Virtualising the Nine Hole Peg Test of Finger Dexterity

Jonathan Collins¹, Simon Hoermann¹², Holger Regenbrecht¹

¹Department of Information Science, University of Otago, Dunedin, NEW ZEALAND
²Department of Medicine (DSM), University of Otago, Dunedin, NEW ZEALAND

jonnymcollins@gmail.com, simon.hoermann@otago.ac.nz, holger.regenbrecht@otago.ac.nz

¹www.otago.ac.nz/informationscience, ²www.otago.ac.nz/dsm-medicine,

ABSTRACT

Using Virtual and Augmented Reality (VR/AR) approaches in physical rehabilitation can lead to better controlled, more client motivating, and more flexible forms of therapy. The Nine Hole Peg Test (NHPT) is a standard instrument in physiotherapy to practice and assess a patient's hand motor control abilities. A physical, wooden or plastic board with nine holes and cylindrical shaped pegs are used to perform this task. There are only limited ways of varying the degree of difficulty or to precisely measure progress with this physical setup.

This study presents the development of a VR/AR version of the NHPT and evaluates the usability of three versions: (1) the real life wooden version, (2) a video-mediated version and (3) a computer-generated AR version built from low-cost off-the-shelf components.

Our results show that all three conditions were successfully completed by all participants with the highest measured performance and perceived usability still achieved in the real life situation. This indicates that the implementation of currently available low-cost, off-the-shelf components is not yet reliable enough to suggest its use for therapeutic exercises or assessments that require very fine finger level interaction.

1. INTRODUCTION

Is a virtualised Nine Hole Peg Test as convincing in terms of usability, as a real version? This is the primary question investigated in this study. The Nine Hole Peg Test is a tool for the therapeutic assessment of finger function and is commonly used with people who suffer from impairments after stroke (Mathiowetz, Weber, Kashman, & Volland, 1985). The purpose of a virtual version of the test is to allow a broader range of therapeutic applications as well as a more patient-based adaptation than the traditional test. For example, the difficulty could be adjusted based on the patients' performance and frustration tolerance as well as their motivation, this also allows patients with severe impairments to be treated who otherwise would not be able to perform the test.

The development of the virtual Nine Hole Peg Test (vNHPT) requires new hardware as well as software components. The general concept is based on Augmented Reflection Technology (ART) introduced by Regenbrecht et al. (2011) and used for a number of studies with healthy participants (Hoermann, Franz, &
Regenbrecht, 2012; Regenbrecht et al., 2012; Regenbrecht, McGregor, et al., 2011) as well as with clinical participants (Hoermann, Hale, Winser, & Regenbrecht, 2012). For the specific implementation of the vNHPT however, more sophisticated tracking and rendering approaches are necessary.

In current rehabilitation, there are several approaches to help the patients gain back some of their motor functions. Among the most common is physiotherapy following the Bobath concept (Lennon, 2003), which often includes the use of external devices to support the patients in their execution of movement tasks. Another approach is Constraint-Induced Movement Therapy proposed by Taub, Uswatte & Pidikiti (1999). This involves restraining the healthy limb of the patient, and having them perform actions with their impaired limb. Doing so for extensive periods of time (i.e. up to 90% of waking hours) has been shown to improve motor deficits of patients suffering from impairments after stroke (Miltner, Bauder, Sommer, Dettmers, & Taub, 1999).

A less restraining approach is one which takes advantage of the manipulability of human perceptions, beliefs and even sensations. It was in fact shown that psychotherapies such as Cognitive Behaviour Therapy, involving only talking, have effects on the brain (Straube, Glauer, Dilger, Mentzer, & Miltner, 2006). Similar changes in the brain were also shown in a stroke patient treated with Mirror Visual Illusions (Michielsen et al., 2011).

This phenomenon is commonly referred to as “neuroplasticity” and is described as the brain’s ability “to respond to intrinsic and extrinsic stimuli by reorganizing its structure, function and connections” (Cramer et al., 2010). In order to make best use of it, therapy approaches should focus on providing environments that allow meaningful therapeutic movements, with adequate intensity and repetitions, as well as motivating the patient and providing appropriate feedback (Holden, 2005). Virtual and Augmented Reality Environments have the potential to be used in this context.

In this paper an implementation of such an environment is presented and compared with its real life counterparts.

2. SYSTEM

There are three main technical components and the physical apparatus itself that contribute to the system. (1) An off-the-shelf webcam with a built in 3D depth sensor with a resolution of 320x240, and an HD 720p RGB image sensor (Interactive Gesture Camera, Creative Technology Ltd) mounted on a custom build frame (Fig. 2), (2) a tailor-made plugin to process the data from the webcam for delivery to the application, and finally (3) a virtual reality application created using the Unity3D game engine (version 4.2, unity3d.com) which provides the final environment in which the users perform their tasks in.

The webcam’s functions are accessed from the plugin using the Intel Perceptual Computing SDK 2013 (software.intel.com/en-us/vcsource/tools/perceptual-computing-sdk). This provides access to the raw data from both the depth and the colour sensors and provides features such as basic finger tracking.

The hardware “therapy frame” (Figure 2 left) where the webcam is mounted, consists of a flat board with a metallic frame attached to the front of it. On the top of the frame, the webcam (described above) is attached and points toward the board at a 45 degree angle. A black curtain in front of the frame prevents the user from seeing the real interaction (Fig. 2 right). This is to direct the participants’ attention to the interaction shown on the screen and to maintain the illusion of interacting in the virtual space during the tasks. A blue fabric is used to cover the base.

![Figure 2. Metal Frame used to position the depth cam without curtain (left) and with the curtain to prevent the direct view of the hand during use (right)](image-url)
2.1 Finger Tracking

The target action required for task completion in this study is a grabbing action where the participant grabs a peg and places it in the board using the index finger and the thumb (Figure 1). For this, only two coordinates need to be tracked which are the x, y, and z coordinates of the thumb and index finger. First the blue background (the fabric covering the table) is being subtracted from the video image leaving only the pixels representing the hand. The colour blue is used because in the (HSV) colour space, blue is the closest opposite to the average skin colour. Then we traverse the remaining image (which is now containing only the hand), starting with the top left pixel moving right, and then down until finding an opaque pixel (not made transparent by the background subtraction method). With this, the first fingertip is found, then by ignoring all pixels below the initial point found, and either side for a threshold of 45 pixels, resuming the search will result in finding the second fingertip. The coordinates of these two points are stored and their depth values are retrieved using the Intel SDK. The Unity3D plugin uses these computed coordinates to control the interaction with the virtual environment.

2.1 Virtual Environment

The graphic engine Unity3D was used to create and display the environment and handle the interactions with the objects in this environment. Within Unity3D, C# scripts were programmed which retrieve the coordinates of the fingers and import the video image of the hand into the virtual scene from the plugin. For each frame, the plugin function is called and copies the image data of the users hand as a texture to a virtual plane, and at the same time the two 3D coordinates of the finger and thumb are retrieved. Since the blue background of the hand images was removed the user gets the impression of seeing the own hand in the virtual environment.

The virtual NHPT model in Figure 1 was created in Google Sketchup Make (version 13). This model was exported as a Collada model and then imported directly into Unity3D. The way we use the fingertip data to interact with the peg models is by checking three conditions. First we find the midpoint between the two fingertips, and we cast a virtual, invisible ray through that point and check if that ray collides with any peg. If it does, we then calculate the Euclidean distance between the two fingertips, and if the distance is small enough (to represent the grabbing gesture), then the third check is performed which is testing if the depth coordinate of the two fingertips is equal to that of the peg which the ray is colliding with. When all three of these conditions are satisfied, the peg will attach itself to the midpoint and will move with the fingertips. Placing the peg in the hole of the virtual board utilises a sphere collider (invisible/un-rendered) placed in each hole, and if the peg that was being moved collided with the sphere collider in the appropriate hole then the peg releases itself in that hole.

In order to prevent the pegs from being moved outside of the visible area, a condition was added that limits the working environment and if this condition is violated the peg that violates this condition is returned back to its initial starting position.

3. METHOD

The virtualized Nine Hole Peg Test (vNHPT) was implemented and compared to the original wooden Nine Hole Peg Test (NHPT). In three experimental conditions the vNHPT was compared to two conditions of the traditional NHPT: (1) the original NHPT performed with direct vision, and (2) the NHPT mediated through the webcam and computer system but using the original wooden components.

3.1 Participants

Eighteen participants were recruited from the University of Otago. The sample consisted of 9 male and 9 female students from a range of disciplines, and between the ages of 18 and 25 years. All participants provided written informed consent and received a $10 grocery voucher as compensation for their time.

3.2 Measures

The traditional Nine-Hole Peg Test (NHPT) kit used for comparison was made from a piece of wooden board and has nine holes drilled in it evenly spread apart. The nine pegs were cut to equal length from a piece of wooden dowel. The test kit was made according to the standard described in Mathiowetz et al. (1985).

There were two questionnaires involved in this experiment, a demographics and a usability questionnaire. The demographics questionnaire was first given to the participants requiring information such as age, gender, handedness, possible vision impairments, physical well-being, previous augmented reality experiences, and previous involvements in similar experiments. After completing the tasks the usability questionnaire was presented evaluating their experience with the system. This questionnaire was divided into three sections to be filled out after each condition.
The usability questionnaire was composed of questions from the Mixed Reality Experience Questionnaire (Regenbrecht et al., 2013). Some questions were modified slightly as to fit the nature of the experiment. The questionnaire can be divided into two main parts. There were 13 questions in total, nine of which can be categorized as direct usability assessment of the condition, and 4 of which are assessing the environment surrounding the condition. There were five questions to assess the task of physically reaching, grabbing, moving, placing and releasing the pegs when performing the test. Each of the questions were measured on a Likert scale (1 – 7) with 1 being “strongly disagree” and 7 being “strongly agree”. As well as having a questionnaire to evaluate user performance, each condition was timed using a stopwatch to measure the completion time.

3.3 Design

The experiment uses a within-subject design with the 18 participants pre-randomised and counterbalanced across the three conditions. The independent variable consists of the three conditions of the NHPT, and the dependent variables are time to complete the task, user satisfaction, and perceived performance.

3.4 Procedure

Experiments were run in a controlled lab environment (Computer-Mediated Realities Lab) to reduce unnecessary distraction for the participants. In total three conditions were evaluated: real life (RL), video mediated (ME), and augmented reality (AR) versions of the NHPT.

Upon their arrival participants were greeted and given an information sheet detailing the experiment and what they should expect. After reading this, they were presented with a consent form to give their formal consent. They were then shown their first condition and timed with a stopwatch. After each condition participants had to complete the usability questionnaire regarding their experience. Participants repeated this procedure for all three conditions.

In the RL condition, the wooden board was placed on a table in front of the participant (see Figure 3 (left)) and the users were instructed to use their left hand to transfer the pegs one by one to the holes. In contrast to the original NHPT, the holes on the board were numbered in the order in which the participants were to move the pegs to. The reason for this was to keep the tasks as similar as possible for each condition and in this case slightly adapt the real world NHPT procedure to the virtualised version. When the user picked up a peg, a hole would light up (green) on the board to show which hole to place the peg in.

Another small modification from the original NHPT protocol, again to retain tasks as similar as possible between conditions, was that the pegs starting position was standing upright in a second real board. This board replaced the box where the pegs would be lying in the original version of the test and the users are meant to grab the pegs from that box. Pegs in both the virtual and the real space were constrained to an upright starting position.

Figure 3. Photos of a participant exercising in the three conditions: real life RL (left), video mediated ME (centre) and virtual VR (right).

The Video-Mediated (ME) condition involved having the real NHPT placed in the exact same manner within the apparatus as the virtual one (see Fig. 3 centre). The participants were instructed to complete the test by moving the pegs from the initial board to the final peg board one by one, again using their left hand, except for this condition they are allowed to move the pegs to any hole they choose. This was because it was too difficult to see the number labels on the peg board, and it was decided that it was less confounding than to ask the participant to remember the order of the holes. In this condition the user were allowed to observe only the scene on the monitor, see Fig. 4 (left), while the NHPT was hidden from their direct view.

The AR condition, see Fig. 3 (right) had again the participant sitting at the apparatus and referring only to the scene shown on the monitor. The task was the same as in the other conditions; participants had to place all pegs
one by one into the board. When a peg was grabbed, the peg turned green, and a hole lit up to indicate where to place the peg see Fig. 4 (right). Before users were to complete the AR condition, they were shown the environment, and given a small time to navigate the space and interact with 3 virtual pegs. This was to accustom the user to the new environment and reduce a possible so called “wow-effect” with new technologies.

After completion of the third and final condition and after filling in the usability questionnaire, participants were thanked, compensated with the grocery voucher and released.

Figure 4. Monitor screenshots of ME (left) and AR (right)

3.5 Statistical Analysis

Data analysis was carried out in SPSS version 21. A 95% confidence interval was used. First the questionnaire data was checked for normal distribution using the Shapiro-Wilk method. This test returned a significant result for the real life condition ($p < .001$), but not for the video mediated and virtual conditions of ($p = .875$, $p = .970$), showing the real life condition is not normally distributed. This was expected because almost all of the questions were designed to cater for all three conditions. The distribution of the values in the real life condition showed that they were very lopsided with a large majority of usability questionnaire answers resulting in an answer of 7. Following this, non-parametric tests needed to be applied on the data. First a Related-Samples Friedman’s Two-Way Analysis of Variance by Ranks was applied across all questions for each condition. If significance was found, Related-Samples Wilcoxon Signed Rank test was applied to the data to determine if the differences between conditions were significant. The analysis of bivariate correlations used one-tailed Kendall’s tau-b correlations coefficient. The ratings for Q13, “I had the impression of seeing the pegs as merely a flat image”, were inverted prior to the data analysis to align it with the other questions.

4. RESULTS

4.1 Overall Combined Scores

As expected the RL condition returned the highest values with $M = 6.69$ ($SD = 0.368$, $IQR = 7–7$). The ME condition closely followed with ($M = 5.01$, $SD = 1.023$, $IQR = 4–6$). Questions for the AR condition returned lower values with ($M = 3.88$, $SD = 0.824$, $IQR = 3–5$). The non-parametric tests applied to this data showed significant differences ($\chi^2 = 2$, $p < .001$).

4.2 Task

Similar to the overall questionnaire results, RL returned the highest values for the nine questions regarding the task itself with values of $M = 6.70$ ($SD = 0.393$, $IQR = 7–7$). The ME and AR returned values of ($M = 5.18$, $SD = 0.954$, $IQR = 4–6$) and ($M = 3.89$, $SD = 1.01$, $IQR = 3–5$) respectively. When the non-parametric test is applied to the task questions we receive results of ($\chi^2 = 2$, $p < .001$). Again a strong significance value is found which supports a large difference in the performance of each task.

4.3 Environment

The four questions regarding the participants’ perception of the environment returned results in the same order with RL > ME > AR (RL: $M = 6.68$, $SD = 0.451$, $IQR = 6–7$; ME: $M = 4.73$, $SD = 1.40$, $IQR = 3–6$; AR: $M = 3.88$, $SD = 0.710$, $IQR = 3–5$) respectively. Non-parametric results give ($\chi^2 = 2$, $p < .001$). When using Related-Samples Wilcoxon Signed Rank Test to compare each of the conditions, there was significance found between all of the conditions with both RL-ME and RL-AR giving values of ($p < .001$). There was however less significance found between the ME and AR condition as the graph in Figure 6 suggests with a value of ($p = .015$).

4.4 Single Question Comparison
The results for each individual question of the three conditions is shown in Table 1. It shows that in the AR condition, participants rated Q1, Q2, Q6, Q8, Q9, and Q10 significantly below ($p < .05$) the neutral midpoint at level “4”. In contrast Q3, Q12 and Q13 were rated significantly positive by the participants. This could indicate that they did not have any negative experiences in these parts.

### Table 1. Results of questionnaire (results significantly above neutral midpoint are highlighted in green and results significantly below in red)

<table>
<thead>
<tr>
<th>Question</th>
<th>RL</th>
<th>ME</th>
<th>AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1 - It was easy for me to reach the pegs</td>
<td>Mean 6.89</td>
<td>SD 0.32</td>
<td>IQR 7–7</td>
</tr>
<tr>
<td>Q2 - It was easy for me to grab the pegs</td>
<td>Mean 6.83</td>
<td>SD 0.38</td>
<td>IQR 7–7</td>
</tr>
<tr>
<td>Q3 - It was easy for me to move the pegs in the board</td>
<td>Mean 6.94</td>
<td>SD 0.24</td>
<td>IQR 7–7</td>
</tr>
<tr>
<td>Q4 - It was easy to perform the task overall</td>
<td>Mean 6.44</td>
<td>SD 0.78</td>
<td>IQR 6–7</td>
</tr>
<tr>
<td>Q5 - It was easy for me to release the pegs</td>
<td>Mean 6.83</td>
<td>SD 0.38</td>
<td>IQR 7–7</td>
</tr>
<tr>
<td>Q6 - It was easy for me to grab the pegs</td>
<td>Mean 6.67</td>
<td>SD 0.77</td>
<td>IQR 6.75–7</td>
</tr>
<tr>
<td>Q7 - It was fast in completing the task</td>
<td>Mean 6.61</td>
<td>SD 0.78</td>
<td>IQR 6–7</td>
</tr>
</tbody>
</table>

* *inverted values*

### 4.5 Comparison of Times

The completion times were checked for normality using the Kolmogorov-Smirnov test with both the RL and ME conditions sitting within a normal distribution with values of ($p = .157$) and ($p = .066$) respectively, however the AR condition resulted outside of normal distribution with a significance value of ($p = .002$). The AR condition returned by far the highest values with ($M = 167.94$, $SD = 116.73$) followed by the ME task with values of ($M = 48.34$, $SD = 19.28$) and finally the lowest values in the RL condition with ($M = 13.55$, $SD = 2.3$).

Given that one condition was outside of normal distribution, we used Related-Samples Friedman’s Two-Way Analysis of Variance by Ranks to analyse the data. This returned values of ($\chi^2 = 2$, $p < .001$) showing again significant difference between conditions.

### 4.6 Correlations between conditions

The analysis of correlation between the more similar conditions showed a tendency with a positive correlation of the time used between the RL and the ME condition ($\tau = .262$, $p = .065$) and the ME condition to the VR condition ($\tau = .255$, $p = .07$). The correlation between RL and VR was not significant ($\tau = .170$, $p = .162$).

### 5. DISCUSSION AND CONCLUSIONS

In this study we demonstrated that the NHPT can be virtualised, although it is not yet as convincing as the real world test in terms of usability. The results show significant differences between each of the conditions. Participants found the RL condition easier than performing the ME condition. This could be due to the positioning of the camera and screen (see Fig. 3) as well as the fact that users see a 2D version of their own hand performing the test. This could have made it hard for them to see the holes on the board. Furthermore, when the users perform the RL scenario, they have the test directly in front of them, whereas the viewing angle (due to the position of the monitor) could contribute to further disorientation/difficulties when completing the ME and VR conditions. It was observed that users would face their body towards the monitor and perform the actions holding
their arm out to the left (see Fig. 3). When comparing the users’ view of the ME and VR scenarios (see Fig. 4), there is a slight difference between the perspectives. The boards appear to be at different angles which could also be contributing to users’ difficulties due to inaccurate depth perception.

When performing the virtual version of the test, it was observed that when participants tried to move their arm in depth to reach the pegs, they would move horizontally forward in real space. Due to the angle of the camera relative to the table top, the depth sensor does not sense the users’ forward action as purely moving away from the user. This causes the virtual “fingertip spheres” to move within the environment in a perceptually incorrect way. For example, the spheres will not move in as much depth in the virtual space as the user is moving in real life. For this reason, some participants had difficulties picking up pegs and placing them. Results showed that users found placing the peg on the board much easier than grabbing the peg. Furthermore, the camera used is developer hardware and software which meant that in this case, the data retrieved from the SDK was somewhat unreliable. To the participants it was noticeable in the AR condition when the depth camera temporarily faulted, because if a depth coordinate was not supplied, then some default value was used. Unfortunately, this just made the peg move back to its starting location.

The time required to complete the conditions showed that there was a large variance between participants when they used the vNHPT. The real life NHPT was significantly easier to perform than the vNHPT. There is evidence though that not all parts of the vNHPT conditions contributed equally to this difference. This was shown by the results of the ME condition which were not significantly different from the vNHPT condition in terms of the environmental perception questions. In fact the mean values of the environment questions in the ME condition were only slightly higher than in the AR condition. Therefore the display and execution of the task by just observing the screen seemed to possibly have negatively influenced the performance of the participants. This should be addressed in future research by optimizing the display condition.

The results from the questionnaire suggested various areas of possible future improvements of the virtualised condition. Apart from the task of placing the peg in the virtual board, most tasks were identified to be significantly harder compared to the other conditions, notably the RL condition. It was easier for participants to place the pegs in the virtual board than it was to place them in the board in the ME condition. The question that gave the lowest response was the more general question about the handling of the pegs and whether it felt natural to the user. There were some positive aspects such as the task of moving the pegs from one location to another. This was expected given that the peg attaches itself to the midpoint between the fingertip spheres once the conditions for picking up the peg are satisfied. The 3D aspect of the condition was also identified easily by users.

It is important to note that a possible limitation of such an implementation is the obvious lack of haptic feedback within the augmented environment. With question 10 “The handling of the pegs felt natural to me” gaining the lowest score with regards to the AR condition, it is likely that the aforementioned limitation of not being able to feel the peg had an effect on the results of this question. Either directly or indirectly, this is could also have affected the users’ performance in the AR environment.

The hardware setup for this research placed the users’ monitor off to the side next to the camera-frame, tracking the users’ hand. This meant that the participants were looking in a different direction to where the action was occurring, which could potentially have affected the users’ feeling of presence, comfort, and performance. This could be overcome by using a hardware setup similar to the ART system (Regenbrecht et al. 2011) which places the monitor directly in front of the user and therefore helps the users have the experience as if they are looking at their hands more directly.

There is also considerable potential for improvements to be made at the technical and implementation levels of the virtualisation of the NHPT. As stated above, the depth information retrieved through the Intel SDK was somewhat unreliable. Also, the finger tracking module could be improved, e.g. by making better use of the depth information in conjunction with the colour image. The difficulty here is that the colour image provided by the SDK is not only of a higher resolution (1280 X 720) than the depth image resolution (320 X 240) but they are even different in their aspect ratios. There are various other tracking methods available which could potentially provide more reliable tracking data, however, most of these devices or methods require the users’ hand(s) to have an instrument attached in some way (i.e. data gloves). The idea of our rehabilitation scenario is that it provides the users with a natural interface so to facilitate the users’ feeling of presence in the environment. Data gloves could provide a reliable stream of data but then the user is “wired” to the computer. An advantage of having an un-instrumented system as presented here is that users are able to observe their “real” hands in the virtual environment, which potentially facilitates the users’ presence in the augmented environments.

As a virtual environment is adaptive in nature, this could be utilised to modify the NHPT for different users. For example, the board and pegs could be made bigger to make picking them up and placing them much easier for a user with less mobility and motor control. It would also be possible to scale movement so that it appears that they are moving the peg further than they are really moving their arm. Different tasks could be implemented...
such as changing the order of the holes which the pegs should be placed in, or increasing and decreasing the number of holes. These are just examples of adaptations which can be made to the vNHPT application.

Time and distance measures can also be put in place in the application which can accurately record both completion time, and distances. These kinds of data can be analysed further by physiotherapists and used for motivation of patients. It is also possible to record the task being completed so it can be further observed and analysed.

Hybrid approaches can also be implemented with the possibility of using for example the real NHPT board but virtual pegs. The camera approach also comes with its flaws, most of which are of a technological nature. The Intel development software is still flawed and is still being updated. The background subtraction could also be improved as the current version is compromised if there is too much natural sun light on the apparatus.

Acknowledgements: We would like to thank the participants for taking part in this study as well as the staff who helped us. Thanks also to the Department of Information Science for funding the research. Thanks to Patrick Ruprecht for his input and technical support. This study was part of the thesis work of the first author, supervised by the last author. This study was approved by the University of Otago Ethics Committee.

5. REFERENCES


