**Table 1.** Paired t-test for training performance of children at advanced rehabilitation stage

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Evaluations** | **Mean ± Standard deviation** | | **Difference Value** | **t** | **p** |
| **Before**  **experiment** | **After**  **experiment** |
| **Auditory memory** | | | | | |
| Auditory memory1 | 4.56±0.51 | 4.75±0.45 | -0.19 | -1.861 | 0.083 |
| **Articulation and Expression** | | | | | |
| Articulation performance1 | 20.88±3.98 | 22.38±2.68 | -1.5 | -3.674 | 0.002\*\* |
| Auditory description with hints2 | 3.69±0.48 | 3.88±0.34 | -0.19 | -1.861 | 0.083 |
| Auditory description without hints2 | 2.75±0.77 | 2.88±0.72 | -0.13 | -1.464 | 0.164 |
| Speech intelligibility2 | 4.44±0.51 | 4.63±0.50 | -0.19 | -1.861 | 0.083 |
| **Training Engagement** | | | | | |
| Language for expression2 | 7.31±1.30 | 7.63±0.96 | -0.31 | -2.611 | 0.020\*\* |
| Language as game tool2 | 7.63±1.41 | 7.69±1.30 | -0.06 | -1 | 0.333 |
| Language for coordination2 | 7.56±1.31 | 7.63±1.31 | -0.06 | -1 | 0.333 |
| \*p < 0.05, \*\*p < 0.01 | | | | | |

1 represents the analysis results are obtained based on game performance data

2 represents the analysis results are obtained based on the auditory-verbal assessment

**Table 2.** Paired t-test for training performance of children at advanced rehabilitation stage in lip syncing game

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Evaluation metrics** | **Mean ± Standard deviation** | | **Difference Value** | **t** | **p** |
| **The 1st training** | **The 2nd training** |
| Completion time of Task A | 7.31±4.25 | 4.50±0.97 | 2.81 | 2.505 | 0.024\* |
| Errors in Task A | 1.13±1.50 | 0.19±0.40 | 0.94 | 2.7 | 0.016\* |
| Completion time of Task B | 10.31±14.08 | 4.50±1.37 | 5.81 | 1.767 | 0.098 |
| Errors in Task B | 1.88±2.36 | 0.31±0.60 | 1.56 | 2.854 | 0.012\* |
| Completion time of Task C | 6.69±6.49 | 3.88±0.89 | 2.81 | 1.733 | 0.104 |
| Errors in Task C | 1.19±1.52 | 0.31±0.60 | 0.88 | 2.333 | 0.034\* |
| \*p < 0.05, \*\*p < 0.01 | | | | | |

**Table 3.** Paired t-test for training performance of children at advanced rehabilitation stage in dubbing game

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Item** | **Mean ± Standard deviation** | | **Difference value** | **t** | **p** |
| **The 1st training** | **The 2nd training** |
| Completion time of Task A | 57.94±52.29 | 45.00±37.84 | 12.94 | 2.97 | 0.010\*\* |
| Errors of Task A | 1.56±1.86 | 1.88±2.13 | -0.31 | -1.431 | 0.173 |
| Completion time of Task B | 29.19±52.43 | 15.88±15.50 | 13.31 | 1.361 | 0.194 |
| Errors of Task B | 0.75±1.57 | 0.75±1.24 | 0 | 0 | 1 |
| Completion time of Task C | 80.19±53.02 | 67.69±43.09 | 12.5 | 3.619 | 0.003\*\* |
| Errors of Task C | 3.06±2.52 | 3.06±2.17 | 0 | 0 | 1 |
| \*p<0.05, \*\*p< 0.01 | | | | | |

**Table 4.** Paired t-test for training performance of children at advanced rehabilitation stage in direction game

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Direction Cognition** | **Mean ± Standard deviation** | | **Difference value** | **t** | **p** |
| **The 1st training** | **The 2nd training** |
| Completion time | 141.50±63.02 | 106.50±26.26 | 12.94 | 35 | 0.005\*\* |
| \*p<0.05, \*\*p< 0.01 | | | | | |

**Table 5.** Paired t-test for training performance of children at intermediate rehabilitation stage

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Evaluations** | **Mean ± Standard deviation** | | **Difference Value** | **t** | **p** |
| **Before**  **experiment** | **After**  **experiment** |
| **Auditory memory** | | | | | |
| Auditory memory1 | 3.67±0.49 | 4.00±0.53 | -0.33 | -2.646 | 0.019\* |
| **Articulation and Expression** | | | | | |
| Articulation performance1 | 16.20±1.57 | 18.93±4.28 | -2.73 | -4.771 | 0.000\*\* |
| Auditory description with hints2 | 2.73±0.59 | 3.07±0.70 | -0.33 | -2.646 | 0.019\* |
| Auditory description without hints2 | 1.20±0.86 | 1.33±0.72 | -0.13 | -1.468 | 0.164 |
| Speech intelligibility2 | 3.47±0.74 | 3.53±0.64 | -0.07 | -1 | 0.334 |
| **Training Engagement** | | | | | |
| Language for expression2 | 5.00±1.81 | 5.60±1.68 | -0.6 | -3.154 | 0.007\*\* |
| Language as game tool2 | 5.60±2.29 | 5.73±2.15 | -0.13 | -1.468 | 0.164 |
| Language for coordination2 | 5.40±1.76 | 5.53±1.85 | -0.13 | -1.468 | 0.164 |
| \*p < 0.05, \*\*p < 0.01 | | | | | |

1 represents the analysis results are obtained based on game performance data

2 represents the analysis results are obtained based on the auditory-verbal assessment

As shown in Table 5, “Auditory memory” was significantly improved. Different from children at advanced rehabilitation stage, the auditory-verbal capability of such participants is still in a preliminary stage. Therefore, the training appears very helpful to enhance their basic auditory abilities. Moreover, the dimensions of “Articulation performance”, “Auditory description with hints” and “Language for expression” got significantly improved, which further demonstrates that the gamified training can effectively enhance their articulation and expression performance.

However, "Auditory description without Hints" shows no significant difference. The dimension appears more challenging for children at the intermediate rehabilitation stage to get quick improvement. Besides, no significant differences were observed with "Speech intelligibility," "Language as a game tool," and "Language for coordination”. Similar with children at advanced rehabilitation stage, such dimensions require better mastering of the functional use of language, and are not easy to be influenced by limited training. Therefore, the influence on participants’ effective training engagement is limited.

For Lip syncing game, the performance data is shown in Table 6. It can be found that there is a significant decrease in completion time for all the tasks, which suggests the training performance can be effectively improved in mouth position learning and the pronunciation of single consonants. However, the influence on task error is not very clear, which indicates that the training is not useful enough to improve the training effect.

**Table 6.** Paired t-test for training performance of children at intermediate rehabilitation stage in lip syncing game

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Item** | **Mean ± Standard deviation** | | **Difference**  **value** | **t** | **p** |
| **Before the**  **Experiment** | **After the**  **Experiment** |
| Completion time of Task A | 16.80±14.01 | 9.20±5.23 | 7.6 | 3.111 | 0.008\*\* |
| Errors of Task A | 1.47±1.81 | 0.80±1.01 | 0.67 | 2.197 | 0.045\* |
| Completion time of Task B | 4.80±2.14 | 3.73±0.59 | 1.07 | 2.256 | 0.041\* |
| Errors of Task B | 0.07±0.26 | 0.13±0.52 | -0.07 | -0.435 | 0.67 |
| Completion time of Task C | 8.47±7.39 | 4.67±1.05 | 3.8 | 2.153 | 0.049\* |
| Errors of Task C | 0.40±0.63 | 0.27±0.59 | 0.13 | 1.468 | 0.164 |
| \*p<0.05 \*\*p< 0.01 | | | | | |

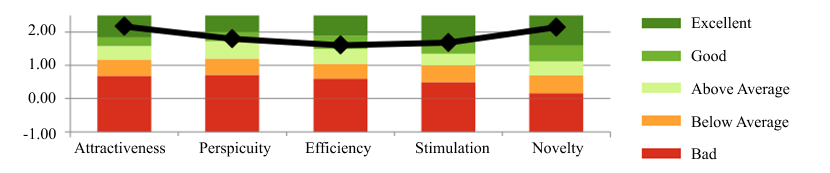
For Dubbing game, the performance data is shown in Table 7. Significant influence was observed only on the completion time of two tasks. As mentioned above, dubbing games have higher requirements on users’ scenario understanding and descriptive expression abilities. It is difficult for the children at intermediate rehabilitation stage to achieve marked improvement in their training performance.

**Table 7.** Paired t-test for training performance of children at intermediate rehabilitation stage in dubbing game

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Item** | **Mean ± Standard deviation** | | **Difference value** | **t** | **p** |
| **The 1st test** | **The 2nd test** |
| Completion time of Task A | 78.00±37.18 | 72.07±33.54 | 5.93 | 1.771 | 0.098 |
| Errors of Task A | 4.13±2.47 | 3.73±1.83 | 0.4 | 1.702 | 0.111 |
| Completion time of Task B | 117.40±86.44 | 98.87±87.39 | 18.53 | 2.875 | 0.012\* |
| Errors of Task BB | 4.80±2.88 | 4.53±2.77 | 0.27 | 1.293 | 0.217 |
| Completion time of Task C | 66.27±53.79 | 44.53±28.82 | 21.73 | 2.311 | 0.037\* |
| Errors of Task C | 3.33±3.15 | 2.67±2.44 | 0.67 | 1.128 | 0.278 |
| \*p<0.05 \*\*p< 0.01 | | | | | |

**Table 8.** UEQ evaluation results

|  |  |  |  |
| --- | --- | --- | --- |
| **Scale** | **Mean** | **Comparison to benchmark** | **Interpretation** |
| Attractiveness | 2.17 | Excellent | In the range of 10% best results |
| Perspicuity | 1.79 | Good | 10% of results better,75% of results worse |
| Efficiency | 1.60 | Good | 10% of results better,75% of results worse |
| Stimulation | 1.68 | Good | 10% of results better,75% of results worse |
| Novelty | 2.15 | Excellent | In the range of 10% best results |



**Figure 7.** Benchmark Bar Chart of UEQ results

**Table 9.** Paired t-test for training performance of children at advanced rehabilitation stage in Scenario understanding test game

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Game performance** | **Mean±Standard deviation** | | **Difference**  **value** | **t** | **p** |
| **Before the**  **Experiment** | **After the**  **Experiment** |
| Completion time of Task A | 13.00±5.74 | 13.00±4.55 | 0 | 0 | 1 |
| Completion time of Task B | 19.56±16.35 | 14.00±7.28 | 5.56 | 2.014 | 0.062 |
| Completion time of Task C | 12.81±6.36 | 11.94±3.59 | 0.88 | 0.76 | 0.459 |
| Completion time of  Task D | 11.50±3.46 | 12.50±3.10 | -1 | -3.464 | 0.003\*\* |
| Completion time of Task E | 15.88±12.27 | 12.38±3.98 | 3.5 | 1.475 | 0.161 |
| \*p<0.05 \*\*p< 0.01 | | | | | |

**Table 10.** Paired t-test for training performance of children at intermediate rehabilitation stage in Scenario understanding test game

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Game performance** | **Mean±Standard deviation** | | **Difference**  **value** | **t** | **p** |
| **Before the**  **Experiment** | **After the**  **Experiment** |
| Completion time of Task A | 20.87±13.29 | 15.53±6.35 | 5.33 | 2.442 | 0.028\* |
| Completion time of Task B | 23.20±10.88 | 18.40±6.78 | 4.8 | 3.091 | 0.008\*\* |
| Completion time of Task C | 17.93±11.66 | 13.87±6.49 | 4.07 | 2.8 | 0.014\* |
| Completion time of  Task D | 20.73±13.42 | 15.60±6.27 | 5.13 | 2.464 | 0.027\* |
| Completion time of Task E | 24.87±16.90 | 15.67±6.95 | 9.2 | 3.03 | 0.009\*\* |
| \*p<0.05 \*\*p< 0.01 | | | | | |

**Table 11.** Paired t-test for training performance of children at advanced rehabilitation stage in scenario-matching game

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Item** | **Mean±Standard deviation** | | **Difference**  **value** | **t** | **p** |
| **Before the**  **Experiment** | **After the**  **Experiment** |
| Prompts play in Task1 | 2.50±1.03 | 2.13±0.72 | 0.38 | 2.423 | 0.029\* |
| Completion time of Task1 | 94.63±43.35 | 81.31±29.34 | 13.31 | 3.247 | 0.005\*\* |
| Click count in Task1 | 3.75±1.18 | 3.44±0.51 | 0.31 | 1.159 | 0.264 |
| Prompts play in Task2 | 2.56±1.03 | 2.25±0.77 | 0.31 | 1.775 | 0.096 |
| Completion time of Task2 | 126.31±58.47 | 111.13±32.95 | 15.19 | 1.373 | 0.19 |
| Click count in Task2 | 5.06±0.93 | 4.88±0.89 | 0.19 | 1.145 | 0.27 |
| \*p<0.05 \*\*p< 0.01 | | | | | |

The performance of children in the intermediate stage of rehabilitation in the situational game reveals notable improvements. In Task 1, there were significant reductions in the number of times prompts were needed, completion times, and the number of clicks required, compared to their initial test. Moreover, completion times in Task 2 also showed a significant decrease compared to the first test.

In contrast to the results seen in advanced-stage children, intermediate-stage children did not initially possess strong skills in this area. Therefore, after training with the system, most aspects of their performance displayed significant enhancements, indicating the effectiveness of this system as a training tool.

**Table 12.**Paired t-test for training performance of children at intermediate rehabilitation stage in scenario-matching game

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Item** | **Mean±Standard deviation** | | **Difference**  **value** | **t** | **p** |
| **Before the**  **Experiment** | **After the**  **Experiment** |
| Prompts play in Task1 | 2.93±2.05 | 2.13±0.83 | 0.8 | 2.175 | 0.047\*\* |
| Completion time of Task1 | 113.73±61.30 | 82.47±18.00 | 31.27 | 2.396 | 0.031\* |
| Click count in Task1 | 4.13±1.73 | 3.13±0.64 | 1 | 2.958 | 0.010\*\* |
| Prompts play in Task2 | 3.27±1.28 | 3.07±096 | 0.2 | 1.146 | 0.271 |
| Completion time of Task2 | 161.20±69.66 | 99.33±46.22 | 61.87 | 3.328 | 0.005\*\* |
| Click count in Task2 | 4.80±0.94 | 4.27±0.70 | 0.53 | 2.086 | 0.056 |
| \*p<0.05 \*\*p< 0.01 | | | | | |