):523–533, 2015. doi: 10.1002/hfm.20566

[38] R. F. Lin and H. Cheng. Accessing effects of various depth-cue combinations on hand control movement in a virtual environment. In M. Kurosu, ed., *Human-Computer Interaction. Novel User Experiences*, pp. 527–537. Springer International Publishing, Cham, 2016.

[39] X. Liu, J. Zhang, G. Hou, and Z. Wang. Virtual reality and its applica-

tion in military. In *IOP Conference Series: Earth and Environmental Science*, vol. 170, p. 032155. IOP Publishing, 2018.

[40] T. Luger, T. Bosch, M. Hoozemans, M. de Looze, and D. Veeger. Task variation during simulated, repetitive, low-intensity work–influence on manifestation of shoulder muscle fatigue, perceived discomfort and upper-body postures. *Ergonomics*, 58(11):1851–1867, 2015.

[41] I. S. MacKenzie and W. Buxton. Extending fitts' law to two-dimensional tasks. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 219–226, 1992.

[42] I. S. MacKenzie and P. Isokoski. Fitts' throughput and the speed-accuracy tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1633–1636, 2008.

[43] I. S. MacKenzie and A. Oniszczak. A comparison of three selection techniques for touchpads. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 336–343, 1998.

[44] P. Mallery and D. George. SPSS for windows step by step: a simple guide and reference. *Allyn, Bacon, Boston,*, 2003.

[45] Meyend. Meyend training, September 2022.

[46] S. C. Michalski, A. Szpak, D. Saredakis, T. J. Ross, M. Billinghurst, and T. Loetscher. Getting your game on: Using virtual reality to improve real table tennis skills. *PloS one*, 14(9):e0222351, 2019.

[47] H. C. Miles, S. R. Pop, S. J. Watt, G. P. Lawrence, and N. W. John. A review of virtual environments for training in ball sports. *Computers & Graphics*, 36(6):714–726, 2012.

[48] M. H. Mughrabi, F. Kaya, A. U. Batmaz, A. Aliza, W. Stuerzlinger, B. Borazan, E. Tonyali, and M. Sarac. On the effectiveness of virtual eye-hand coordination training with head mounted displays. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 36–43, 2023. doi: 10.1109/VRW58643.2023.00014

[49] A. K. Mutasim, W. Stuerzlinger, and A. U. Batmaz. Gaze tracking for eye-hand coordination training systems in virtual reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–9, 2020.

[50] E. Nagorsky and J. Wiemeyer. The structure of performance and training in esports. *PloS one*, 15(8):e0237584, 2020.

[51] NIKE-SPARQ. Nike introduces SPARQ Sensory Performance (SSP), October 2011.

[52] H. H. Olafsdottir, Y. Guiard, O. Rioul, and S. T. Perrault. A new test of throughput invariance in fitts' law: Role of the intercept and of jensen's inequality. In *26th BCS Conference on Human Computer Interaction (HCI 2012)*, 2012.

[53] E. W. Pedersen and K. Hornbæk. An experimental comparison of touch interaction on vertical and horizontal surfaces. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, NordiCHI '12, pp. 370–379. ACM, New York, NY, USA, 2012. doi: 10.1145/2399016.2399074

[54] D. Poltavski and D. Biberdorf. The role of visual perception measures used in sports vision programmes in predicting actual game performance in division i collegiate hockey players. *Journal of sports sciences*, 33(6):597–608, 2015.

[55] Quotronics Limited. Batak fitness training solutions, 2019.

[56] B. M. Schout, A. Hendrikx, F. Scheele, B. L. Bemelmans, and A. Scherpbier. Validation and implementation of surgical simulators: A critical review of present, past, and future. *Surgical endoscopy*, 24:536–546, 2010.

[57] B. N. Smetanin and K. E. Popov. Effect of body orientation with respect to gravity on directional accuracy of human pointing movements. *European Journal of Neuroscience*, 9(1):7–11, 1997. doi: 10.1111/j.1460-9568.1997.tb01347.x

[58] R. J. Teather, A. Pavlovych, W. Stuerzlinger, and I. S. MacKenzie. Effects of tracking technology, latency, and spatial jitter on object movement. In *2009 IEEE symposium on 3D user interfaces*, pp. 43–50. IEEE, 2009.

[59] J. Teizer, T. Cheng, and Y. Fang. Location tracking and data visualization technology to advance construction ironworkers' education and training in safety and productivity. *Automation in construction*, 35:53–68, 2013.

[60] S. Thompson, J. Slocum, and M. Bohan. Gain and angle of approach effects on cursor-positioning time with a mouse in consideration of fitts' law. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 48(5):823–827, 2004. doi: 10.1177/154193120404800517

[61] L. Wang, K. Krasich, T. Bel-Bahar, L. Hughes, S. R. Mitroff, and L. G. Appelbaum. Mapping the structure of perceptual and visual–motor abilities in healthy young adults. *Acta psychologica*, 157:74–84, 2015.

[62] N. J. Washington, S. Dogramaci, K. A. Steel, and E. Ellem. The use of simulation training to accelerate the rate of forward ice skating skill acquisition. *International Journal of Kinesiology and Sports Science*, 4(2):11–17, 2016.

[63] T. Whisenand and H. Emurian. Analysis of cursor movements with a mouse. *Computers in Human Behavior*, 15(1):85 – 103, 1999. doi: 10.1016/S0747-5632(98)00036-3

[64] T. G. Whisenand and H. H. Emurian. Effects of angle of approach on cursor movement with a mouse: Consideration of Fitt's law. *Computers in Human Behavior*, 12(3):481 – 495, 1996. doi: 10.1016/0747-5632(96)00020-9

[65] Y. Zhang, H. Liu, S.-C. Kang, and M. Al-Hussein. Virtual reality applications for the built environment: Research trends and opportunities. *Automation in Construction*, 118:103311, 2020.