Bachelor’s thesis

Information and Communications Technology

2025

Nelli Heinonen

Exploring the Effective Use of Visualization in Virtual Reality Training Environments

TURKU AMK
TURKU UNIVERSITY OF APPLIED SCIENCES

Bachelor’s Thesis | Abstract

Turku University of Applied Sciences

Information and Communications Technology

2025 | 36 pages

Nelli Heinonen

Exploring the Effective Use of Visualization in Virtual Reality Training Environments

As educational technology keeps advancing, new methods of teaching and learning are appearing with Virtual Reality (VR) being one of them. As VR differs from traditional teaching, it is vital to research good practices in utilizing this technology. A previous study by the thesis commissioner found that the visualized summative feedback of a fire extinguisher VR application did not offer trainees sufficient insight into their learning experience. This thesis developed an alternative visualization aiming to offer a more effective form of feedback. The graphical feedback presented in the training environment and the graphical feedback alternative developed for this thesis were compared to determine their effectiveness in communicating learning outcomes. Forty-six participants consisting of Turku University of Applied Sciences students and staff took part in the between-groups study that asked each group to examine one of the visualizations and report their understanding in the form of questionnaire responses. The participants reported higher levels of perceived understanding with the original visual feedback, while the group with the reworked visualization were more accurate in their quantitative assessment. The study found that the absence of clear-cut written feedback increased confusion when interpreting visualized summative feedback.

Keywords:

Virtual Reality, visualization, summative, feedback

Opinnäytetyö (AMK) | Tiivistelmä

Turun ammattikorkeakoulu

Tieto- ja viestintätekniikka

2025 | 36 sivua

Nelli Heinonen

Visualisoinnin tehokkaan käytön tutkiminen virtuaalitodellisuuden koulutusympäristöissä

Opetustekniikan kehittyessä uusia opetus- ja oppimismenetelmiä ilmaantuu, ja virtuaalitodellisuus (VR) on yksi niistä. Koska VR eroaa perinteisestä opetuksesta, on tärkeää tutkia hyviä käytäntöjä tämän teknologian hyödyntämisessä. Opinnäytetyön toimeksiantajan aiemmassa tutkimuksessa havaittiin, että palosammuttimen VR-sovelluksen visualisoitu summatiivinen palaute ei antanut harjoittelijoille riittävästi tietoa oppimiskokemuksestaan. Tässä opinnäytetyössä kehitettiin vaihtoehtoinen visualisointi, jonka tarkoituksena oli tarjota tehokkaampi palautemuoto. Harjoitteluympäristössä olevaa graafista palautetta ja tätä opinnäytetyötä varten kehitettyä graafista palautevaihtoehtoa verrattiin niiden tehokkuuden määrittämiseksi oppimistulosten viestimisessä. 46 Turun AMK:n opiskelijoista ja henkilökunnasta koostuvaa osallistujaa osallistui ryhmien väliseen tutkimukseen, jossa jokaiselle ryhmälle esitettiin valkotaululla yksi visualisoinneista analysoitavaksi ja heidän ymmärrystään mitattiin kirjallisen kyselyn kautta. Osallistujat kertoivat ymmärtävänsä alkuperäisen palautteen paremmin, kun taas muokattua visualisointia tarkastelevan ryhmän jäsenet osoittivat parempaa ymmärrystä kyselyssä. Tutkimuksessa todettiin, että selkeän kirjallisen palautteen puuttuminen lisäsi hämmennystä visualisoitua summatiivista palautetta tulkittaessa.

Asiasanat:

Virtuaalitodellisuus, visualisointi, summatiivinen, palaute

Contents

List of abbreviations **6**

1 Introduction **7**

2 Theoretical framework and literature review **10**

2.1 Visualization in classrooms 11

2.2 Virtual Reality 13

3 Developing the research prototype **15**

4.1 Protocol 21

4.2 Analysis 22

5 Results **23**

6 Discussion **27**

7 Conclusion **30**

References **31**

Appendices

Appendix 1. Survey page 1.

Appendix 2. Survey page 2

Figures

Figure 1. Fire extinguishing VR training environment, tutorial location. 15

Figure 2. Fire extinguishing VR training environment, extinguishing zone. 16

Figure 3. Fire extinguishing VR training environment, original feedback graph, graph B. 16

Figure 4. Snippet of results csv depicting one completed attempt. 17

Figure 5. First iteration of developed feedback graph. 18
Figure 6. Second iteration of developed feedback graph. 19
Figure 7. Graph A, Final iteration of developed research graph. 20

List of abbreviations

CSV	Comma Separated Values
ID	An Identifier
VR	Virtual Reality

1 Introduction

In the modern era of taking advantage of the internet to assist them in their studies [1], many students turn their attention to sites such as YouTube to fill in the blanks in their teachers' or tutor's explanations. As many as 94.67% [2] of college students reported using YouTube in their education. The participating students also reported the platform having "a positive effect on college course instruction [2] ". A significant number of students turn to YouTube videos for additional support, as these videos allow them to revisit challenging content at their own pace. Moreover, YouTube effectively utilizes visual elements such as color, shapes, and language to enhance understanding and engage learners in the educational process.

A simple copy-and-paste of a wiki page on to multiple PowerPoint slides may have been enough in the past, but in modern higher education and even later in life this does not suffice. Whether being a teacher or a student in higher education, it is essential to know how to best present the wanted information to peers and/or students [3] while remaining true to the data being condensed [4]. This applies especially when presenting complex or copious amounts of data. It is critical to be aware of the elements that could facilitate comprehension and those that may transform information into an incomprehensible mess of text, shapes, and colors [3]. That is, effectively utilizing graphs, colors and the flow of text is critical in forming an understanding of information and concepts [5]. Efficient use of these methods applies especially in scenarios where students are expected to learn independently without the guidance of a teacher as -is often the case with VR-based training or teaching.

Although VR education is a relatively new phenomenon, numerous studies show that such training can be effective [6] [7] [8]. In recent years VR has become a more widespread teaching or training option for schools [9] and industry [10] When a learner is isolated inside a VR headset and can only rely on the information before their eyes, without the possibility to ask for clarification, it is crucial that that information is clear and easy to understand.

While a student may be able to process a ton of information by reading materials in real life, the added cognitive load of VR [11] makes processing substantial amounts of written, or otherwise visualized, data difficult. All these factors paint a picture about just how important it is to be precise and mindful of what and how information is presented to students and trainees.

With the evolving capabilities and increasing popularity of VR to offer highly immersive experiences, it is no wonder that it is being examined as a safer training environment for potentially dangerous situations [12]. This thesis is a case study at Turku University of Applied Sciences as part of the Virtual Training Certifications project (funded by Business Finland) that examines the utilization of VR-based education tools and their analytics visualization techniques in teaching the proper procedures when extinguishing a fire. Ongoing research [13] with the VR-based training software showed that a larger portion of trainees missed or ignored a carefully constructed analytics graph at the end of training. The research concluded that the graph may have been too complex to read immediately, and that future development should consider simplifying it. The commissioner of the thesis is the Futuristic Interactive Technologies research group, and they have requested this study to investigate whether more simplified analytics would benefit the application.

For this thesis study, the researcher created a second (simplified) version of the analytics graph and conducted a “between subjects” study to determine whether the altered version is at least equally informative. In the analysis, one group will be asked to interpret the original graph, and another group will do the same for the simplified version. The study compares the two groups in how quickly and accurately they are able to interpret their respective graphs. This thesis aims to provide empirical evidence that in certain situations, simplified analytics offer a more valuable learning experience than comprehensive analytics that require complex visualization.

This thesis document will begin by presenting the literature that discusses this topic and different aspects of similar cases (Chapter 2). This is to obtain a better understanding of the industry practices as well as to acquire a better grasp on the encouraged and unadvised practices of visualization in Virtual Reality training environments. After that, this thesis will present the development and baseline research that took place to develop a viable alternative feedback visualization tool to be later used in a case-study using 46 volunteered participants. The thesis will then analyze and report the findings of the case-study as well as report suggestions for future research and utilization.

2 Theoretical framework and literature review

Throughout their academic lives, students will be presented with a multitude of various sources offering them information in trying to educate them. Whether that would be a teacher presenting information in a physical or digital classroom or from a book the student is required to understand. Whatever the source the student interacts with, they all aim to achieve the exact same goal: educating those interested. However, even with the same goal in mind, the way in which these sources present their information can have a significant impact on how well the students understand the information and concepts they are presented with [5].

It is not sufficient to merely state that visualization as an important and effective tool in ensuring a learner can best comprehend the information that is wanted of them to learn. Simply slapping images and colors on a PowerPoint presentation and presenting it for learners has a high chance to backfire [14]. When aiming to best utilize visualization, one must not forget the importance of text as well. A mismatched wall of seemingly unrelated images and colors, and a block of hard-to-read text are both the opposite ends of the same spectrum, where neither leaves the learner with a deeper, or even adequate, understanding of what they just saw [15]. The best results lie somewhere in the middle, where the text and other visuals complement and enhance each other, while not trying to overwhelm the learner [14]. One effective method for achieving this is by utilizing informational graphics (infographics). Infographics strategically make use of images and other forms of data visualization, such as pie charts and line graphs, and in doing so “add value by increasing understanding and the reach of research [16].”

One of the key considerations when designing infographics is to focus on the intended target audience [16]. Other notable factors to consider in the conceptualization and designing process include providing the reader with a narrative via clear start and end points, to ensure no information is missing [16]. In addition, while a text’s key points can be emphasized by utilizing larger font sizes and bold colors, one must pay close attention to ensure these methods are in harmony with the text and other elements around them [16]. To keep students from losing interest in their learning material, it was found that "all infographics features including the use of images and symbols, good design, attractive colors, concise texts and diagram or chart can encourage learners to understand better any learning information delivered through it [17]." Table 1 provides a summary of the key points of various studies conducted on the encouraged practices in visualization. The mentioned tips were heavily utilized in the research and development process of this thesis study.

Table 1. Authors providing tips for effective infographic design. [4]

Author	Effective Inforgraphic Strategies
Harris, 2013	Maintain a simple and focused infographic.
Richard, 2015	Allow time for design and communication.
Khoja, 2017	Select a topic that people care about.
Fain, Laihow & Claveria, 2017	Keep a specific audience in mind.
Martin, 2018	Present information clearly and truthfully.
Toporoff, 2018	Stick to one topic.
Walton, 2018	Include explanatory narrative in support if comprehensio
Compact Creative, 2019	Create a visual style and clear information flow.
Jordan, 2019	Include credible sources.

2.1 Visualization in classrooms

Keeping students interested in the information they are being taught instead of their phones has become especially difficult in the past few years, due to the alarmingly rapid decrease in the attention spans of students [18]. With social media platforms like TikTok offering unlimited amounts of content at any second of the day with rapidly changing trends, the unending bombardment of information is rapidly diminishing the cognitive capacity to process new and larger chunks of information [19]. With the need to keep grabbing students' attention constantly, teachers must adapt to these ubiquitous circumstances. This is where the utilization of various visualization methods and tools enter the fray. By utilizing short bursts of images and videos in their lessons, teachers and institutions can better keep students satisfied and focused for extended periods of time [20].

Visualization has also been discovered to be effective in developing a learners' critical thinking skills, which is a crucial step on the road of ensuring the learners are prepared for their future fields of work or study. Being able to visualize a problem or information is a crucial part of the solution [21] when it comes to developing critical thinking skills. A heavy visualization approach has been noted to increase communication as well as enabling students to take more analytical approaches to various problems [21]. A learner is also much more likely to retain visual information over written information [22]. In cases where visualization was utilized during the teaching process, students achieved better performances in their exams [22]. The possibility to expand on their material by adding visual elements such as images and videos to their lectures can also reduce the burden of the teachers, giving them greater opportunities to focus on feedback.

Throughout the teaching and testing processes, the teacher should offer assessment and feedback [23] to their students. Feedback can be formative, summative or an amalgamation of the two. Formative assessments and feedback refer to "frequent, interactive assessments of students' development and understanding to recognize their needs and adjust teaching accordingly [24]". Summative assessment on the other hand is mostly applied as tests and other tasks that provide the student with a numerical score of their performance, often with very limited additional feedback [25]. These results in summative assessments are mainly used to measure learning instead of the process of learning [24]. While summative assessment is a teacher-driven tool designed to measure knowledge transfer [26], summative feedback on the other hand serves students by allowing them to reflect on a recent learning performance themselves [27]. Summative feedback provides learners with a summary of their accomplishments and sets the tone for their further development under independent learning conditions [28]. A key application area of summative feedback lies in independent virtual reality head-mounted display learning environments [29]. This thesis will mainly focus on discussing summative feedback that takes place at the end of a VR learning scenario and how such feedback can be communicated most effectively.

2.2 Virtual Reality

With the development of virtual reality (VR) head-mounted display technology, it is no wonder it's being researched and tested for its potential to offer a safer and more repeatable solution to train and test trainees in high-risk situations, such as medical education [30], construction [7] and fire-extinguishing [13]. In the medical field, it not only addresses the concern for patient and trainee safety but also alleviates the shortage of training personnel and scarcity of equipment [6] that drive medical training. In construction, a primary push for the utilization of VR-training environments was caused by the need to re-train construction workers to know how to operate newly emerged construction robotics [7]. The goal of these training environments would be to allow trainees to safely and effectively train in dangerous situations without requiring the presence of a real-life trainer and without the fear of disaster should they fail their training lesson. Luckily, thanks to the broad evaluation tools VR offers, such as tracking a trainees' button operation, head, eye and hand movements, as well as automatically analyzing and processing data using embedded evaluation criteria, trainees can get instant summative feedback at the end of each training session [30].

With the intentionally limited supervisor presence during testing, the only form of feedback a trainee receives is what the program presents to them. Therefore, the feedback must be well presented, clear and offer valuable information about the trainee's performance in both successes and failures [31].

Notable issues that arise when only presenting feedback through the VR training environment include: (a) ensuring that the trainee notices their evaluation when presented; (b) the trainee's ability to read, process and understand the information given; (c) that the trainee finds value in the feedback; and (d) that trainers can accurately verify the occurrence of learning. Theoretical research shows that “novice students learn better with explanatory feedback (providing students with a principle-based explanations of why his or her answer was correct or incorrect) than with dichotomous feedback alone (simply informing the student whether their response was correct or not) or giving a score with no further explanation [32]”. Additionally, it is recommended that feedback “must only respond to specific and objective assessment criteria” and be expressed “in a language already known and understood by the learner” [33]. These examples put forward a strong point for in-depth written feedback being the best way to inform a trainee of their performance, but there is a hiccup with this when it comes to VR training environments. Reading in VR is challenging, and research shows that participants relying on VR-based reading situations were outperformed by their control group counterparts who were reading on more conventional media [34]. This puts the case forward for VR feedback being offered in a slightly more abstract form, such as with graphs and other visuals. It is still important to remember when pivoting to these non-written feedback alternatives, that they must still present the same key information as their written counterparts.

This thesis is going to evaluate how effectively participants can process different feedback graphs containing multiple pieces of varying information as well as how well participants retain the information they saw. This thesis will also collect participants’ evaluations on the clarity of different elements on the graphs they were shown.

3 Developing the research prototype

To conduct meaningful research into the visual feedback graph of the fire extinguishing training environment, the first step was to become more familiar with the application. This was done by playtesting the environment, which included completing the tutorial and performing multiple training rounds with differing levels of success. Additionally, this testing allowed the researcher to record the necessary video material for participant testing that would be conducted in a later stage of research, which followed in the footsteps of previous studies of the same training environment [13].

The application narrative starts with a tutorial (Figure 1) on how to use the virtual fire extinguisher. After the tutorial, the trainee would move on to the test site containing a simulated fire (Figure 2). The trainee was tasked with collecting and correctly utilizing a fire extinguisher while being aware of their distance to the fire as well as where they were aiming the spray of their fire extinguisher.

Figure 1. Fire extinguishing VR training environment, tutorial location.

Figure 2. Fire extinguishing VR training environment, extinguishing zone.

Upon completion or failure, the distance and spray statistics as well as (a) the content of the fire extinguisher; (b) time taken; (c) when the pin was pulled; (d) fire level; as well as (e) optimal distance and spray recommendations, were then displayed on a graph (Figure 3) for the trainee to inspect.

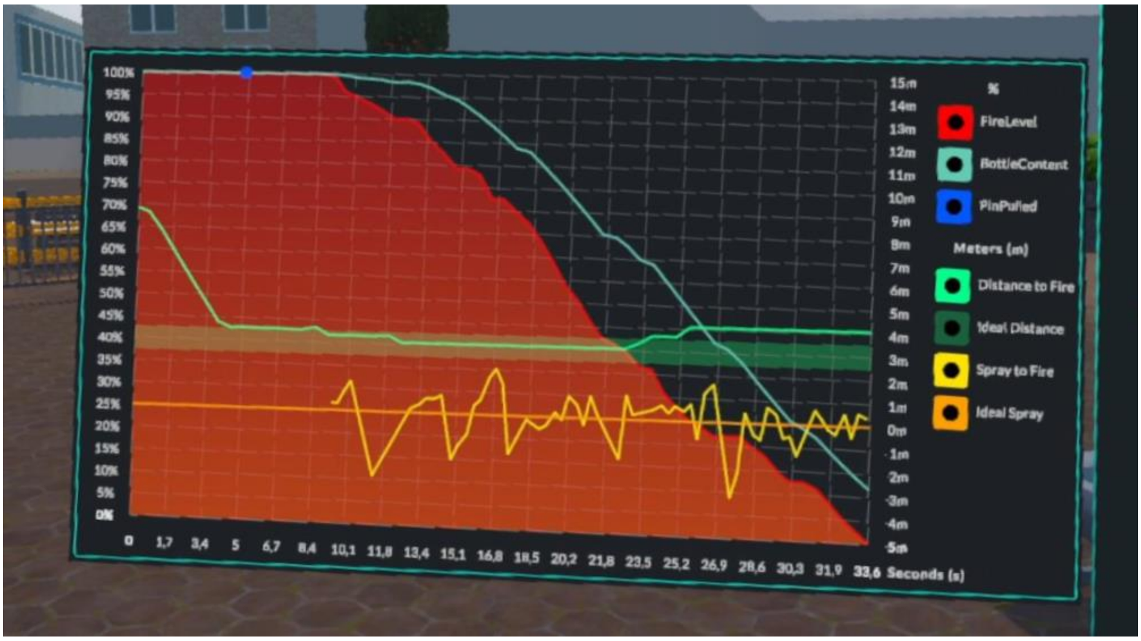

Figure 3. Fire extinguishing VR training environment, original feedback graph, graph B.

Before developing a new data visualization, the first task was to analyze and understand the structure of the pre-existing visual feedback in the fire extinguishing training environment. This consisted of acquiring the data from previous testing cycles that was saved and stored in a csv file within the game's directory. Each run's data was recorded on an appending list as shown in Figure 4. Each run had an ID showing who had logged into the system to complete the run and their attempt number from number 1 onwards. This ID accompanied all the stats of a run; an example of a statistics ID would be "K_S_015.1.TimeFromStart". A run is started by picking up the virtual fire extinguisher within the testing environment. The program would then record the desired statistics to the csv. The environmental data was logged ten times per second and later used by the program to create the feedback graph of the run. After understanding how the game records and displays its data, the next step was to start work on creating an alternative visualization method for the purposes of group testing and this thesis.

	A	B	C	D	E	F
1	K_G_037.1.TimeFromStart,0,0.1113756,0.2088903,0.306158,0.4					
2	K_G_037.1.DistanceToFire,8.893373,8.892408,8.891832,8.89143					
3	K_G_037.1.PinSuccess,False,False,False,False,False,False,False,Fa					
4	K_G_037.1.ExtinguisherEngaged,False,False,False,False,False,Fals					
5	K_G_037.1.HeightOfNozzle,0.5329066,0.5629109,0.5224768,0.4					
6	K_G_037.1.SprayToFireDistance,-10,-10,-10,-10,-10,-10,-10,-10,-10,-1					
7	K_G_037.1.BottleContents,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,					
8	K_G_037.1.FireIntensity,150,150,150,150,150,150,150,150,150,1					
9	K_G_037.1.ApproachWithRespectToWind,44.57121,44.57752,44					
10	K_G_037.2.TimeFromStart,0,0.1115319,0.2085488,0.3065405,0.					
11	K_G_037.2.DistanceToFire,7.961113,7.964087,7.967648,7.97041					

Figure 4. Snippet of results csv depicting one completed attempt.

The training program recorded the specific metric of a run as individual pieces into a singular cell in the csv, only separated by a comma. Once the data was made clear by utilizing Excel's organizing tools, it was time to create the first draft of the experimental graph. The first iteration (Figure 5) was created to distinguish the key statistics to display back to the trainee. This was necessary because the original csv also contained datapoints that offered information that the developer and project researcher found valuable—these data points had to be identified and ignored so that they would not confound the data collected for this thesis. It is important to note that the first draft was created without applying research findings of the good and bad habits of visualization. Once the draft was deemed as an acceptable starting point, it was sent for review.

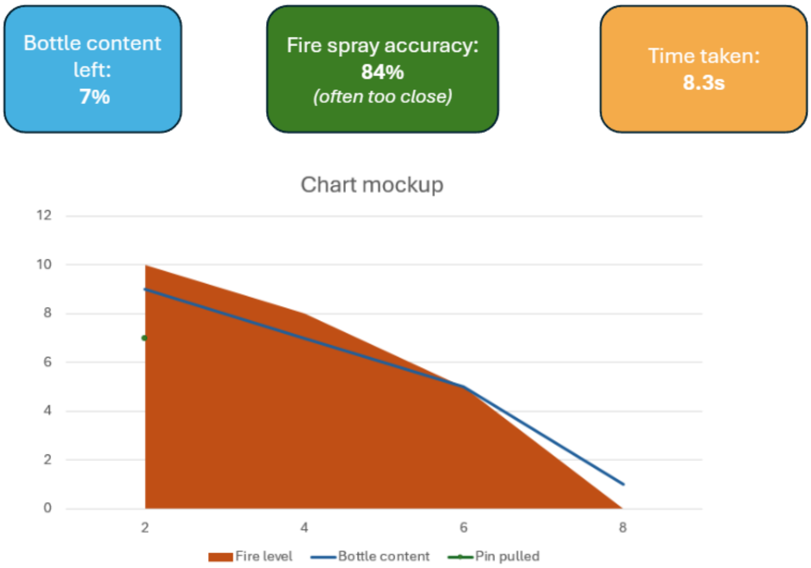

Figure 5. First iteration of developed feedback graph.

After the project researcher had reviewed the first draft, the thesis researcher created an updated version (Figure 6) that applied the suggested guidelines of other researchers in the fields of deferred feedback. This second iteration, following prescribed guidelines, echoed the pre-existing hypothesis of this thesis about the importance of the balance between color and text in all visual media. After applying many of the theoretical aspects found in research, the draft was updated to better reflect the findings and sent for a second round of review.

Figure 6. Second iteration of developed feedback graph.

The review process was followed by deeper research focused on finetuning the graph to its final form (Figure 7). The final iteration of the graph was developed to include the same pieces of information as the second iteration (Figure 3), while making it easier to understand. While it would be advisable to limit text-based feedback in Virtual Reality due to its nature [34], the researcher hypothesized the amount of information loaded onto a visuals-only graph to be overwhelming and difficult to process. The construction of the original graph, while containing a zone of optimal performance thus making it easy to tell an excellent performance from a terrible one, made it difficult to process performances that fell somewhere in the middle [3]. This led to the implementation of a scoreboard-like statistics screen being added to the right side of the graph. This was done to offer a summary of sorts to offer clarity for the middle-of-the-road performances. By color coding as well as giving optimal performance statistics as a comparison, the researcher hypothesized that the trainee would be able to tell how well they did at a quick glance and in full detail.

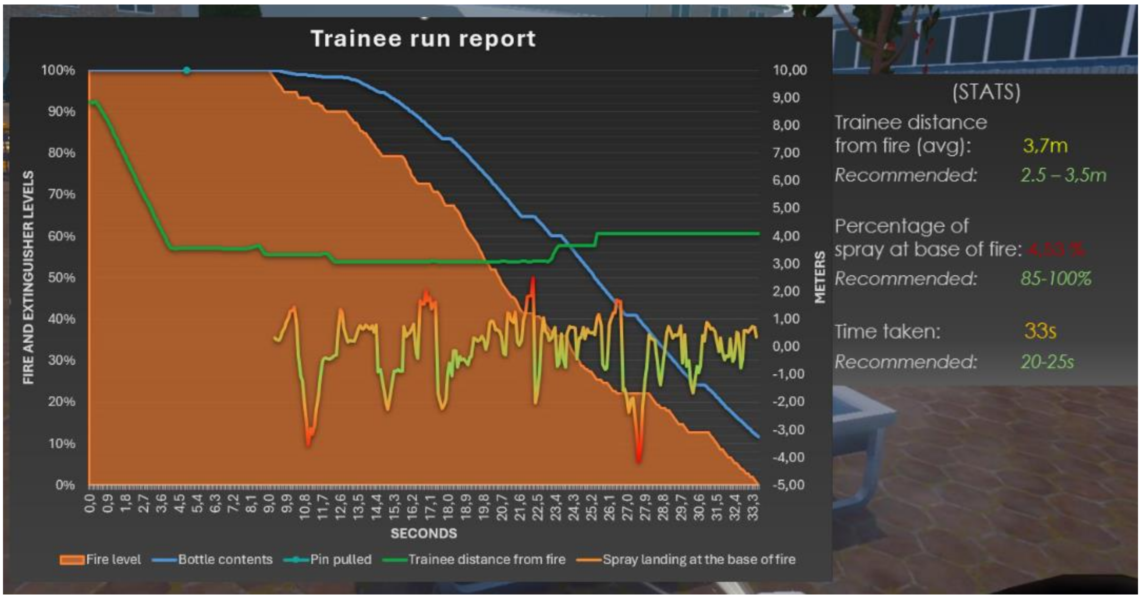

Figure 7. Graph A, Final iteration of developed research graph.

4 Methodology

This study used a between groups A/B testing approach whereby one group of participants attempted to interpret the original graph, and another group set about interpreting the visualization result of this thesis. This approach required participants to understand the application, but not to the point that they had to use the application. The reason for this decision was to avoid participants grappling with the controls and mechanics of the VR application, taking away from their interpretations of the graphs. To achieve this, the researcher allowed participants to watch a demonstration screen recording of a trainee using the VR application. The video was filmed in VR inside the fire extinguishing training environment. The recording was done with OBS Studio and with an Oculus Rift S – VR headset. An accompanying survey was created as a way to measure the participants' understanding of the graphs and consequently making it possible to compare the effectiveness of both the pre-existing graph and the graph created for testing purposes. The run shown on the video was for informational use only and does not align with the run data used for participant testing graphs.

The aim of the experiment is to determine which of the visualizations is faster to understand and to research how well testers grasp the information in each of the graphs.

4.1 Protocol

A total of 46 Participants were gathered on campus and presented with a short explanatory video of the developed fire-extinguisher training environment, along with a brief verbal clarification of what they were seeing. The video contained a brief playthrough of the in-game tutorial as well as a summarized playthrough of the main training process. Participants were informed that they would have one minute to analyze a visualization that was unrelated to the video they just watched before answering a survey. Participants were also informed that the questions would be simple analytical questions about the performance of the trainee as well as a section asking about the clarity of visual elements on the graph.

After the introductory phase, the participants were randomly assigned into either the A (n = 22) or B (n = 24) category. Group A participants were presented with the visualization that was developed for the purposes of this thesis. Group B was shown the original graph from the training software. The data on both charts were made from the collected performance data of the same playthrough.

Once each group had completed their visual inspections of their respective visualizations, they were presented with a short questionnaire (appendix 1 & 2) that asked them how well they thought the trainee performed based on the graph shown as well as how easy the graph was to understand. Participants were also offered a text box to give their reasoning for their responses as well as a way to leave any additional comments.

4.2 Analysis

Once testing had concluded, the surveys were then labeled, and their data was transferred into a csv sheet. After all of the answers and comments were documented, the amounts of correct, incorrect and unsure answers were calculated. The work then continued to calculate the opinion-based scores of the graphs' understandability.

5 Results

The total number of participants in the study was 46, 22 participants in group A (modified visualization developed for this thesis) and 24 participants in group B (the original visualization). The first page of the questionnaire (appendix 1) asked specific questions about the performance of the trainee. These questions had correct answers and tested the understanding of the participants. These questions included questions such as ‘how long did it take to extinguish the fire’, ‘was the trainee staying at an ideal distance from the fire’ as well as asking the participants to rate the trainee’s spray accuracy on a scale from 1 – 10 (10 being highest). Table 2 shows how the participants answered the first part of the questionnaire.

Table 2 Participant answers to trainee performance questions

Question	Number of responses	
	Group A (n = 22)	Group B (n = 24)
When was the pin pulled (correct)	4	7
When was the pin pulled (incorrect)	3	6
When was the pin pulled (unsure)	15	11
How long the extinguishing lasted (correct)	11	6
How long the extinguishing lasted (incorrect)	8	12
How long the extinguishing lasted (unsure)	3	6
Was the trainee staying at an idea distance (correct)	15	3
Was the trainee staying at an idea distance (incorrect)	2	19
Was the trainee staying at an idea distance (unsure)	5	2
Rating for trainee spray accuracy (correct = 1 to 3)	7	2
Rating for trainee spray accuracy (incorrect = 7 to 10)	3	11
Rating for trainee spray accuracy (unsure = 4 to 6)	12	11

The second page of the questionnaire (appendix 2) asked the participants to answer how well they thought they understood the different elements of the graph they were shown. The participant responses to indicate their perceived understanding are shown in Table 3. Additionally, participants were given the opportunity to leave any additional feedback about the test.

Table 3 Participant answers to understandability of the elements on the graph

Visualization element	Average / (Median) understanding Scale of 0 to 5	
	Group A (n = 22)	Group B (n = 24)
Fire level	3,8 (4,0)	4,17 (5,0)
Contents of the fire extinguisher	3,0 (3,0)	3,65 (4,0)
Distance to fire	3,5 (3,5)	3,96 (4,0)
Spray accuracy	2,7 (3,0)	3,33 (4,0)
How long extinguishing took	3,4 (3,5)	3,87 (4,0)
Performance compared to recommended stats	3,3 (3,0)	3,73 (4,0)
AVERAGE	**3,3**	**3,8**

From the tables above we can see that participants in both groups struggle to determine when the pin of the fire extinguisher was pulled with a large portion of both groups answering “I am not sure” when asked (group A = 68,2% and group B = 45,8%).

As seen in table 2, 50,0% of members of group A answered correctly and 36,36% answered incorrectly when asked how long the trainee took to extinguish the fire. In comparison, 25,0% of group B members answered correctly, and 50,0% of group B members answered incorrectly to the same question.

An apparent difference can be seen when participants were asked to analyze how well the trainee stayed at an optimal distance from the fire. Most members from group A answered correctly (68,2%) that the trainee was standing too far away while 12,5% of group B members answered correctly. Contrastingly 75,0% of group B members, but no one from group A, answered incorrectly ‘yes, the trainee was staying at the optimal distance’.

When participants were asked to rate on a scale from 1 to 10 (10 being perfect) how good the trainee’s spray accuracy was, the percentage of correct answers was low. Even though only 31,8% of group A members answered in the correct range, this result was still better than for group B, being only 8,3%. Conversely, the percentage of members from group B who answered in the completely incorrect range (45,8%) was over three times higher than those in group A (13,6%). Overall, group A (with the revised visualization) gave a much higher percentage of correct answers than group B (with the original visualization).

Moving onto the perceived understanding portion of the survey, members of group B reported higher levels of perceived understanding than their counterparts in group A by a clear margin. All participants were asked to rate how well they understood the different metrics of the graph on a scale from 0 – 5 (5 indicating the highest understanding). Members of group B reported a median understanding of four, whereas Group A with the original visualization reported a median understanding of between three and 3.5.

Out of 46 participants, 26 gave additional written feedback. A common comment present on both graphs was them being too overloaded with information. Both groups also felt that 1 minute was not enough time to fully understand and memorize the information. Participants further noted there were too many lines on the graph and also commented that there were too many similarly colored lines as well as some color inconsistencies making interpretation challenging.

6 Discussion

One of the key points of interest when conducting this thesis study was how much phrasing and word selection affect the understanding of written material. Poor phrasing can make even the simplest pieces of information difficult to understand. Consider these two phrases: ”*Mean baseline HbA1c of 73 diabetic patients before intervention was 8.9% and mean HbA1c after intervention was 7.8%* [3]*”* and “*Mean HbA1c of 73 of diabetic patients decreased from 8.9% to 7.8% after an intervention* [3].” The former displays the information as it was found in the conducted study. The means and results are fully displayed, but the reader must come to their own conclusion after processing the sentence. Meanwhile the latter offers the same pieces of information, but its form is much clearer and doesn't force the reader to come to their own conclusion while still truthfully presenting the same information. This emphasis on wording was echoed in some of the acquired feedback collected in the surveys. Some participants commented on the overly complicated wording on the metrics shown on the graphs.

In the graph shown to members of group A, an additional statistics board was shown, displaying in text the recommended distance as well as the average distance of the trainee. This statistics board was not present in the graph shown to group B and was instead replaced by a highlighted section on the graph. This alternative, or supplemental, mode of delivering key information is one of the primary differences between the two visualizations in the experiment. The researcher feels that with an alternative mode of data representation; more participant preferences were being catered to. That is, the participants who were overwhelmed by the technical appearance of the graph could still acquire learning through a less daunting effort.

A curious thing to notice is that while participants reported all aspects of graph B (the original visualization) to be easier to understand across the board, a staggeringly low percentage of participants in group B answered correctly to any performance analysis questions. This gives the idea that while graph B was perceived to be easier to understand, the story it offered to participants was misinterpreted. This may have been due to the fact that while both visual reports held the same pieces of key information, graph B condensed all its information only to a single thing to look at, without the need to find correlation between two visualizations. This may have caused a better sense of understanding, since the graph had fewer differing elements to process. Contrastingly graph A's lower score may be due to some developmental oversight that was caught once testing had already started. A portion of the legend for trainee spray accuracy did not correlate clearly with the data it represented on the graph. Another factor may have been the color choice of the spray accuracy on the additional score board being difficult to read. Both these theories were echoed in the written feedback of some group A members.

Another common theme was that participants of both groups struggled to pinpoint when the pin of the fire extinguisher was pulled. A few theories that would explain this include the color of the 'pin pulled' metric being difficult to notice and participants treating this statistic as less important than, for example, the 'trainee distance' metric. If future development deems the pin pulled metric more important than it currently is or more significant than other metrics of the graph, a change of color to a striking magenta could be advisable. The researcher would still like to note that while a color that pops out more could invoke a sense of importance, the person viewing it may still deem the information it represents as less valuable than other metrics. In order to fairly balance out all metric of the graph, it would be advisable to conduct further research into color theory and the different associations of different colors.

In future development and testing, it would be key to consider if this type of graphical feedback is the best way to go. It should also be taken into consideration that many participants reported that they would have performed better if they had had more time to analyze the graphs. Giving participants 2 – 5 minutes may have offered more interesting insights.

By conducting the testing personally on graph A, the researcher determined that one minute offered barely enough time for a once-over of the whole graph, let alone having additional time left to re-read sections or for any additional steps to commit the information to memory. Typically, when going over a performance report, a person would first take in the big picture before focusing on and comparing the specific details. With the experiment offering barely enough time for a once-over, it is not surprising that the sentiment of having too little time was echoed both in written feedback as well as in the test scores. By removing the possibility to confirm their understanding by re-reading a section, it is no wonder that test scores varied. An additional explanation for the comments about time feeling limited is the fact that participants were not given a visible timer. This decision was made to remove the possible added stress that a countdown could cause. While timer stress remains plausible, offering a way for participants to manage their limited time may have yielded more stable results. If participants were aware of the time they had left, they would have known when to do that final recap of the information they wanted to remember. Taking all of this into account, if similar tests were to be conducted, adding a timer that counts up from zero instead of counting down, could remove the stress that a countdown could cause while still offering participants a sense of time.

While all of the aforementioned points ring true in real life experiments, they would be amplified if similar tests were to be conducted in VR. If the same task of analyzing a visual performance graph was given to be done inside the virtual reality training environment, the analysis time should be reevaluated. With the added strain of a VR headset and the struggle with reading in a VR environment, a single minute for analysis would probably not be feasible.

7 Conclusion

This thesis set out to investigate the visual representation of summative feedback in a Virtual Reality training environment by comparing an existing visualization with a revised version that adheres to theoretical guidelines found in academic literature. While condensing data on a singular graph was found to be easier to understand, providing an additional written statistics screen was found to boost correct understanding of given information. Even considering the difficulties that reading in Virtual Reality brings, it may still be advised to include more key information as written feedback. This would ensure that trainees take away correct information from their training feedback. In the future, it would be wise to research what led to the participant's perceived understanding of the trainee's fire extinguishing case performance and why the documented understanding differed so widely between the groups.

If similar studies are conducted in the future, certain changes to the protocol may be advisable. Based on given written feedback, it may be beneficial to increase the time the participants were given to understand and memorize the information on the graphs. Another promising future change would be to give participants the opportunity to look at the questions beforehand or to give them a clearer briefing on what information is and is not important on the performance feedback graph. These suggested changes are of a pedagogical nature and were not a set objective for this thesis, but are nevertheless advisable to take place, should the development team wish to focus on finetuning the feedback visuals to be even more effective.

References

[1] O. D. Apuke and T. O. Iyendo, "University students' usage of the internet resources for research and learning: forms of access and perceptions of utility," *Heliyon,* vol. 4, no. 12, 2018.

[2] S. Greeves and M. Oz, "YouTube in higher education: comparing student and instructor perceprtions and practices," *Frontiers in Education,* vol. 8, 2024.

[3] T. S. Fah and A. F. A. Aziz, "How To Present Research Data?," *Malaysian family physician : the official journal of the Academy of Family Physicians of Malaysia,* vol. 1, pp. 82-85, 2006.

[4] J. Steen and K. Markert, "Bias in News Summarization: Measures, Pitfalls and Corpora," Department of Computational Linguistics, Heidelberg, Germany, 2024.

[5] S. Schukajlow, J. Schoenherr and A. R. Strohmaier, "Learning with visualization helps: A meta-analysis of visualization in mathematics education," *Educational Research Review,* vol. 45, 2024.

[6] E. Makled, A. Yassien, P. Elagroudy, M. Magdy, S. Abdennadler and N. Hamdi, "PathoGenious VR: Vr medical training," in *the 8th ACM International Symposium*, 2019.

[7] P. Adami, P. B. Rodrigues, P. J. Woods, B. Becerik-Gerber, L. Soibelman, Y. Copur-Gencturk and G. Lucas, "Effectiveness of VR-based training on improving construction workers' knowledge, skills, and safety behavior in robotic teleoperation," *Advanced Engineering Informatics,* vol. 50, 2021.

[8] A. Menin, R. Torchelsen and L. Nedel, "The effects of VR in training simulators: Exploring perception and knowledge gain," *Computers & Graphics,* vol. 102, pp. 402-412, 2022.

[9] C. Daily, "All About the History of Virtual Reality Education," Agile Meridian, 2022.

[10] B. Xie, H. Liu, R. Alghofaili, Y. Zhang, Y. Jiang, F. Lobo, C. Li, W. Li, H. Huang, M. Akdere, C. Mousas and L.-F. Yu, "A Review on Virtual Reality Skill Training Applications," *Frontiers Virtual Reality,* 2021.

[11] J. Juliano, S. N. and S. Liew, " Increased cognitive load in immersive virtual reality during visuomotor adaptation is associated with decreased long-term retention and context transfer," *Journal of NeuroEngineering and Rehabilitation,* vol. 19, 2022.

[12] Khairul, I. Purnama, A. Prananta, S. R. Megahati and A. Risnawati, "Advantages and Disadvantages of Virtual Reality in Science Learning Systems in the 21st Century," *Jurnal Penelitian Pendidikan IPA,* vol. 11, pp. 35-41, 2025.

[13] W. Ravyse, T. Tarkkanen, M. Luimula and J. Smed, "A Data-driven Approach to Identify Feedback Elements for Effective VR Training," in *15th International Conference on Applied Human Factors and Ergonomics (AHFE 2024)*, 2024.

[14] J. R. Yarbrough, "Infographics: In Support of Online Visual Learning," *Academy of Educational Leadership Journal,* vol. 23, no. 2, 2019.

[15] S. Stokes, "Visual Literacy in Teaching and Learning: A Literature Perspective.," *lectronic Journal for the integration of Technology in Education,* vol. 1, 2001.

[16] I. R. Murray, A. D. Murray, S. J. Wordie, C. W. Oliver, A. W. Murray and A. H. R. W. Simpson, "Maximising the impact of your work using infographics," *Bone & Joint Research,* vol. 6, no. 11, pp. 619-620, 2017.

[17] M. A. Mohd Noh, M. S. H. Mohd Fauz, H. F. Jing and M. F. Alias, "Infographics: Teaching and Learning Tool," *ATTARBAWIY Malaysian Online Journal of Education,* vol. 1, pp. 58-63, 2017.

[18] E. Getahun, "Are attention spans decreasing?," *Andrews University,* vol. 104, no. 09, 2022.

[19] G. Yuanyuan, H. Ying, W. Jinlian, L. Chang, I. Hohjin, J. Weipeng, Z. Wenwei, G. Wei, Z. Guang, Y. Qiong, W. Pinchun, Z. Manman, N. Xin, H. Qinghua and W. Qiang, "Neuroanatomical and functional substrates of the short video addiction and its association with brain transcriptomic and cellular architecture," *NeuroImage,* vol. 203, 2025.

[20] P. Lee, "Video to play growing role in higher education - Survey," *University World News,* 2016.

[21] K. Shatri and K. Buza, "The Use of Visualization in Teaching and Learning Process for Developing Critical Thinking of Students," *European Journal of Social Sciences Education and Research,* vol. 9, pp. 71-74, 2017.

[22] S. Dineva, "The importance of visualization in e-learning courses," in *The 14th International Conference on Virtual Learning VIRTUAL LEARNING – VIRTUAL REALITY*, 2019.

[23] T. Tutunaru, "Improving Assessment and Feedback in the Learning Process: Directions and Best Practices," *Research and Education,* 2023.

[24] S. M. Ismail, D. R. Rahul, I. Patra and E. Rezvani, "Formative vs. summative assessment: impacts on academic motivation, attitude toward learning, test anxiety, and self-regulation skill," *Language Testing in Asia,* vol. 12, 2022.

[25] H. Korkmazoglu, "The importance of Summative Feedback," 2022.

[26] N. Messier, "Summative Assessments," University of Illinois Chicago, Chicago, 2022.

[27] A. Adediwura and I. Ojediran, "A Comparative Study of the Effect of Formative and Summative Evaluative Feedbacks on Students’ Self-Efficacy," *African research review,* vol. 4, no. 1, 2010.

[28] S. University, "Summative Assessment and Feedback," Stanford University.

[29] J. Hayes, "The Importance of VR Simulations with Analytics for Instant Feedback," Virtual Medical Coaching, 2024.

[30] Y. Wang, Y. Li, C. Chen, Zhang, Y. Wang, K. Sha and S. Wang, "Research on virtual reality-based assessment framework and application path in medical education," *PLOS ONE,* vol. 19, 2024.

[31] A. M. Stiller, "Interactive summative assessment in an educational VR application," RWTH Aachen University, 2023.

[32] G. Makransky, R. Mayer, A. Nørremølle and L.-C. Ainara, "Investigating the Feasibility of Using Assessment and Explanatory Feedback in Desktop Virtual Reality Simulations," *Educational Technology Research and Development,* 2019.

[33] A. Potter and K. Lynch, "Quality feedback on assessment: apple for the teacher? How first year student perceptions of assessment feedback affect their engagement with study," 2008.

[34] M. Çoban, A. Topuz, A. Karabulut and T. K. Chiu, "Comparing Realities: A Study on the Impact of Virtual Reality versus," *Participatory Educational Research (PER),* vol. 11, pp. 19-36, 2024.

[35] C. Umayam, "Five (Easy-to-Implement) Ways Videos Can Have a Powerful Impact on Teaching and Learning," *Edsurge.com,* 2018.

Survey

1. Which diagram did you analyze?

☐ A
☐ B

2. When was the pin pulled?

☐ At __________ seconds
☐ I'm not sure when the pin was pulled
☐ The pin was not pulled
☐ I'm not sure if the pin was pulled

3. How long did it take to extinguish the fire?

☐ 0 – 10s
☐ 10 – 20s
☐ 20 – 30s
☐ 30 – 40s
☐ 40 – 50s
☐ The fire was not extinguished
☐ I'm not sure how long it took

4. Was the tester staying at an ideal distance from the fire?

☐ Yes
☐ No, they were too close
☐ No, they were too far away
☐ I'm not sure

5. On a scale of 1 – 10 (10 being perfect) how accurate was the spray?

1 2 3 4 5 6 7 8 9 10

☐ I'm not sure how good their accuracy was

Appendix 1. Survey page 1

6. Please rate on a scale of 0 – 5 how well you understood each of the following metrics in the graph (0 = I did not get it at all, 5 = Very easy to understand).

	0	1	2	3	4	5
Fire level	O	O	O	O	O	O
Contents of the extinguisher bottle	O	O	O	O	O	O
Trainee's distance from the fire	O	O	O	O	O	O
Percentage of spray landing at the base of the fire	O	O	O	O	O	O
How long it took to extinguish the fire	O	O	O	O	O	O
How well the trainee did regarding the recommended stats	O	O	O	O	O	O

7. Please give any additional feedback regarding your answers in question 6.

Appendix 2. Survey page 2