Dynamically Transparent Ghost Instructors and Their Effect on Learning and Skill Retention in Virtual Reality Environments

Comparing Static and Dynamic Transparency

Cassandra Visser[1]

Supervisor(s): Ricardo Marroquim[1]**, Amir Zaidi**[1]

[1]**EEMCS, Delft University of Technology, The Netherlands**

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 23, 2024

Name of the student: Cassandra Visser
Final project course: CSE3000 Research Project
Thesis committee: Ricardo Marroquim, Amir Zaidi, Chirag Raman

Abstract

Virtual Reality's uniquely immersive nature can facilitate skill transfer between an instructor and a student. Through the use of statically transparent ghost instructors that are superimposed on the student avatar, the student can learn skills by observing the ghost instructor's movements from a first-person perspective.

This paper aims to investigate the effect that changing the transparency values of the ghost avatar based on student performance has on learning and retaining skills. A small-scale user study has been conducted to contrast these dynamically and statically transparent ghost instructors. While the group taught by statically transparent ghost instructors displayed better performance improvement between training trials, the performance drop between the training and test trials was much smaller for the group taught by dynamically transparent ghost instructors.

1 Introduction

Virtual Reality has great potential for skill transfer in less traditional settings. It can allow physically separated users to have live lessons in which body language can be observed and followed. It can also benefit users that are in the same physical location by providing guidance to users by showing them the target movements from a first-person perspective or by displaying additional information in the field of view of the user. Virtual Reality has also been used for tai chi systems [1].

Various guidance systems for virtual reality have been developed thus far. Virtual co-embodiment systems [2] are an example of one such guidance system. In these systems, the teacher and student share and control the same virtual avatar, and their movement is averaged. Ghost-metaphor systems are another example. Systems such as "Just Follow Me" [3] have been shown to be effective in skill transfer, but these systems often rely on a statically transparent ghost instructor, and the effect of dynamic transparency has not been widely investigated.

Ghost metaphor-based instructors are generally of the statically transparent variety, meaning that the instructor avatars are displayed at a certain level of transparency and that this transparency is not varied. In contrast, the transparency level of a dynamically transparent ghost instructor, as introduced in this paper, changes depending on how well the student is performing the skill that is being taught. These skills are taught through imitation. The ghost instructor is superimposed on the avatar of the student to allow the student to observe the desired motions from a first-person perspective, and to be able to mimic these motions in real time. User studies in this domain may take the form of a training phase with guidance from the instructor followed by a test phase where the student must perform the skill without any guidance.

This research contributes new information as to the effectiveness of dynamic transparency in motor task learning in virtual reality using ghost metaphor-based instructors. It specifically contrasts the dynamic and static equivalents using a user study. Furthermore, the code base and models are publicly available and can contribute to further research into dynamically transparent ghost instructors in virtual reality.

The main question of this research is "How do dynamic transparency values impact the learning and retention of motor skills in collaborative virtual reality systems using a first-person perspective in combination with a ghost metaphor?" Three hypotheses were formed in relation to this main question.

- H1 Using a dynamically transparent ghost instructor, as compared to a statically transparent ghost instructor, leads to a smaller performance drop between the training and testing phase.
- H2 Using a dynamically transparent ghost instructor, as compared to a statically transparent ghost instructor, leads to a smaller performance improvement between trials of the training phase.
- H3 Using a (static or dynamic) ghost instructor leads to an increase in performance between sequential trials of the training phase.

The rest of this paper covers all aspects required to investigate this research question. This includes the full setup of the developed system, the user study used to evaluate it, and a section on responsible research.

2 Related Work

"Just Follow Me" [3] is a system developed to transfer skills from a skilled instructor to a student in Virtual Reality. It utilises a 'ghost metaphor' to display the instructor as a semi-transparent avatar superimposed on that of the student in combination with a swept volume and an acceleration vector indicator. The expert performs motions required to complete a certain task that the student is expected to follow. Evaluation of this system considered tracing a 3D trajectory and showed that trainees in a Virtual Reality setting performed better or as well as those in a real setting where a screen was used to display the motions from a third-person perspective instead.

Skill transmission and instruction utilising a mixed view has also been investigated through the use of a view-sharing system [4]. This system blends the real-world views of an expert and a student and displays this view to the participants through head-mounted displays. The study compared this blended view to various different assignments of a participant's view to the participant itself or to the other participant. The blended view was shown to produce the smallest errors in position compared to most conditions. However, in terms of velocity, the condition in which participants only saw the view of the other participant seemed to outperform the blended view, and the difference between the two was statistically significant. While this blended format does not strictly use virtual reality, it is interesting to note that a blended view may not always produce the most ideal results.

Another way to offer guidance is an Augmented Reality solution named Vishnu, which has been proposed to allow non-experts to perform a task without a locally available expert to help them [5]. In this system, a remote expert controls two virtual arms that are attached to the shoulders of the non-expert. The expert can use these arms to interact with virtual objects that resemble the real-world objects in the non-expert's vicinity. By viewing these interactions, the non-expert can follow the actions of the remote expert to perform the task. The study concluded that the AR solution was much faster than a traditional monitor-based solution in the context of complex tasks.

A more traditional VR solution called 'virtual co-embodiment' has been used for transferring skills between an expert and a student [2]. In this type of system, a teacher and a student are embodied in a single avatar, and the movement of the shared avatar is determined by the weighted average of the movement of both users.

A study comparing the learning and retention of motor skills through a no-guidance condition, a first-person perspective ghost metaphor (1PPGM) guidance condition, and a virtual co-embodiment guidance condition [6] came to the conclusion that the virtual co-embodiment performed significantly better in the testing phase than both other conditions. Virtual co-embodiment and the 1PPGM both showed significantly higher performance improvement between trials as compared to learning alone. However, the performance drop between the training phase with guidance and the testing phase without guidance was largest in the virtual co-embodiment condition, followed by the 1PPGM condition, and lastly the no-guidance condition. This study, however, does not consider a dynamically changing transparency value of the 1PPGM, which is investigated in this paper under the hypothesis that this may improve learning in addition to reducing the performance drop once the guidance is removed.

3 Methodology

The developed system aims to explore the impact of statically and dynamically transparent ghost instructors on skill learning and retention. The system consists of two vital parts: transparency control, and task representation. The system also contains logging functionality, and a session control module that runs trials without researcher input.

3.1 Transparency Control

The transparency control system consists of two components: error calculation and transparency value calculation. Its objective is to determine how transparent the ghost avatar should be based on how well the student is performing the taught skill. In the case of a statically transparent ghost instructor, the calculated error value is not taken into account in the transparency calculation as the transparency value is static throughout all training trials. Therefore, this section is applicable only to the dynamically transparent ghost instructors.

First, this system calculates the error between the locations of the right-hand index finger tip of the ghost avatar and the student avatar. This process is outlined by Algorithm 1. To compute the error, weighted Manhattan distance is used, where the global x and y-axis values have the same weight, and the global z-axis value has a weight half that of the other two. The reason that the z-axis is weighed differently is that this axis relates to depth, which may be more difficult to match. The same reasoning applies to the difference between the maximum values for the global x, y, and z axes. However, Euclidean distance is an alternative option that could be used to compute the distance.

Figure 1: Left: static transparency. Middle: dynamic transparency, high alpha value. Right: dynamic transparency, invisible.

```
weights = Vector3(1f, 1f, 0.5f);
maxErrors = Vector3(0.1f, 0.1f, 0.15f);

CalculateError(studentPos, ghostPos) {
   posError = abs(ghostPos - studentPos);

   // If pairwise '>' is greater for any dimension:
   if (posError > maxErrors) return 1.0f;

   // Scale to [0,1]. Pairwise '/'
   scaledError = posError / max;

   // Add weight scaling.
   weightedError = Dot(weights, scaledError)

   // Normalise to an error in the range of [0,1].
   return weightedError / (weights.x + weights.y +
       weights.z).
}
```

Algorithm 1: The error calculation method.

The error returned by Algorithm 1 is then used to determine the target transparency (in this case alpha value) of the ghost avatar through the use of Equations 1 and 2, where mod is the modifier. The transparency is interpolated between the minimum and maximum alpha values allowed in the current trial using a transparency modifier.

The transparency modifier is set to 0 if the error is below or equal to the error threshold. Otherwise, it is calculated by using Equation 1. That said, this modifier is incorrectly scaled in the code version that was used for the user experiments as the maximum value that can be obtained using this equation and the maximum error of 1 is less than 1.

$$mod = (e^{error} - 1)/e \tag{1}$$

The mentioned threshold is used for two reasons. The first is that the objective of this study is to investigate the effect of the dynamically and statically transparent ghost instructors on learning the general movement of a skill. An error below this threshold is thus "low enough". A similar reasoning applies to the use of the exponential function. If the student's movement is still close to the intended movement, the ghost's visibility should not dramatically increase. The second reason is to further incentivise the student to internalise the movement rather than to just follow the ghost hand.

$$transparency = (1 - mod) \times alphaMin + mod \times alphaMax \quad (2)$$

where mod is the modifier, and alphaMin and alphaMax are the minimum and maximum alpha value, respectively

3.2 Task Representation

The motion skill that the ghost instructor should teach the student is represented as a collection of cubic Bézier curves. Cubic Bézier curves are represented by a start point, end point, and two control points that determine the shape of the curved path between the start and end point (see Equation 3). A point on the curved path can be obtained by filling in a value for the t parameter between 0 and 1, inclusive, where 0 represents the start of the curve, and 1 the end.

$$P = (1-t)^3 P1 + 3t(1-t)^2 C1 + 3t^2(1-t) C2 + t^3 P2 \quad (3)$$

While the use of Bézier curves allows for the construction of elaborate paths, the t-value is not a parameterisation over the arc length of the curve due to the nature of the curve computation. This is visualised in Figure 2. If time is naively used as the t-value after scaling, this leads to lower speeds around curves as compared to the rest of the trajectory.

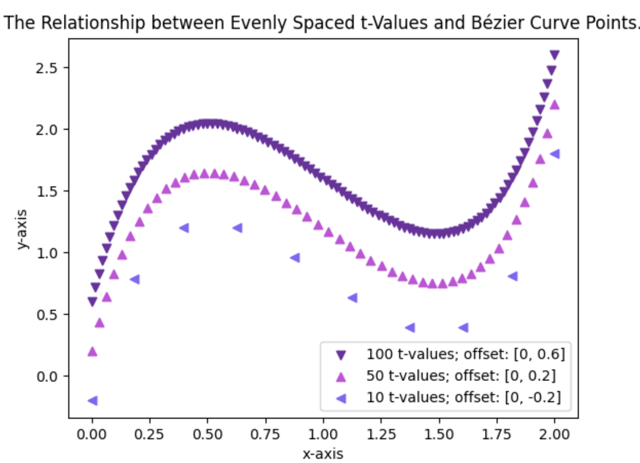

Figure 2: Evenly spaced out t-values were taken to plot points of a Bézier curve with P1 = (0,0), P2 = (2,2), C1 = (0.5, 4), and C2 = (1.5, -2). The points are in more crowded areas around steep parts of the curve. The offset was added to all plotted points for that specific number of t-values.

Due to this uneven spacing, the input parameter must be reparameterised to allow the ghost instructor to move at a constant speed throughout the curve. This is achieved through piece-wise linear approximation of the curve using line segments. The lengths of these segments can be summed to obtain the approximate total length of the curve. To even out the speed over all the curves in the path, the total lengths of the individual curves are used in a similar fashion.

3.3 Implementation Details

The system was developed as a Unity application using Unity version 2022.3.26f1 [7]. The code is available in a public Github repository [8]. This repository also includes the human and ghost models that were made using the MakeHuman plugin [9] (version mpfb2-20240602) for Blender [10]. Several MakeHuman asset packages [11], [12], [13], [14] were used in the creation of the models.

4 Evaluation

The developed system has been evaluated through a user study with a free-hand single-task set-up.

4.1 Experiment

For this experiment, all participants wore a HTC Vive Pro 2 Head-Mounted Display (HMD) and used two Vive controllers to map their real-life hand movements to that of their virtual human avatar. The experiments were all run on the same desktop computer. This computer had 32 GB of RAM installed, and was equipped with an RTX 2080Ti graphics card, complemented with an AMD Ryzen 7 3700X processor. The hardware was capable of handling the software without any issues.

A total of 12 participants were recruited for this user study through personal channels.

The participants were divided evenly into two groups. Group A (6 participants) was taught by a dynamically transparent ghost metaphor-based instructor. Group B (6 participants) was taught by a statically transparent ghost metaphor-based instructor, and functioned as the control group of the experiment. Both groups were asked to perform the same identical task, and the initial 3 trials were run with the same settings for both groups. This was done to make sure that group A has an initial understanding of what the motion should look like before the ghost avatar's transparency starts to vary. Learning multiple tasks to allow for a within-subject setup was out of the scope of this work to avoid sessions longer than 40 minutes.

The user study itself can be best explained by first giving the step-by-step procedure of the full experiment. This will be followed by a description of the procedure of an individual trial, and a description of the task itself. Lastly, the experiment settings will be discussed, including the logging frequency.

Step-By-Step Procedure

The experiment had an approximate duration of 30 minutes and followed the following step-by-step procedure:

1. The participant is handed a physical copy of the opening statement (Appendix A.1) and is asked to read the statement carefully.

2. Once they finish reading the opening statement, the researcher asks the participant if they have any questions related to the task. Answers are given only if they will not bias the results of the experiment.
3. A short explanation is given about the experiment objective.
   - The participant will be learning a motion task.
   - A ghost instructor will perform the motion to show what the motion looks like.
   - The ghost instructor is a semi-transparent, or see-through, green avatar.
   - The objective is to use your right virtual (the avatar's) index finger to trace the same curve as that traced by the index finger of the ghost instructor's right hand.
   - The motion is to be performed at the same time as it is showcased by the ghost instructor.
   - At the end, the participant should hold their hand in the end position for a short while to successfully register the trial as completed.
4. A short overview of the structure of the experiment is given.
   - The experiment starts with a training phase consisting of 10 trials. After completing the first 5 of these, there will be a 1-minute break in which the participant will be asked to perform a different task.
   - After the training phase, a 10-minute break is held in which the participant is, again, asked to perform a different task.
   - The test phase consists of 3 trials in which the participant must complete the task without any help from the ghost instructor.
5. The participant takes a seat on a chair that is placed in the same position for each individual experiment and is asked to put on the Head-Mounted Display (HMD). They are then either handed the controllers or they pick the controllers up by themselves.
6. The participant's location is recentred using SteamVR's Recenter functionality.
7. The Unity application containing the task, models, and program is launched, and either group A's or group B's profile is loaded, in accordance to the group assignment mentioned previously.
8. The participant is told that they can now start the trial, and that there will be text displayed on the wall that tells them what to do. The text itself says "Hold your hands in the turquoise areas for 2 seconds to start the trial".
9. If the participant is from group A, after the first 3 trials, they are informed that they may experience differences in the transparency of the ghost avatar depending on how well they perform the motion from this point onwards. This step was added following feedback from the pilot study stating that the change in transparency was quite jarring without being informed of it in advance.
10. After the first 5 trials, the participants are asked to take off the HMD. Once the HMD is placed, a 1-minute timer starts. They are then told to recite the alphabet from Z to A repeatedly until asked to stop. This is done to distract them from the task that they had been learning.
11. Once the minute passes, the participant is asked to put the HMD back on, and the controllers are handed to the participant or picked up by the participant.
12. The participant is then recentred, and the rest of the training trials are then conducted.
13. After the training trials are completed, the participant is asked to take off the HMD and to leave it on a surface. A 10-minute timer is activated once the head-mounted device is put down. They are then brought to a nearby desk with a pen and handed a paper with 3 Sudoku puzzles on them. A copy of the puzzles can be found in Appendix A.2. They are asked to start with the one labelled by "Expert-1". The objective of this activity is to distract the participant from the task that they had been learning. After the conclusion of the experiment, the puzzles are taken home by the participant or are disposed of.
14. Once 10 minutes have passed, the participant is asked to return to the chair and put the HMD back on. Again, the controllers are either picked up or handed over.
15. The participant is then recentred, and is reminded that the ghost instructor will not be visible anymore.
16. The participant completes the 3 test trials.

Trial Procedure

The trial procedure itself is fully handled by the software. Once the software is launched, the researcher can select the static or dynamic configuration. At this point in time, the participant sees the default Unity skybox. Once the researcher starts the experiment, the participant will load into the test environment.

The instructions on the wall of the test environment mentions turquoise areas that are visible from the moment the trial is loaded, but are generally not visible at eye level. The participant must hold their hands inside of the turquoise cubes for two seconds to start the trial.

Once a (training) trial starts, the ghost avatar will appear, and the participant must match the end point of their virtual right-hand index finger to that of the ghost instructor. The ghost instructor takes 20 seconds to complete its path. After the path completes, the participant has 5 more seconds to hold their hand in the ghost avatar's ending position for 0.5 seconds to complete the trial before the trial is marked as unfinished.

A test trial is identical to a training trial in every regard except for the visibility of the ghost instructor. In the test trials the ghost avatar is fully transparent and cannot be seen, leaving the participant to perform the motion based on their recollection of it.

Task

The task is, as mentioned in Section 3, represented through a list of cubic Bézier curves. The task is set up as a 3-

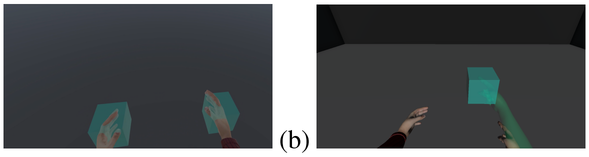

Figure 3: These images show what the turquoise cubes look like and where they are positioned. (a) Shows the starting cubes with the student's right hand in the correct position. (b) Shows the position of the end cube near the ghost avatar's final position. This cube is not visible to the student.

dimensional flower, and its front, top, and side views can be seen in Figure 4

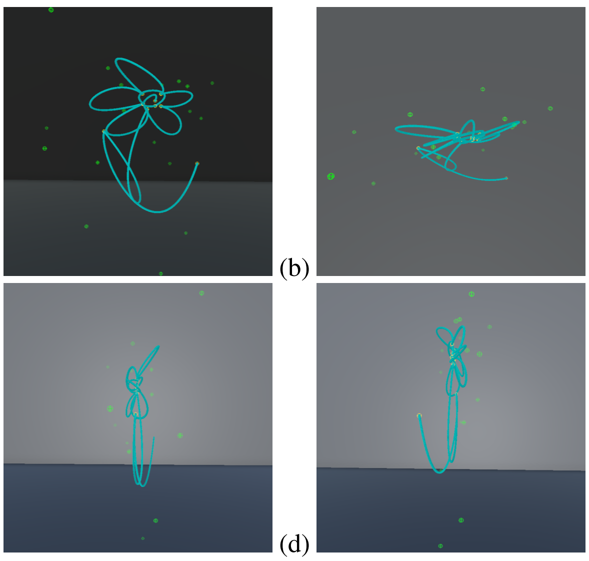

Figure 4: These images show what the task pathing looks like from different angles: (a) front view, (b) top view, (c) left side view, (d) right side view. The green spheres represent control points, and the yellow spheres represent start and end points.

Experiment Settings

As mentioned in Section 4.1, the collection of trials is run with a certain configuration. The differences and similarities between the static and dynamic configurations are shown in Table 1.

The configurations include important common information such as the time that the ghost instructor will take to fully trace the task curve as shown in Figure 4. The ghost instructor finishes showing the motion after 20 seconds. The participant then still has 5 seconds to finish the curve before the trial is marked as incomplete, as shown by the time limit.

The base transparency of the ghost instructor is set to 0.2, which was chosen due to its good visibility. This means that, for all training trials, the alpha value is set to 0.2 for the static configuration. This is overwritten during the test trials to render the ghost completely invisible. The dynamic configuration takes a slightly different approach. For each trial, a minimum and maximum transparency is defined, and the transparency is set according to the error and transparency calculations described in Section 3.

Parameter	Static	Dynamic
Number Training Trials	10	10
Number Test Trials	3	3
Path Tracing Time	20	20
Time Limit	25	25
Transparency Type	base	dynamic
Min. Transparency	[]	[0.2, 0.2, 0.2, 0.1, 0.1, 0.05, 0.05, 0, 0, 0, 0 0, 0]
Base Transparency	0.2	0.2
Max. Transparency	[]	[0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0, 0, 0]
Error Threshold	0.3	0.3
Data Logging Frequency	50 Hz	50 Hz

Table 1: This table shows the settings for the static and dynamic configurations. Rows that contain differences are highlighted in turquoise.

4.2 Results

A total of 10 out of the 12 sets of data gathered from the experiment described in Section 4.1 were analysed. 2 of the sets were disregarded due to breaks in procedure, resulting in one or more breaks being taken at the wrong point in the trial sequence. In one more set, the student assigned to Group A (dynamic) was told 1 trial late that the ghost instructor would now display varying transparency. This trial was not disregarded as this was not deemed a severe procedure violation.

A visual representation of the target position of the right virtual index finger can be seen in Figure 5. This figure also shows an example of what the gathered positional data of a single student during a single trial may look like. The data for both of these plots has been mean-shifted as the students' drawings were not always centred around the same value as that of the ghost instructor. By comparing both plots, it is easily noticeable that the shape drawn by the student is not fully accurate. Furthermore, it is fairly noticeable that the student takes much less time to go through the curve as compared to the ghost instructor, and ends up around their final position much earlier. This figure thus identifies an important difficulty in analysing the data: the students did not follow the exact same timing as the ghost instructor, making it difficult to directly compare their locations to determine the students' performance.

Due to this timing difference, two different ways to determine similarity were adopted in further analysis: Euclidean distance and Dynamic Time Warping (DTW). Euclidean distance was applied in a straightforward manner. First, the distance between the student's index finger position and that of the ghost instructor is computed per timestamp. Then, these computed distances are summed to represent the distance between the student's and the ghost's positions throughout a single trial. For the DTW computation, a python library called "dtaidistance" was used [15]. DTW is often used to find the most optimal alignment of time-series data. In this case, it

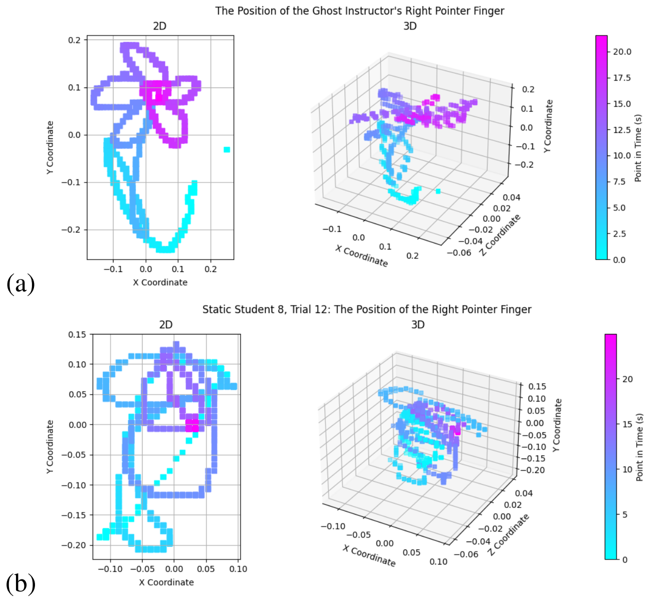

Figure 5: These images show the position data in time of: (a) the ghost, (b) a student in the static group during a test trial.

is used to find an optimal sequence of positions such that the distance between the ghost instructor's and student's movements is as small as possible.

Figure 6 shows the mean distance and standard deviation thereof between the mean-shifted positions of the student and that of the ghost per trial index. The bars are separated per experiment group (dynamic and static).

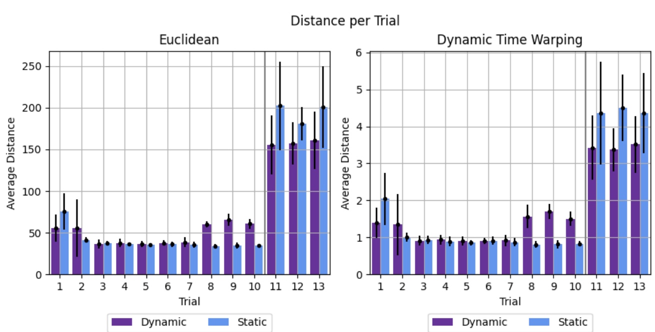

Figure 6: On the left, the average distance per trial computed using Euclidean distance is shown. On the right, the average distance per trial computed using Dynamic Time Warping is shown. For both plots, the standard deviation is shown through black error bars. The trials in the training and testing phases are separated by a black vertical line.

While there are visible differences between the dynamic and static groups during the first 3 trials, the ghost avatar's transparency was exactly the same for both groups during these trials. The difference is thus likely caused by natural variance. It should also be noted that the relative average distance computed for trial 12 differs by a relatively large amount. This likely indicates that the positions of the ghost instructor's index finger and that of the student may be better aligned in time, but aligned worse in space as compared to trials 11 and 13.

Between trials 8 and 10, the average distance per trial is much larger for the dynamic group than for the static group. The most likely cause for this is that the static group can still follow the visible ghost instructor, while the dynamic group may be struggling to remember the path when the ghost instructor becomes mostly or fully invisible, causing the ghost avatar's transparency to oscillate between visible and invisible. This is further supported by Figure 7, in which a low mean transparency paired with large standard deviation is pictured for trials 8 through 10.

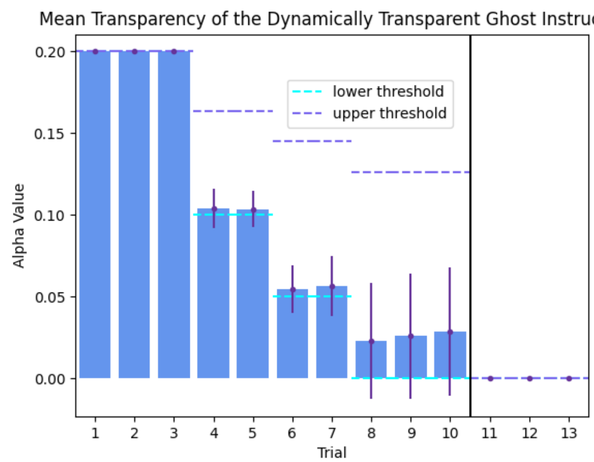

Figure 7: This figure visualises the mean alpha value (1 denotes fully visible, 0 denotes invisible) of the ghost avatar per trial. The standard deviation thereof is shown through purple error bars. Horizontal bars above the trial indexes mark the lower and upper transparency thresholds, and the training trials are separated from the test trials through a black vertical line.

Of particular interest are the test trials. For both Euclidean distance and DTW, the average distances of the test trials is much larger than that of the training trials. The absence of the ghost instructor, paired with the relatively complex task led to students struggling to reproduce the target motion. Furthermore, many students failed to complete the test trials (Figure 8), showing that they could not accurately find the end position within the given time. While most students seemed to struggle during the test phase, the static group shows higher mean distances and greater standard deviations for all test trials, indicating worse performance.

Using the mean distances per trial, the learning effect can be calculated per group. The learning effect is defined as the improvement compared to the first trial. In Figure 9 it can be seen that the training trial performance of the static group is never worse than that in the first trial although the performance does not always increase between two successive trials. This is not the case for the dynamic group as their performance does drop below that of the first trial. This is likely due to the lack of guidance caused by the fact that the ghost avatar can become invisible during these trials even if the imitated motion is not perfect due to the use of the error threshold.

The performance drop is negatively correlated with the

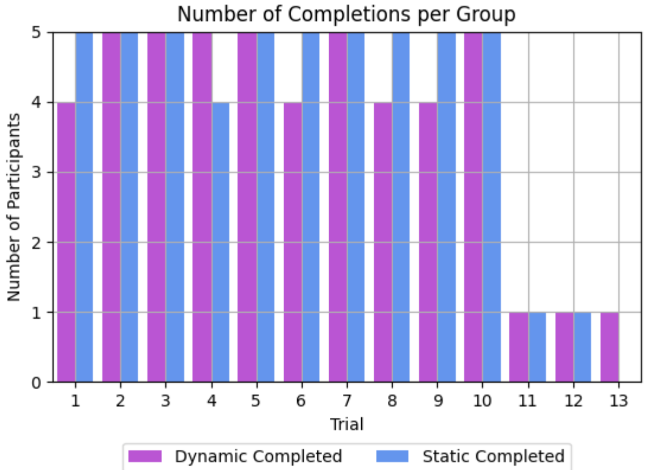

Figure 8: This figure visualises how many participants out of the 5 assigned to each group completed a trial.

Figure 9: This figure show the improvement with regards to the first trial per group per trial, based on the mean distances in Figure 6. It also shows the standard deviations of said figure as coloured areas around the improvement line. Negative y-values constitute an improvement.

group mean distance between the last training trial and the best test trial. The reason that the best test trial is used rather than the average of all test trials is so that the performance drop is related to the student's best test performance. From Table 2, one can draw the conclusion that the performance drop of the dynamic group is lower than that of the static group due to the negative correlation between performance drop and distance increase. This is the case for both training trial 7, and the last training trial.

	Dynamic (Trial 7)	Dynamic	Static
Euclidean	301.0%	154.9%	427.9%
DTW	268.1%	123.7%	424.1%

Table 2: The percentage distance increase between the last training trial and the best test trial. The percentage increase between the 7th trial and the best test trial for the dynamic group can also be seen.

5 Responsible Research

A large part of this research consists of the user study for which human subjects were recruited. To ensure that this was done in an ethical way, approval from the TU Delft Human Research Ethics Committee was obtained prior to starting the experiments. No personally identifiable data was collected, and participants were informed of the goal of the study and of their right to abort the study at any point before its conclusion. Furthermore, the log file names consisted of a Version 4 Universally Unique Identifier followed by a trial index, and did not contain any information that could be linked back to the participant.

As described in Section 4, 12 sets of log files were collected through the experiments. Of these 12 sets, 2 were linked to experiments that did not have one or more of the breaks at the designated time and were therefore not used in the results section. The traced curve, represented by a collection of virtual index finger positions, was transformed by mean-shifting the data. This was done because tracing the right shape was the primary task and precise location data was not required. Furthermore, not all users traced the shape in the exact same 3-dimensional area as the ghost instructor, so mean-shifting allowed for a more accurate comparison.

The figures created to visualise the results of the user study heavily relied on averages to ensure the figures did not become cluttered. While this makes identifying outliers quite difficult, an indication of the standard deviation has been included in relevant figures to ensure transparency.

The collected data has been uploaded to 4TU data repository [16]. The models, and the created C# and Python code files are available in a public Github repository [8].

6 Discussion

The results obtained from the user study were analysed and visualised in order to provide all necessary information to determine whether the 3 hypotheses posed at the start of the paper hold:

- H1 Using a dynamically transparent ghost instructor, as compared to a statically transparent ghost instructor, leads to a smaller performance drop between the training and testing phase.
- H2 Using a dynamically transparent ghost instructor, as compared to a statically transparent ghost instructor, leads to a smaller performance improvement between trials of the training phase.
- H3 Using a (static or dynamic) ghost instructor leads to an increase in performance between sequential trials of the training phase.

The obtained results lead to the following conclusions. Firstly, the performance drop of trials taught by a dynamically transparent ghost instructor are indeed lower than those taught by a statically transparent ghost instructor. Thus, H1 holds in the context of this study. Secondly, the improvement based on the first trial was much smaller for participants assigned to the dynamic group, even when comparing the 10th static trial to the 7th dynamic trial to avoid the distance spikes caused by low ghost transparencies in trials 8 through 10. Therefore, H2 holds as well. H3, however, does not hold for either group. The performance does not increase (distance does not decrease) between each consecutive trial although there does seem to be a generally increasing performance trend for the static group.

While interpreting these results, it is very important to keep in mind that the user study was a small-scale user study consisting of only 5 people per group. Therefore, the data is likely not representative of the wider population, and the distribution may change considerably if the sample size were to increase. The amount of virtual reality experience that each participant had was also not taken into account, but may explain part of the variability in the results.

The comparison study [6] mentioned earlier in the paper has a relatively similar experimental procedure. The procedure required participants to complete a trial alone first to set the baseline performance for the participant. This was followed by 5 training trials, a short break, and another 5 training trials followed immediately by 3 test trials. The lack of a second break makes these results unsuited to measuring retention. However, it is interesting to compare the performance drop between the training and test trials for the comparison study with the results obtained in Section 4.2. According to the comparison study [6, Figure 8], the improvement compared to the baseline of the last training trial and the test trial of the guidance type most similar to the ghost instructor (perspective sharing) is equal to 17 or 18, and 10, respectively. This leads to a percentage decrease in improvement equal to 41.1%-44.5%. This is much smaller than the percentage distance increase values (where distance is negatively correlated with performance and improvement) obtained for all groups listed in Table **??**. This may be due to a number of reasons: the presence of an extra break in between training and testing, the difference in task types (free hand as opposed to touching vertices in a specific order), or a difference in task complexity.

The developed system will require some corrections before future studies can be run. Firstly, as discussed in Section 3.1, the error that is used to vary the ghost avatar's transparency is based on the Manhattan distance between the positions of the student and ghost avatars' index fingers. This distance should instead be calculated using Euclidean distance to avoid overestimating the actual shortest-path distance, leading to lower transparencies. Secondly, the transparency modifier that is used to interpolate between the minimum and maximum transparency per trial is incorrectly scaled, causing the maximum transparency set per configuration mentioned in Section 4.1 to be higher than the maximum achievable transparency.

In addition to these corrections, there are some further adjustments that could be done to improve the quality of the obtained results and the usability from the student's perspective. One of these adjustments could be to reduce the error threshold to allow participants to more easily notice when they are not performing the motion correctly when making small mistakes. Ideally, this should be fine-tuned based on the results of a user study. There are also adjustments that could be made to the task difficulty and representation. As is evident from the low test scores for both the dynamic and the static groups, the participants did not have enough time to learn the relatively complicated motion. The complexity of the task should thus be reduced. Furthermore, the chained Bézier curves used to represent the target motion should be adjusted so that the ghost instructor does not make a jerking motion near the boundaries of curved segments, making the motion much easier to follow.

While there are many ways in which this system could be improved, these preliminary results seem to suggest that there may be merit in further exploring the topic of dynamically transparent ghost instructors in skill transfer.

7 Conclusions and Future Work

This research aimed to explore the effect of dynamically transparent ghost instructors on skill transfer in a virtual reality environment. Three hypotheses were developed:

- H1 Using a dynamically transparent ghost instructor, as compared to a statically transparent ghost instructor, leads to a smaller performance drop between the training and testing phase.
- H2 Using a dynamically transparent ghost instructor, as compared to a statically transparent ghost instructor, leads to a smaller performance improvement between trials of the training phase.
- H3 Using a (static or dynamic) ghost instructor leads to an increase in performance between sequential trials of the training phase.

To test these hypotheses, a system was developed with which 12 user experiments were conducted. The results of these experiments supported hypotheses H1 and H2, but not H3. Although the user study was small-scale, the support of hypothesis H1 shows that dynamically transparent ghost instructors may have a real benefit over the statically transparent variety.

Due to the small sample size, the results may not be generalisable to a larger population. Therefore, more user tests should be run with a similar setup to verify if this trend emerges with larger sample sizes as well. Furthermore, future research should compare different tasks with various difficulty levels to investigate if dynamically transparent ghost instructors can be applied to tasks in all, or only to tasks in a subset of, difficulty classes for maximum performance.

References

[1] X. Chen *et al.*, "ImmerTai: Immersive motion learning in VR environments," *Journal of Visual Communication and Image Representation*, vol. 58, pp. 416–427, Jan. 2019, ISSN: 10473203. DOI: 10.1016/j.jvcir.2018.11.039.

[2] R. Fribourg *et al.*, "Virtual co-embodiment: Evaluation of the sense of agency while sharing the control of a virtual body among two individuals," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 10, pp. 4023–4038, 2021. DOI: 10.1109/TVCG.2020.2999197.

[3] U. Yang and G. J. Kim, "Implementation and Evaluation of Just Follow Me: An Immersive, VR-Based, Motion-Training System," *Presence: Teleoperators and Virtual Environments*, vol. 11, no. 3, pp. 304–323, Jun. 2002. DOI: 10.1162/105474602317473240. eprint: https://direct.mit.edu/pvar/article-pdf/11/3/304/1623762/105474602317473240.pdf.

[4] H. Kawasaki, H. Iizuka, S. Okamoto, H. Ando, and T. Maeda, "Collaboration and skill transmission by first-person perspective view sharing system," in *19th International Symposium in Robot and Human Interactive Communication*, 2010, pp. 125–131. DOI: 10.1109/ROMAN.2010.5598668.

[5] M. L. Chenechal, T. Duval, V. Gouranton, J. Royan, and B. Arnaldi, "Vishnu: Virtual immersive support for helping users an interaction paradigm for collaborative remote guiding in mixed reality," in *2016 IEEE Third VR International Workshop on Collaborative Virtual Environments (3DCVE)*, 2016, pp. 9–12. DOI: 10.1109/3DCVE.2016.7563559.

[6] D. Kodama, T. Mizuho, Y. Hatada, T. Narumi, and M. Hirose, "Effects of collaborative training using virtual co-embodiment on motor skill learning," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 5, pp. 2304–2314, 2023. DOI: 10.1109/TVCG.2023.3247112.

[7] Unity Technologies. "Unity Real-Time Development Platform — 3D, 2D, VR AR Engine." (2024), [Online]. Available: https://unity.com/ (visited on 06/22/2024).

[8] Cassandra Visser. "thesis-vr-dynamic-ghost-instructors-2024." (2024), [Online]. Available: https://github.com/C-Hoek/thesis-vr-dynamic-ghost-instructors-2024.

[9] MakeHuman Community. "MPFB Downloads :: MakeHuman Community." (2024), [Online]. Available: https://static.makehumancommunity.org/mpfb/downloads.html (visited on 06/22/2024).

[10] Blender. "blender.org - Home of the Blender project - Free and Open 3D Creation Software." (2024), [Online]. Available: https://www.blender.org/.

[11] bogdan666, culturalibre, dariush086, Joel Palmius, MTKnife, MargaretToigo. "MAKEHUMAN_SYSTEM_ASSETS." (2024), [Online]. Available: https://static.makehumancommunity.org/assets/assetpacks/system_hair_materials01.html (visited on 02/06/2024).

[12] makehuman_system. "SYSTEM_HAIR_MATERIALS01." (2024), [Online]. Available: https://static.makehumancommunity.org/assets/assetpacks/makehuman_system_assets.html (visited on 02/06/2024).

[13] Elvaerwyn, Joel Palmius, namuhekam, skalldyrssuppe, MargaretToigo. "Shirts01." (2024), [Online]. Available: https://static.makehumancommunity.org/assets/assetpacks/shirts01.html (visited on 02/06/2024).

[14] Cortu, MargaretToigo. "Pants01." (2024), [Online]. Available: https://static.makehumancommunity.org/assets/assetpacks/pants01.html (visited on 02/06/2024).

[15] W. Meert, K. Hendrickx, T. Van Craenendonck, P. Robberechts, H. Blockeeland, and J. Davis, "DTAIDistance," Aug. 11 2020.

[16] C. Visser, *Data underlying research: Dynamically transparent ghost instructors and their effect on learning and skill retention in virtual reality environments*, Jun. 23, 2024. DOI: 10.4121/abb0629c-9f42-4a0f-917f-d5d666554738.

[17] EasyBrain. "Hardest sudoku puzzles - play sudoku expert level online for free." (2024), [Online]. Available: https://sudoku.com/expert/.

[18] EasyBrain. "Extreme sudoku - solve extremely difficult puzzles for free." (2024), [Online]. Available: https://sudoku.com/extreme/.

A Evaluation Documents

In this appendix, the opening statement and the Sudoku puzzles handed to the participants can be found.

A.1 Opening Statement

The opening statement has been approved by TU Delft's Human Research Ethics Committee and can be found below.

Opening Statement

You are being invited to participate in a research study titled "***Dynamically Transparent Ghost Instructors and Their Effect on Learning and Skill Retention in Virtual Reality Environments***". This study is being conducted by ***Cassandra Visser from the TU Delft, supervised by Amir Zaidi (supervisor) and Prof. Dr. Ricardo Marroquim (responsible professor) from the TU Delft.***

The purpose of this research study is to evaluate the ***effect of dynamically transparent ghost instructors on skill transfer through an experiment in Virtual Reality.*** The experiment will take you approximately 25 minutes to complete, ***and will focus on teaching you to perform a skill. To measure the effect of the ghost instructor, data such as in-game hand movement/position and task performance will be logged throughout the experiment. The data will be used to evaluate the developed VR system in addition to being used for a scientific paper with the purpose of obtaining a Bachelor's degree in Computer Science and Engineering, and may be submitted to a conference.***

We will minimise any risk of identification ***of individuals by collecting no personally identifiable information, and only logging details provided through the VR system's interface. The data will be logged in CSV format which will be uploaded to TUD TopDesk servers.*** The best efforts are made to ensure that the data is ***anonymised*** and handled in a secure manner, but there always remains a risk that a data breach ***occurs or that*** the participants are identified.
The data you provide may be published in openly available academic journals, **and will be archived in a data repository that can be shared with the public**, in accordance with the TU Delft Research Data Framework Policy and TU Delft Open Science Guidelines.

Participation in this study requires you to use a Virtual Reality device. Whilst these devices are considered safe for general use, you may experience some side effects such as nausea, headaches, or dizziness. If you experience any discomfort, you should alert the researcher(s) immediately so the experiment can be stopped and the device can be removed.

Your participation in this study is entirely voluntary and you can withdraw at any time. ***You may abort the study to have all collected data removed and to ensure that no information is saved.*** Once the experiment has ended, and you have submitted a response, it will no longer be possible to withdraw your data from the study.

By participating in this study, you are agreeing to this Opening Statement and the methods of processing your submitted data as outlined.

A.2 Sudoku Puzzles

All three Sudoku puzzles were obtained from Sudoku.com. The expert puzzle [17] and the two extreme puzzles [18] are shown in the figure below. Two puzzle trios were printed on the same page, which was cut in half so that each participant received one of each unique puzzle. The type of the puzzle was written at the top left of the puzzle in handwriting. The types were: Expert-1, Extreme-1, and Extreme-2. The puzzles can be seen on the right side of this page.

The puzzle on top is the Expert-1 puzzle. The ones below the top puzzle are Extreme-1 and Extreme-2, in that order.

None of the participants managed to finish all three puzzles within the 10-minute time limit.