Motor Learning in a Goal-Oriented Visuospatial Task

Written by

Erin Williams

Research Advisor and Mentor: Lewis Wheaton, Ph.D.

Presented to the College of Sciences

Georgia Institute of Technology

In Partial Fulfillment of the Requirements

For the Research Option

17 December 2021

Motor Learning in a Goal-Oriented Visuospatial Task

Approved by

Dr. Lewis Wheaton, Principal Investigator

School of Biological Sciences

Georgia Institute of Technology

Dr. Mary Holder, Advisor

School of Neuroscience

Georgia Institute of Technology

Table of Contents

ABSTRACT

There have been numerous studies that investigate motor learning at large, but there is a lack of research focusing on three-dimensional visuospatial learning and action observation in the setting of a goal-oriented motor task. There are even fewer that test these variables while introducing a social component in which the subject must execute motor control based off another person's directed movements. The objective of this study is to investigate how factors, such as action observation, social intention, motor control, and goal-oriented behavior impact motor learning of a subject during a structure building task. Twelve right-hand dominant subjects engaged with the researcher during two rounds (each consisting of 15 trials) with each round resulting in a static structure that the subject constructed by imitating the movements of the researcher. Half of the subjects completed this paradigm while wearing a transradial body-powered prosthetic simulator device on their right arm (experimental group) to test if the prosthesis altered motor learning. The research questions aimed to assess subjects' ability to display evidence of motor learning throughout the task, if this evidence is associated with increased gaze position in the researcher's quadrant, and if these trends remain consistent when subjects are wearing the prosthesis. Eye-gaze patterns, task completion time, performance and task errors, and behavioral observations were used as methods of data collection, and statistical analyses, including t-tests, sample means, and surface distributions, were performed to evaluate the hypotheses. Results revealed subjects in both groups demonstrated motor learning between round one and two, and while the addition of the prosthesis increased task completion time and error values, subjects within the experimental group were also able to demonstrate significant decreases in the latter round. In addition, eye-tracking data revealed increased gaze patterns on the researcher's path of movement vs. the quadrant itself for the experimental group compared to the control group. This suggests that social intent and action observation are likely facilitating an increase in motor learning in subjects tasked with completing goal-directed movements with the unfamiliar prosthesis.

Keywords: cognitive motor control, action observation, social intention, motor learning

1. Introduction

The ability to learn and successfully perform new sequences of movements, known as motor learning, is kinematically and neurologically complex yet can appear innate at surface level. For example, if learning how to cook eggs on the stove, the learner would observe the chef using the spatula to flip the eggs until fully cooked. When it is time for the learner to take over and try this action for themselves, they would likely grasp the spatula in the same way and imitate the movements of the chef. This task may appear as if it requires little thought, but it involves the acquisition of a new skill. Acquisition is a key measurement of motor learning and refers to the initial performance of a new motor skill (Muratori et al., 2013). Acquiring a new skill often involves observing someone perform the action before attempting to imitate their movements. Previous studies have indicated that visual cues strengthen motor learning (Cusack, 2014) and can lead to increased performance in a task (Scorolli et al., 2014; Bayani et al., 2019). Another key component of motor learning is the transfer of skills. Referring to the egg example, if the learner were to next make pancakes with a spatula, then this would require similar movements but with a different end goal. They would likely still grasp the spatula in the same way and alternate their hand in prone and supine positions to cook the pancake, but the timing and outcome of the movements would be different. If the learner applied the initially acquired skill set to cooking the pancakes, then they demonstrated motor learning. This transfer of skills involves completing a movement similar but different than the original task practiced in the acquisition step (Muratori et al., 2013). These principals of observing, acquiring, and transferring a skill are the foundation for the present study.

The goal of this study is to evaluate principals of spatial learning (action observation, social intent, mental transformation) in the setting of a motor learning paradigm and to see if adding the prosthesis will also interfere with the degree of learning. This, in part, will be evaluated through two rounds of testing with each round resulting in the construction of a unique static structure. The change in structure from round one to round two requires subjects to display motor learning as altering the structure after several trials may interfere with the slope of learning (Zacks & Michelon, 2005). The slope of learning can be evaluated through the task completion time and frequency of errors. In addition, subjects will be spatially challenged as they must build each non-mirrored structure by imitating the researcher with ~180° degree of rotation. Social intent is a widely studied concept, though its specific mechanism in motor learning is understudied (Scorolli et al., 2014; Wheaton, 2017). However, it is known that the impact of perceiving the actions of others in order to replicate similar movements is essential for successful skill transfer (Muratori et al., 2013). The present study aims to test the impact of social intent and its impact on task completion time.

In attempt to address the lack of data investigating motor learning in the context of a three-dimensional visuospatial task, the present study sought to answer four specific questions: Do subjects display evidence of motor learning from round one to round two when comparing average task completion times? Does the observed trend in average task completion time remain consistent for the experimental group in which subjects are wearing an upper limb prosthesis? Do subjects in the experimental group display evidence of motor learning from round one to round two when comparing total number of errors? Do subjects in the experimental group demonstrate increased gaze position on the path of movements to and from their quadrant and the researcher's quadrant in comparison to the control group? It is predicted that subjects in both the control and experimental group will show a decrease in the average task completion time between round one and round two, suggesting there is a similar trend for both groups, and thus, displaying evidence of motor learning in this paradigm. It is predicted that subjects in the experimental group will have a significant decrease in the total number of errors between round one to round two, while there is not expected to be a significant decrease in total error count for the control group. It is also predicted that the subjects in the experimental group will show increased gaze position on the path of movements to and from their quadrant as well as the researcher's quadrant when compared to subjects in the control group, who are expected to not have significant gaze positions in areas where the path of movement occurs.

2. Materials and Methods

2.1. Subjects

There were a total of 12 healthy right-handed subjects (ages 18-65) that voluntarily participated in this study. Nine of the subjects were female. Subjects had a mean age of 34.25 years with four of twelve between the ages of 55-65 and eight of twelve between the ages of 18-24. All the participants reported having normal or corrected to normal vision and have not suffered from central nervous system injuries. All participants were unaware of the purpose of the experiment, and each attended one experimental session with a three-hour duration. Signed informed consent was obtained from every participant according to the procedures of the Institutional Review Board at the Georgia Institute of Technology.

2.2. Materials

A binocular Pupil Labs Eye Camera (Berlin, Germany; 200 Hz) was used to collect gaze pattern data throughout the course of the experiment. The two pupil cameras were attached to each side of the eye tracking glasses, and subjects were fitted for this device prior to beginning the study. The pupil cameras tracked the movements of each pupil in the subject’s field of vision. Additionally, a world view camera was attached to the apex of the device, and it displayed a view of the workspace containing markers that were placed to assist in collecting data of the subject’s gaze path. Manual marker calibration was used to calibrate each of the three cameras relative to the workspace to ensure accurate data was obtained while the subject completed the task. The program PupilCapture (Kassner et al., 2014) was used to record data from the eye tracking device. PupilCapture uses the cameras to detect the pupils, track gaze, track markers in the workspace, and record real-time events (Kassner et al., 2014). An associated program, PupilPlayer (Kassner et al., 2014), was used to replay the recordings, define areas of interest for gaze patterns, and output the data into an Excel file.

Subjects in the experimental group performed the paradigm while wearing a transradial body-powered prosthesis (Cusack et al., 2014) that fit over their sound (intact) right upper limb. This device was created to mimic the type of prosthesis an individual with an upper limb amputation at the forearm would use to complete basic prehension tasks (Bayani et al., 2019). The prosthesis contained a harness loop that was placed through the left arm of subjects, thus, making it a body-powered design. The device was mechanically opened with glenohumeral flexion and scapular abduction (Bayani et al., 2019). The prosthesis itself was worn on each subject’s right arm and had three velcro pads located over the hand, wrist, and upper forearm and a voluntary opening hook (Figure 1).

Figure 1a. Full view of prosthesis, including harness

Figure 1b. Side view of right arm in the prosthesis with the hook partially opened

2.3. Cognitive assessments

Subjects underwent a series of assessments prior to completing the experimental paradigm. These assessments are designed to measure visuospatial abilities, mental representation, working memory, and the ability to manipulate perceptual context. Each subject completed the four following cognitive assessments: mental rotation test (Ganis & Kievit, 2015), mental transformation test (Ehrlich et al., 2006), Stroop test (Rouger, 2020), and Sternberg memory task (Padgett, 2018). A fixed set of instructions for each test were read to the subjects prior to the beginning of every assessment. Administration of cognitive assessments and completion of the paradigm occurred in a singular session.

2.3.1. Mental Rotation Task

The mental rotation task is designed to test the visuospatial capabilities by having subjects determine if the image on the right is a rotation of the image on the left or if the two structures cannot be brought into alignment through means of rotation (Ganis & Kievit, 2015) (Figure 2). There were two images per page with a total of 32 images for subjects to assess. Each image was categorized as easy (0° rotation), intermediate (50° rotation or °100 rotation), and hard (150° rotation). Subjects were scored on correctness of each response, and there was no time limit for each response.

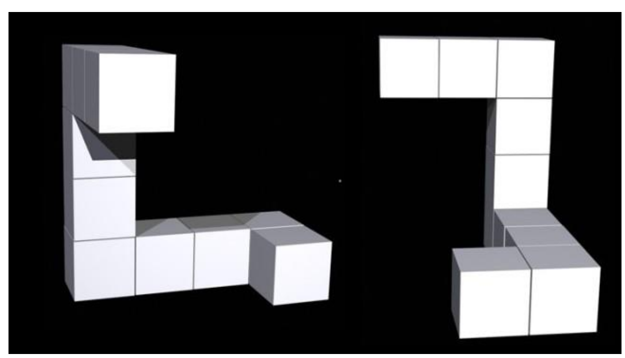

Figure 2. Mental rotation task from Ganis & Kievit (2015)

2.3.2. Mental Transformation Task

The mental transformation task is also designed to test the visuospatial capabilities of subjects (Ehrlich et al., 2006). Subjects were shown four images on one page and asked to identify which image is created when two separate pieces of a whole image are combined. Subjects viewed the four images horizontally on one sheet of paper and their attention was directed to the sheet of paper directly below that contained two separated pieces of one of the four images displayed above (Figure 3). Subjects had to indicate which of the four images above would be created when the two pieces are combined into one whole image. Figure 4 demonstrates the types of rotation or translation of the split images (direct translation & rotation and diagonal translation & rotation). There was a total of 32 sets of images for the subject to identify. Subjects were scored on the correctness of their response, and there was no time limit for each response.

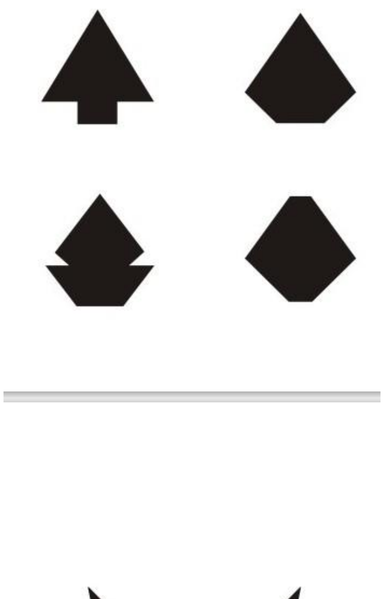

Figure 3. Mental transformation task choice array from Ehrlich et al., 2006

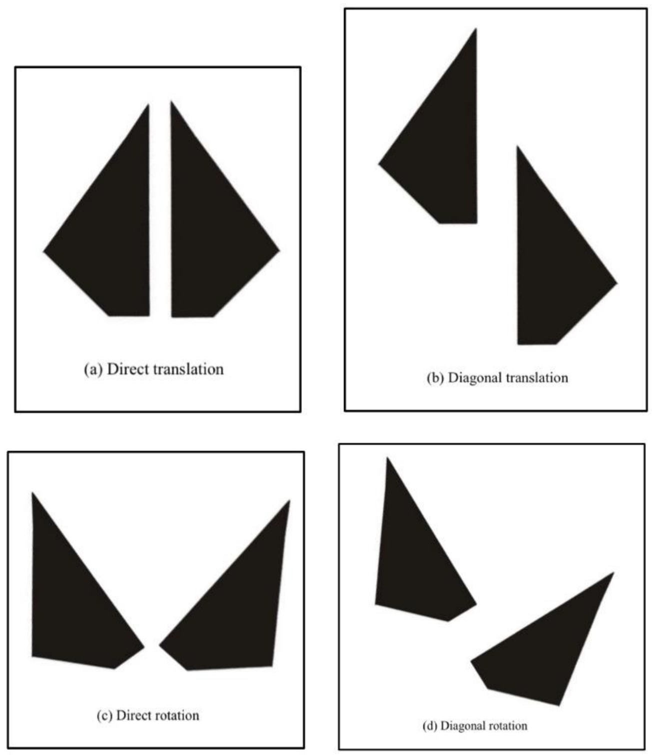

Figure 4. Mental transformation task rotation options (a-d) from Ehrlich et al., 2006

2.3.3. Stroop Test

The Stroop test is a common neuropsychological test that was used to measure the subject’s ability to overcome cognitive interference occurring when processing one type of stimuli in the presence of an additional conflicting stimulus (Scarpina & Tagini, 2017). Three individual Stroop tests were given to subjects: the congruency condition in which the font color and meaning of the word is the same, the incongruency condition in which the font color of the word does not match the meaning of the word (subjects response should match the font color), and a control condition in which subjects are presented with rows of colored rectangles without words and should respond with the color of each rectangle (Rouger, 2020). For each of the three conditions, subjects were scored on the number of correct responses they could produce in 45 seconds.

2.3.4. Sternberg Memory Task

The Sternberg Memory task was used to assess the ability of subjects to store and retrieve information from short-term memory (Padgett, 2018). This assessment was completed on a desktop computer within the laboratory space. Subjects were presented with a list of random letters in the alphabet on the screen, which then disappeared, and a single letter was shown. Subjects had to indicate whether that single letter was present in the previous letter set or not as quickly as possible. Subjects were only permitted to use their index finger on their right hand to indicate their response. A “yes” response was represented by pressing the period (.) key when the single target letter was present in the previous set. A “no” response was represented by pressing the comma (,) key when the single target letter was not present in the letter set. Subjects were instructed to keep their index finger on the keyboard in between the comma and period keys to ensure their response was made as quickly as possible (Padgett, 2018). Subjects completed 10 unscored practice trials before continuing to the official scored assessment, which consisted of 50 trials. Subjects were scored on the average reaction time of each response (ms) and the total number of errors.

2.4. Procedures

2.4.1. Experimental setup

All subjects were seated in an armchair in a designated testing room within the laboratory space. A standard script was read aloud to each subject detailing the instructions for the study as well as a brief demonstration and only deviated slightly when describing which devices were to be worn for the experiment. Subjects were assigned to one of two groups (control vs. experimental). Subjects in the control group completed the paradigm with their hand, while subjects in the experimental group completed the paradigm while wearing a body-powered transradial prosthesis on their right arm. Other than the addition of the prosthesis, the paradigm remained the same for both groups. Subjects in both groups wore an eye-tracking device that was adjusted for each subject so that it did not interfere with their visual field. Subjects in the experimental group were properly fitted for the prosthesis, and they were verbally instructed on how to use the device.

The instructions for the paradigm involved two rounds of subjects building two unique structures by imitating the movements of the researcher to construct the final design. Each structure was built using colored, cube-shaped blocks measuring 2.2 x 2.2 x 2" on a tabletop grid that was sectioned into four quadrants (9 x 9" for all four quadrants), as shown in Figure 5. The cube-shaped blocks were split into groups of four colors: red, dark blue, teal, and pink. There were a total of 48 blocks with 12 blocks for each of the four colors. Not every block was used in the construction of each structure and not all colors were used evenly. The two unique structures each consisted of seven blocks, and the maximum height for both structures was three blocks (~ six inches heigh). Other than this control, either structure could consist of color patterns of the blocks. Prior to beginning round one, subjects were shown the four block colors and their corresponding names.

Figure 5a. Blocks used to build structure

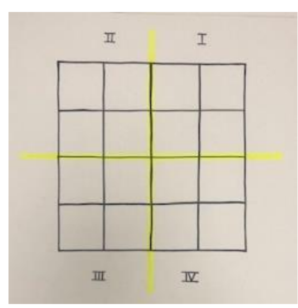

Figure 5b. Tabletop grid split into four, labeled quadrants

The subject and researcher sat on opposite sides of the tabletop grid, and the subject was instructed to build their structure in quadrant II (Q2), while the researcher built their structure in quadrant IV (Q4). On each side of the grid (to the right of subject and the researcher), the blocks were lined up in six rows. Half of the rows consisted of two adjacent stacks of blocks with one red-colored stack and one green-colored stack. Similarly, the other half consisted of one blue-colored stack and pink-colored stack (Figure 6). This setup pattern was consistent for the researcher and the subject

Figure 6. Subject's view of block setup to the right of the grid space

Subjects completed two rounds with each round consisting of 15 trials. Each of the 15 trials in rounds one and two consisted of the subject imitating the researcher's movements block-by-block until the structure was formed. The researcher would begin by picking up a pre-determined colored block and placing it on one of the four squares in Q4. The subject then had to pick up the same-colored block and place it in the corresponding position in Q2. This was repeated until the structure was built. After the structure of seven blocks was constructed, the grid was cleared by the researcher and the blocks were reset back to their marked positions. This procedure was repeated 15 times to represent each of the 15 trials. Once the researcher and subject had built, cleared, and re-built the same structure 15 times, the round was concluded (round 1 = 15 trials = structure 1 built 15 times). The second round followed the same procedure but with the formation of a different structure, though, still consisting of seven blocks (round 2 = 15 trials = structure 2 build 15 times). Figure 7 displays the two different structures produced in the paradigm.

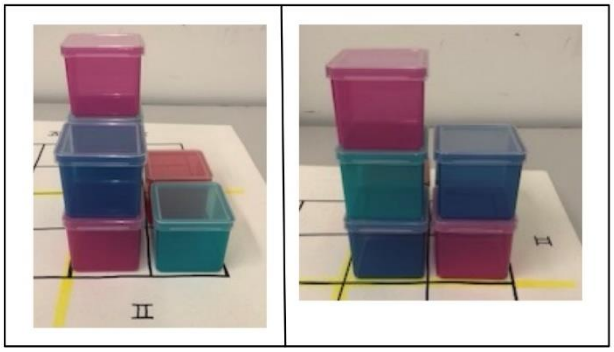

Figure 7a. Front (left) & side (right) view of structure 1 in subject’s quadrant

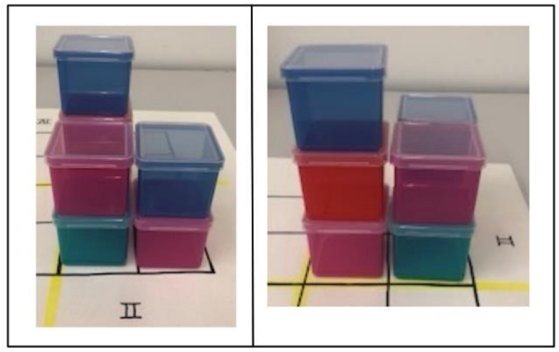

Figure 7b. Front (left) & side (right)view of structure 2 in subject’s quadrant

2.5. Data collection and analysis

2.5.1. Cognitive assessments

Scoring for each assessment was conducted as mentioned above. For each subject, six scoring values were obtained: (1) number of errors from the mental rotation task, (2) number of errors from the mental transformation task, (3) number of word responses and (4) errors from the Stroop test, and (5) reaction time (ms) and (6) number of errors from Sternberg memory task. Subjects’ scores from the assessments were statistically analyzed and compared between the control and experimental group to determine if there were any significant deviations in cognitive capabilities pertaining to the paradigm between the two groups. This was done via a two-sample t-test assuming unequal variances in Microsoft Excel. In addition, each subject’s scores were compared to their average task completion times and performance errors to evaluate corresponding trends in the data. While the data set from the assessments were not representative of direct evidence of motor learning, it was used as a correlative measure for predicted outcomes of subject’s ability to succeed in the paradigm.

2.5.2. Eye tracking

PupilCapture and PupilPlayer (Kassner et al., 2014) were programs used to record and extract data from the eye-tracking device. Within PupilPlayer, surfaces were manually defined according to the areas of interest on and around the grid space. There were two main surfaces defined as Q2/subject in progress (SIP), which included the subject’s quadrant and the proximal area that included the path of the right arm movements, and Q4/researcher in progress (RIP), which included the researcher's quadrant and the proximal area that included the path of the right arm movements.

Data specific to these surfaces was extracted and included: gaze positions on each surface (time of source image frame relative to x and y positions in the frame, gaze presence, & computed confidence), surface position (world view frame relative to its timestamp), surface gaze distribution (gaze count in each defined surface relative to total count), and generated heatmaps representing gaze pattern distribution for each surface relative to the entire grid space.

2.5.3. Task completion time

Task completion time was measured in seconds and represents the total amount of time the subject spent reaching to grasp a block, moving it to the grid space, and placing it down on the grid. The timer was started as soon as the subject initiated movement of their right arm and stopped when the base of the block was placed flat down on the grid. Any minor adjustments made by the subject after the block was placed flat on the grid (post-placement adjustments) were excluded. Post-placement adjustments were defined as any minor shift of the block once three-fourths of its diameter were within the bounds of a singular space in the subject's quadrant. Task completion time was recorded for every trial and resulted in 30 times for each subject. The parameters for task completion time remained constant between the control and experimental group. The average task completion times were calculated in Microsoft Excel for round one, round two, and both rounds as well as the increase or decrease in speed between rounds. The three averages for the control and experimental group were statistically analyzed using a two-sample t-test assuming unequal variances.

2.5.4. Performance and task errors

Performance and task errors were extracted from videography data collected during the study. Each subject's experimental paradigm was videorecorded, and consent for this was obtained and approved by Georgia Institute of Technology's Institutional Review Board. Performance and task errors were created and defined as a method of assessing each subject's performance in the paradigm. Performance errors represented the number of occurrences a block was placed in an incorrect position on the grid in a singular trial. This type of error included a block placed in a mismatched space within a subject's quadrant or the use of a wrong-colored block. The total number of performance errors per round for each subject in both groups was calculated in Microsoft Excel and statistically analyzed using a two-sample t-test assuring unequal variances.

Task errors represented the total number of occurrences for four errors: (1) knocking a block off the structure, (2) dropping a block while attempting to place it on the grid, (3) placing a block on the grid with a position error (block not in upright position or greater than three-fourths of block's diameter is outside bounds of a singular grid space), and (4) number of post-placement adjustments (defined above in task completion time section). The average number of the four task errors per round for each subject in both groups was calculated in Microsoft Excel and compared between the control and experimental group to determine if there was a decrease in the number of errors from round one to round two between subjects in the same group and a trend in the number of errors between the control and experimental group. Both performance and task errors were obtained for every trial and resulted in 30 sets of data for each subject. The parameters for performance and task errors also remained constant between the control and experimental group.

3. Results

3.1. Cognitive assessments

When comparing each group's scores for each of the six cognitive assessment categories, there was no significant difference observed. Figure 8 shows the averages of each category for both groups as well as the associated p-value (each > 0.05). The means in errors for the mental transformation task, Stroop test, and Sternberg reaction time varied the most between groups. The average error count of mental transformation errors for the control group (A) was 1.5, while the average for the experimental group (B) was 0.33. This was likely due to the high number of errors from subject A3, which was identified as an outlier in the data set. The Stroop test yielded an average word count of 59.17 for the control group and 45.83 for the experimental group with corresponding average errors of 1 for the control and 0.67 for the experimental group. This trend can be explained by the trade off between time and accuracy often observed with the Stroop test. On average, the control group reached a higher word count in the allotted 45 seconds, but this resulted in a lower accuracy (higher error count). The opposite trend was observed with the experimental group as they reached a lower word count but maintained a higher accuracy (lower error count). Additionally, the Sternberg memory task reaction time was lower for the experimental group compared to the control group. The experimental group had an average reaction time of 852 ms, while the control group had an average of 921 ms. Despite this difference in reaction times, both groups averaged four errors for this task. Even with these more notable gaps in average scores between the groups, the difference in values were not statistically significant. This is expected and indicates that all subjects possessed the basic cognitive capabilities exercised in the paradigm and that no subject group was more likely to perform significantly better based on cognitive function.

Group	Avg. Rotation Errors	Avg. Transformation Errors	Avg. Stroop Test Word Count	Avg. Stroop Test Errors	Avg. Sternberg Reaction time (ms)	Avg. Sternberg Errors
A	4.33	1.5	59.17	1	921.17	4
B	4.83	0.33	45.83	0.67	852.17	4
p-value	0.764	0.352	0.079	0.449	0.532	1

Figure 8. Table of average number of errors of both groups for each of the six cognitive assessment categories and the p-value for each (difference > 0.05).

3.2. Eye tracking

3.2.1. Q2/SIP surface and Q4/RIP surface heatmaps

The figures below (Figure 9 & 10) are heatmap images generated by PupilPlayer (Kassner et al., 2014). Both figures represent gaze positions for the entire grid space. The colored areas on the heatmap show the frequency of gaze position landing within the defined bounds of the Q2/SIP surface during rounds one and two. The warmer colors (red, orange, yellow) indicate a heavily concentrated area of increased gaze position, while the cooler colors (green, light blue, dark blue) indicate areas where gaze position was not as prominent. In figures 9 and 10 (round 1 & 2), the gaze position of the subject in the experimental group is localized to the lower right quadrant, which is representative of the SIP portion of the surface. In contrast, the gaze position of the subject in the control group is distributed more throughout the Q2 portion and extends more outward on the grid. This trend in gaze position from the heatmaps was observed for majority of subjects and suggests that subjects in the experimental group show increased gaze position on the path of their right arm as they direct their movements towards the quadrant rather than on the quadrant itself.

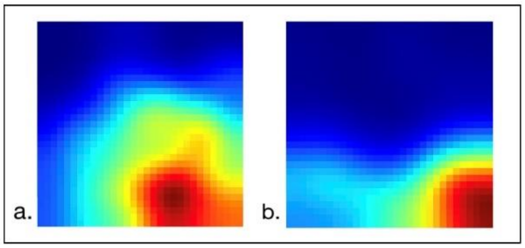

Figure 9. (a) heatmap of grid space representing concentrated gaze patterns in the Q2/SIP surface for subject A1, round 1. (b) heatmap of grid space representing concentrated gaze patterns in the Q2/SIP for subject B5, round 1.

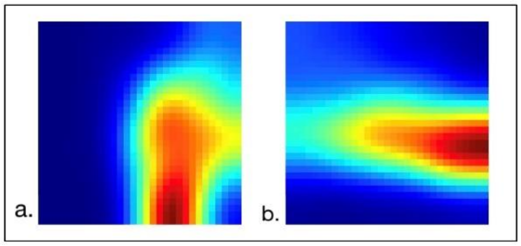

Figure 10. (a) heatmap of grid space representing concentrated gaze patterns in the Q2/SIP surface for subject A2, round 2. (b) heatmap of grid space representing concentrated gaze patterns in the Q2/SIP for subject B2, round 2.

Figures 11 and 12 are two additional heatmap images (Kassner et al., 2014) of gaze positions for the entire grid space. The colored areas on the heatmap show the frequency of gaze position landing within the defined bounds of the Q4/RIP surface during rounds one and two. In both figures (round 1 & 2), the gaze position of the subject in the experimental group is more heavily concentrated on the movement path of the researcher, RIP position of the surface. Compared to Figure 11a, 11b shows a dense region of gaze point adjacent to Q4, which corresponds to the location of the researcher's blocks used to build the structure. Compared to Figure 12a, 12b shows a wider, denser region in front of/adjacent to Q4, which corresponds to path of the researcher's movements. This trend in gaze position from the heatmaps was also largely observed for subjects in each group and suggests that subjects in the experimental group demonstrate increased gaze position on the path of the researcher's right arm, RIP, as they initiate movements to and from Q4 rather than fixating gaze solely within Q4.

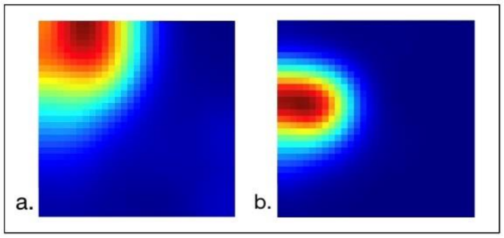

Figure 11. (a) heatmap representing Q4/RIP surface for subject A5, round 1. (b) heatmap representing Q4/RIP for subject B5, round 1.

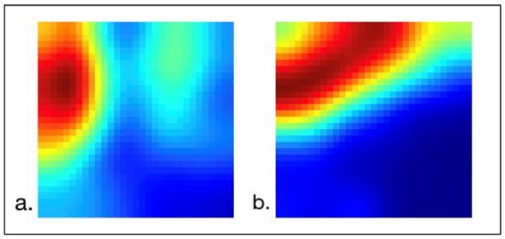

Figure 12. (a) heatmap representing Q4/RIP surface for subject A3, round 2. (b) heatmap representing Q4/RIP for subject B6, round 2.

3.3. Task completion time

To evaluate whether subjects in the control and experimental group were able to show evidence of motor learning from round one to round two, averages of each trial for both rounds were calculated as well as a percentage representing the change in task completion time. The average task completion time (s) for round one and round two as well as the percent decrease in average task completion time (how much quicker the subject completed round two compared to round one) for each subject in the control (A) and experimental group (B) is shown in the two tables below.

Subjects in the control group displayed a 10.94% mean decrease in task completion time from round one to round two (Figure 13). Five of the six subjects in the control group, on average, demonstrated the ability to complete the task for round two more quickly compared to round one. Subject A5 is considered an outlier as they are the only subject to show a minimal increase in their task completion time. Due to this, they were not included in the statistical analyses to ensure the data was not skewed. A t-test was conducted between the average task completion times between round one and two. The t Stat value (1.323) was not greater than the t Critical two-tail (2.365), and therefore the null hypothesis that there is no significant difference between average task completion times between rounds for the control group cannot be rejected. Consequently, there was no statistically significant difference in task completion times between the rounds ($p = 0.227$, $p > 0.05$).

Subject ID	**Round 1 Average (s)**	**Round 2 Average (s)**	**Decrease in task completion time (%)**
A1	13.489	13.482	0.05
A2	11.39	9.991	12.29
A3	14.411	12.962	10.06
A4	19.093	13.702	28.24
A5	11.229	11.327	-0.88
A6	15.601	13.117	15.92

Figure 13. Task completion time averages for the control group with a mean decrease in task completion time of 10.94% from round 1 to round 2

Subjects in the experimental group displayed a 25.02% mean decrease in task completion time from round one to round two (Figure 14). Each subject in the experimental group, on average, demonstrated the ability to complete the task for round two more quickly compared to round one. This observed trend is consistent with the decrease in task completion time for the control group, though the experimental group showed a greater difference in mean task completion time between rounds compared to the control group (25.02% vs. 10.94%). A t-test was conducted between the average task completion times between round one and round two. The t Stat value (2.924) was greater than the t Critical two-tail (2.365), and therefore the null hypothesis that there is no significant difference between average task completion times between rounds for the experimental group can be rejected. Consequently, there was a statistically significant difference in task completion times between the rounds ($p = 0.022$, $p < 0.05$)

Subject ID	Round 1 Average (s)	Round 2 Average (s)	Decrease in task completion time (%)
B1	49.568	36.686	25.99
B2	40.279	33.775	16.15
B3	40.594	33.189	18.24
B4	49.51	40.118	18.97
B5	49.589	32.013	35.44
B6	67.385	43.597	35.3

Figure 14. Task completion time averages for the experimental group with a mean decrease in task completion time of 25.02% from round 1 to round 2

Two additional t-tests were performed to see if there was a significant difference in task completion times between the control and experimental groups for each round. Data comparing round one values between the two groups revealed a t Stat value (8.416) much greater than that of the t Critical two-tail (2.447), and therefore the null hypothesis that there is no significant difference between round one average task completion times between groups can be rejected. Consequently, there was a statistically significant difference in round one task completion times between the control and experimental group ($p = 0.00015$, $p < 0.001$). Data comparing round two values between the two groups also revealed a t Stat value (12.455) much greater than that of the t Critical two-tail (2.445), and therefore the null hypothesis that there is no significant difference between round two average task completion times between groups can also be rejected. Consequently, there was a statistically significant difference in round two task completion times between the control and experimental group ($p = 1.64E-05$, $p < 0.001$).

3.4. Performance and task errors

3.4.1. Performance errors

For both the control and experimental group, the number of performance errors for each round as well as the percent increase/decrease in errors between round one and round two was obtained for each subject. The tables displayed in this section show results for the control (A) and experimental (B) group.

The number of performance errors for the control group (Figure 15) are lower, as expected, and half of the subjects had zero errors for both rounds. Therefore, error status is represented at “No errors”. Subject A4 showed the most significant decrease in performance errors from nine errors in round one to zero errors in round two (100% decrease). Subject A6 also showed a decrease in performance errors from one error in round one to zero errors in round two (100% decrease). The outlier for this data set is subject A1 who showed an increase in performance errors from zero in round one to two in round two. A t-test was conducted between the number of performance errors between round one and two. The t Stat value (0.881) was not greater than the t Critical two-tail (2.445), and consequently, there was no statistically significant difference in performance errors between the rounds ($p = 0.412$, $p > 0.05$). This result was expected for the control group as it was hypothesized that there would not be a significant difference in error count between rounds.

Subject ID	No. of errors in round 1	No. of errors in round 2	Error status (increase or decrease)
A1	0	2	Increase
A2	0	0	No errors
A3	0	0	No errors
A4	9	0	Decrease
A5	0	0	No errors
A6	1	0	Decrease

Figure 15. Performance errors represented by round for the control group.

The number of performance errors for the experimental group (Figure 16) are more numerous than those of the control group with only one subject having zero errors for both rounds. Subject B5 showed the most significant decrease in performance errors from two errors in round one to zero errors in round two (100% decrease). Subject B2 also showed a 50% decrease in errors with two errors in round one and one error in round two, and subject B3 showed a 33.3% decrease in errors with three errors in round one and two errors in round two. Subjects B1 and B4 showed an increase in errors with both having zero errors in round one to one and two errors in round two. A t-test was conducted between the number of performance errors between round one and two. The t Stat value (0.255) was not greater than the t Critical two-tail (2.262), and consequently, there was no statistically significant difference in performance errors between the rounds ($p = 0.805$, $p > 0.05$). This result does not support the initial hypothesis that there would be a significant decrease in total error count between rounds for the experimental group.

Subject ID	No. of errors in round 1	No. of errors in round 2	Error status (increase or decrease)
B1	0	1	Increase
B2	2	1	Decrease
B3	3	2	Decrease
B4	0	2	Increase
B5	2	0	Decrease
B6	0	0	No errors

Figure 16. Performance errors represented by round for the experimental group.

Two additional t-tests were performed to see if there was a significant difference in performance errors between the control and experimental groups for each round. Data comparing the performance error count between the two groups for round one revealed a t Stat value (0.318) much lower than that of the t Critical two-tail (2.445), which indicates that the difference in round one performance errors between the groups is not statistically significant ($p = 0.761$, $p > 0.05$). Data comparing the same error count between the two groups for round two revealed a t Stat value (1.348) lower than that of the t Critical two-tail (2.228), which indicates that the difference in round two performance errors between the groups is not statistically significant ($p = 0.207$, $p > 0.05$).

3.4.2. Task errors

Task errors were categorized into four subdivisions (defined above in *section 2.5.4*) and evaluated for each subject in both groups. Each subject had a table displaying the number of errors for each of the four subdivisions, and a single value was calculated that represented the total number of task errors for the corresponding round. Below is a sample table for subject B1's round one task errors (Figure 17). The green space with the value ‘102’ represents the total number of task errors calculated for round one.

Task errors (R1)	102															
Trial	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Total
(1) knocks block off structure	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
(2) drops block en route to grid	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(3) positional error on grid	1	4	0	0	3	3	1	7	0	1	1	1	2	0	0	24
(4) # of post-placement adjustments	11	8	9	3	5	7	5	6	4	1	4	4	4	3	3	77

Figure 17. Table representing layout of subject B1’s round 1 task errors.

Figure 18 below shows the total number of task errors for round one and round two as well as the percent decrease in errors for subjects in the control group. Five out of six subjects in the control group showed a decrease in task errors from round one to round two. Subject A5 is the outlier of this data set, similarly to this subject’s trend in task completion time, and showed an increase in number of errors in round two. Therefore, subject A5’s data was excluded from the statistical analyses to ensure data was not skewed. Subjects in this group (excluding A5) displayed a 24.19% mean decrease in errors from round one to round two. The overarching trend is subjects’ total error count decreases in the latter task. A t-test was conducted between the number of task errors between round one and two. The t Stat value (0.454) was not greater than the t Critical two-tail (2.306), and consequently, there was no statistically significant difference in task errors between the two rounds ($p = 0.662$, $p > 0.05$). This result was expected for the control group as it was hypothesized that there would not be a significant difference in total error count between rounds.

Subject ID	No. of errors in round 1	No. of errors in round 2	Decrease in errors (%)
A1	9	8	11.11
A2	1	1	0
A3	18	15	16.67
A4	11	9	18.18
A5	1	3	-200
A6	4	1	75

Figure 18. Task errors represented by round for the control group as well as decrease in errors between round 1 and 2 with a 21.19% mean decrease in errors.

Figure 19 below shows the total number of task errors for round one and round two as well as the percent decrease in errors for subjects in the experimental group. Results show all six subjects showed a decrease in task errors from round one to round two. Subjects in this experimental group displayed a 37.57% mean decrease in errors from round one to round two, thus, demonstrating a trend of a lower error count in the second task compared to the first. This trend shows that the experimental group showed a greater decrease in errors from round one to round two compared to the decrease in errors between rounds for the control group; however, the experimental group still had a higher task error count for each round. A t-test was conducted between the number of task errors between round one and two. The t Stat value (2.046) was not greater than the t Critical two-tail (2.262), and therefore the null hypothesis that there is no significant difference between task errors between rounds for the experimental group cannot be rejected. Consequently, there was no statistically significant difference in task error count between round one and two ($p = 0.071$, $p > 0.05$).

Subject ID	No. of errors in round 1	No. of errors in round 2	Decrease in errors (%)
B1	102	84	17.65
B2	111	47	57.66
B3	69	45	34.78
B4	172	119	30.81
B5	202	101	50
B6	194	127	34.54

Figure 19. Task errors represented by round for the experimental group as well as decrease in errors between round 1 and 2 with a 37.57% mean decrease in errors.

Two additional t-tests were performed to see if there was a significant difference in number of task errors between the control and experimental groups for each round. Data comparing round one errors between the groups revealed a t Stat value (5.882) much greater than that of the t Critical two-tail (2.571), and therefore the null hypothesis that there is no significant difference between round one task error count between the groups can be rejected. Consequently, there was a statistically significant difference in round two task error count between the control and experimental group ($p = 0.002$, $p < 0.005$). Data comparing round two errors between the groups revealed a t Stat value (5.501) also much greater than that of the t Critical two-tail (2.571), and therefore the null hypothesis that there is no significant difference between round two task error count between the groups can also be rejected. Consequently, there was a statistically significant difference in round two task error count between the control and experimental group ($p = 0.003$, $p < 0.005$).

4. Discussion

4.1. Conclusions

4.1.1. Task completion time

This study aimed to test how motor learning in a 3D, goal-oriented visuospatial task can be impacted by action observation and social intent through measuring task completion time, performance and task errors, and gaze patterns. Results suggest that subjects in the experimental group show significant decrease in task completion time from round one to round two, while subjects in the control group do not display a significant decrease in task completion time. This data supports the hypothesis that subjects in the experimental group show a decrease in average task completion time, but it refutes the hypothesis of subjects in the control group also showing this same decrease. These findings suggest that subjects completing the paradigm in an unfamiliar upper limb prosthesis were able to show evidence of motor learning, based on task completion time, given their ability to build the structure in round two significantly quicker than in round one. Motor learning is suggested to have occurred as these subjects acquired the skill by imitating the researcher to continuously build and rebuild the same structure in round one, and then showed a transfer of skills by displaying the same abilities in a quicker manner in round two, in which the design of the structure changed. Changing the design of the structure in round two asked motor learning of the subject since they had to enact a new set of movements to build the structure in a different spatial arrangement. Despite the change in structure and its orientation, subjects were able to improve upon their time for this new structure. This means there is some aspect, potentially visuospatial, that impacted their ability to display a degree in learning by continuously completing the task more quickly in round two. The control group did not have a significant decrease in task completion time; however, their task completion times for round one and two were both significantly lower than those of the experimental group. This was expected since they completed the task without a prosthesis or any other cognitive or physical limitations. It was hypothesized that they would overall have lower task completion times when compared to subject's wearing the prosthesis. Additionally, there was a significant difference between task completion time for both rounds between the control and experimental group with the most significant difference observed in round two.

4.1.2. Performance errors

Errors displayed by the subjects were broken into two categories: performance and task errors. The research question and corresponding hypothesis grouped both type of errors together as total error count. Interestingly, results varied based on the type of specific error. For both the control and experimental group, the difference in performance errors between round one and two was not significant. Half of the subjects in the control group did not have any performance errors, while two of the subjects actually showed an increase in performance errors in round two, though the range was low at 0-1 errors in round one. Only one subject showed a 100% decrease in performance errors from round one to round two. These results are greatly varied amongst the control group, which was not expected. However, the difference in errors between rounds was not significant, which supports the hypothesis. An odd occurrence of performance errors in the experimental group was also observed, though different from the control group. Half of the subjects showed a decrease in the amount of performance errors from round one to round two, though these subjects displayed a low range of 2-3 errors in round one. Two subjects had an increase in number of errors between rounds, and one subject had no errors in either round. These results were more supportive of the hypothesis as half of the subjects did show a decrease in errors; however, the variation in trend amongst this group was not expected.

Performance errors measured the subject's ability to overcome the spatial transformation required to build the structure as they had to imitate movements of the researcher in a non-mirrored position. Based on results, subjects in the control generally did not have any difficulties with the spatial transformation required of them to successfully build the structure in the correct positions in their quadrant, since majority of subjects had either zero errors or showed a decrease in errors. Half of the subjects in the experimental group did have more errors than majority of subjects in the control group; however, these errors decreased by round two. Additionally, the errors cannot definitively be attributed to difficulties with spatial transformation due to the success of these subject's cognitive assessment scores. The experimental group subjects that showed a higher number of errors in round one yielded a low number of errors in the mental rotation and transformation tasks, which suggests that there is a separate factor contributing to these performance errors.

4.1.3. Task errors

Task errors were the second category evaluated and comprised of more kinematic-related errors. These included when a subject had an obvious lapse in motor control that involved knocking a block off the structure, dropping a block while en route to the grid space, placing a block on the grid with a positional error (top of block not facing upwards or less than three-fourths of the diameter not within a single space of the quadrant), frequency of post-placement adjustments (minor shift of the block once it was placed within three-fourths of a single space of the quadrant). Each of these four subcategories of errors were summed for each subject for analyses.

When evaluating the total number of task errors, the control group did not show a significant decrease in number of task errors from round one to two. This was hypothesized as the control group completed the task with no physical limitations and did not wear the prosthesis, thus, it was expected that their motor control would not be disrupted. The experimental group also did not show a significant decrease in task errors from round one to round two. However, this does not mean they did not show evidence of adaptation to the device and increasing motor control. Every subject in this group did show a decrease in the number of task errors from round one to round two, but it was not statistically significant. Therefore, the hypothesis is refuted given the decrease was not statistically significant, but there was evidence of improvement from a kinematic standpoint. These results for the experimental group suggest that they could have a higher capacity for improvement due to the decrease in task errors, but it does not directly mean motor learning occurred. To determine this, further studies should be conducted with an additional round of testing to see if a similar decrease in errors occurs, or an additional day of testing should be added one week after the first testing session to observe if errors still decrease with a delay in time. When comparing number of task errors between the control and experimental group, there was found to be a significant difference in error count for both round one and round two with the most significant difference observed in round two.

4.1.4 Gaze positions

Gaze positions of the subjects in both groups were measured using an eye-tracking device and quantified using PupilPlayer (Kassner et al., 2014) in which two types of surfaces were defined. For both groups, Q2/SIP (subject's quadrant & path of movement) and Q4/RIP (researcher's quadrant & path of movement) surfaces were defined on the grid space to evaluate where gaze positions fixated most when completing the paradigm. When investigating positions between the groups for the Q2/SIP surface, the control group showed increasing gaze positions over Q2 itself rather than the path to and from the quadrant (SIP), suggesting their gaze patterns fixated more on building the structure in their quadrant rather than fixating gaze on their path to and from the quadrant. The opposite trend was observed for the experimental group. Subjects wearing the prosthesis showed increased gaze position over the SIP rather than localizing majority of gaze positions to Q2. These results support the hypothesis that subjects wearing the prosthesis will have fixate their gaze positions on the path movements to and from their quadrant compared to subjects in the control. Additionally, this is supported by the trend in task errors for the experimental group. The increased number of task errors, specifically the greater number of errors in dropping the block en route to the grid, can explain the increased gaze positions over the SIP area. When subjects dropped these blocks en route, they were instructed to pick them back up and correct this error until the block is placed on the grid. Therefore, the eye-tracking device detected increased gaze position around the grid space when these errors occurred. Subjects in the control group had significantly less errors en route to the grid, so gaze position was not heavily concentrated in that area.

The trend in Q2/SIP gaze positions was also observed in the Q4/RIP surfaces for both groups. Data showed subjects in the control group had increasing gaze positions over the localized Q4 region of the grid compared to the researcher's path to and from their quadrant (RIP). Once again, the opposite was true for subjects in the experimental group. Subjects wearing the prosthesis demonstrated increased gaze positions over the RIP regions of the surface rather than staying within the Q4 area. This supports the hypothesis of subjects in the experimental group having increased gaze positions on the path of the researcher compared to the control group. This supports the idea that visual cues are important in imitating a new motor skill (Muratori, 2015). The addition of the prosthesis was new for subjects in the experimental group, and they did not get the opportunity to practice movements with the device. This suggests they relied more observational cues from the researcher to increase their own ability in performing the task. It also highlights the role of social intent in the setting of this 3D visuospatial task. Subjects with the prosthesis continuously fixated their gaze positions on the researcher's path throughout both rounds, suggesting they were perceiving the researcher's actions to better adjust their own movements. This has been previously researched when a subject and researcher work together to achieve a shared outcome (Scorolli et al., 2014). Results from the present study indicate perceiving social intent can also lead to increased performance (less errors & decreased task completion time) when a subject and researcher engage in separate, rather than joint, actions. Further research can shed light into how visuospatial cues and social intent impact subject's outcomes when completing motor tasks.

4.2. Limitations of the study

A notable limitation for this study included the presence of data that had a suboptimal confidence interval of eye-tracking accuracy. This was likely due to a combination of device limitations, such as weaker resolution settings of the sensor, errors in fine-tuning pupillary movements during the calibration process, and physiological factors of individual subjects, including eyelid drooping potentially causing camera obstruction. For several subjects, the calibration process was repeated to increase pupil detection for the field of view. However, due to a significant time constraint for collecting data, the above errors were troubleshooted as best as possible and the experiment proceeded. In these cases, data was extracted and corrected to an extent via Pupil Player, though it did lead to a deviation in the accuracy of these subject's gaze patterns. In addition, three subjects had to be rescheduled and two subjects had their experiment's cancelled due to a device failure in which one of the two pupil detectors would not connect. This delay in collecting data proved to be another sizeable limitation as it resulted in lost time, which overall lowered the projected number of participants for this study.

There was a limitation in extracting accurate gaze pattern data for a few subjects due to a less than optimal placement of markers during data collection. PupilPlayer interprets gaze pattern in part by locking pupillary movements to physical markers that were present on and around the workspace. After reviewing data from the first few subjects, it was discovered that one of the surfaces of interest, Q4/RIP, was not properly aligned with the markers. This resulted in a lack of reliable gaze pattern data for the Q4/RIP surface for several subjects. However, despite this error affecting the raw gaze point values, these subjects' pupillary movements and generated heatmaps showed a similar trend as the other subjects in their group. In addition, their corresponding raw values of the Q2/SIP surface remained consistent with the other subjects in their group that did not have errors in marker placement. This suggests that while the raw data was skewed for one subcategory, the overarching conclusion for both the control and experimental group was supported by other modes of data collection.

4.3. Future directions and significance

Future work investigating patterns of gaze strategies in completing a goal-oriented task while simultaneously having to adjust movements based on those of another person could provide valuable data regarding the intricacies of error correction involved in motor learning. This could be further tested in subjects with two sound limbs or a more selective population sample. A future extension to this study could be repeating a similar procedure with one group of subjects (two sound limbs) wearing the prosthesis and the other group consisting of upper limb amputees using their personal prosthetic devices. This would be one way to examine how upper limb amputees must adapt their own movements based off those of another person and how this affects their motor control, as well as their speed and success in performing the task. Another interesting avenue includes using neuroimaging to investigate the same variables between the control and experimental group mentioned in this study. Adding neuroimaging, such as EEG, would provide time-locked electrical responses (Bell & Cuevas, 2012) that could assess the cognitive processes associated with the subject observing the researcher's movements vs. the subject completing the movements themselves. By making EEG the main source for data collection, it would provide a more direct measure of brain activity vs. the emphasized kinematic methods used in this study.

The present study demonstrated that motor learning could be disrupted by executing the paradigm while wearing an upper limb prosthesis. The goal of adding the prosthesis was to determine if using an unfamiliar device that physically challenges the subject would significantly impact motor learning, and to what degree social intention facilitated this learning, which is discussed above. The observed change in motor control shown by adding the prosthesis can be widely applied to the difficulties upper limb amputees face in relearning goal-oriented movements with a prosthesis post-amputation. Currently, there is a lack of research investigating changes in action observation and how social intention/cooperation impacts motor learning with their residual limb specifically for upper limb amputees (Bayani et al., 2019; Scorolli et al., 2014; Wheaton, 2017; Williams et al., 2016). Therefore, future directions should focus on gathering data related to these variables and evaluating how this patient population adapts to their prosthetic device with the goal of improving motor control. Continued research in this area could provide insight into how a physical trauma, such as amputation, changes the interconnected motor networks in the brain and its resulting affects on motor learning and behavior. Subsequently, more clarity in how these factors change in upper limb amputees can lead to improvement of current rehabilitation techniques with the goal of increased functional device utility (Wheaton, 2017) and motor control.

References

Bayani, K.Y., Lawson, R.R., Levinson, L., Mitchell, S., Atawala, N., Otwell, M.,…Wheaton, L.A. (2019). Implicit development of gaze strategies support motor improvements during action encoding training of prosthesis use. *Neuropsychologia, 127,* 75-83. doi: 10.1016/j.neuropsychologia.2019.02.015.

Bell, M.A. & Cuevas, K. (2012). Using EEG to Study Cognitive Development: Issues and Practices. *Journal of cognition and development: official journal of the Cognitive Development Society, 13 (3),* 281-294. https://doi.org/10.1080/15248372.2012.691143

Cusack, W.F, Thach, S., Patterson, R., Acker, D., Kistenberg, R.S., Wheaton, L.A. (2016). Enhanced neurobehavioral outcomes of action observation prosthesis training. *Neurorehabilitation and Neural Repair, 30 (6),* 573-582. https://doi.org/10.1177/1545968315606992

Ehrlich, S., Levine, S.C., Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental Psychology, 42,* 1259-1268. http://dx.doi.org/10.5334/jopd.ai

Excel Easy. (2010). *T-test*. Retrieved December 8, 2021, from https://www.excel-easy.com/examples/t-test.html

Ganis, G. & Kievit, R. (2015). A New Set of Three-Dimensional Shapes for Investigating Mental Rotation Processes: Validation Data and Stimulus Set. *Journal of Open Psychology Data, 3, e3.* http://dx.doi.org/10.5334/jopd.ai

Kassner, M., Patera, W., & Bulling, A. (2014). Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. UbiComp, 1151-1160. https://doi.org/10.1145/2638728.2641695.

Muratori, L.M., Lamberg, E.M., Quinn, L., & Duff, S.V. (2013). Applying principles of motor learning and control to upper extremity rehabilitation. *Journal of hand therapy: official journal of the American Society of Hand Therapists*, *26 (2)*, 94-103. https://doi.org/10.1016/j.jht.2012.12.007

Padgett, R.J. (2018). *Sternberg Memory Task.* The Sternberg Lab. https://rpadgett.butler.edu/nw221/sternberg_lab/index.html

Rouger, E. (2020). Printable Stroop effect game. *Memozor* . https://www.memozor.com/memory-games-to-print/words/stroop-effect-game

Scarpina, F. & Tagini, S. (2017). The Stroop Color and Word Test. *frontiers in Psychology, (8),* 557. doi: 10.3389/fpsyg.2017.00557.

Scorolli, C., Massimiliano, M., Wheaton, L.A., & Borghi, A.M. (2014). I give you a cup, I get a cup: A kinematic study on social intention. *Neuropsychologia*, *57*, 196-204. http://dx.doi.org/10.1016/j.neuropsychologia.2014.03.006

Wheaton, L.A. (2017). Neurorehabilitation in upper limb amputation: understanding how neurophysiological changes can affect functional rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, *14 (41).* doi: 10.1186/s12984-017-0256-8.

Williams, L., Pirouz, N., Mizelle, J.C., Cusack, W., Kistenberg, R., & Wheaton, L.A. (2016). Remodeling of cortical activity for motor control following upper limb loss. *Clinical Neurophysiology*, *127*, 3128-3134. http://dx.doi.org/10.1016/j.clinph.2016.07.004

Zacks, J.M. & Michelon, P. (2005). Transformations of Visuospatial Images. *Behavioral and Cognitive Neuroscience Reviews, 4(2),* 96-118. https://doi.org/10.1177/1534582305281085.