Article

Analyzing Visual Attention of People with Intellectual Disabilities during Virtual Reality-Based Job Training

Sungjin Hong [1], Heesook Shin [1], Younhee Gil [1] and Junghee Jo [2,*]

[1] Intelligent Knowledge Content Research Section, Communication & Media Research Laboratory, Electronics and Telecommunications Research Institute, Daejeon 34129, Korea; sjhong0117@etri.re.kr (S.H.); hsshin8@etri.re.kr (H.S.); yhgil@etri.re.kr (Y.G.)

[2] Department of Computer Education, Busan National University of Education, Busan 47503, Korea

* Correspondence: dreamer@bnue.ac.kr

Abstract: Virtual reality (VR) has proven an effective means of job training for people with intellectual disabilities who may experience difficulties in learning. However, it is unlikely for them to successfully complete a certain task using only VR-based job training contents without receiving supplemental help from others. Accordingly, to increase the effectiveness of virtual job training for people with intellectual disabilities in training situations in which they may experience difficulty and become unable to proceed further, the contents of the training program need to automatically identify such moments and provide support so that they may correctly perform the task. To identity the moment of intervention, we conducted an experiment ($n = 21$) to collect eye tracking data of people with intellectual disabilities while performing VR-based barista training. We measured eye scanning patterns to identify any difference between people with intellectual disabilities who complete a given step independently and those who request intervention. We found that the information about the types of fixated objects did not help to identify any difference, but the information about eye transition, eye movements between two different areas of interest, was useful in identifying the difference. Our findings provide implications for identifying the moment of intervention for people with intellectual disabilities.

Keywords: virtual reality; intellectual disabilities; behavior analysis; visualization; job training; eye tracking; visual attention

1. Introduction

Virtual reality (VR) is an important technology facilitating job training [1–4]. For risky jobs, VR-based training enables students to safely make and correct their mistakes in most situations. This lets them avoid potential risks they might face while providing sufficient opportunities to apply their theoretical knowledge in secure virtual environments. For example, in the medical field, VR allows learning surgical skills by manipulating the virtual anatomic structures of the human body, but avoiding potentially risky or stressful situations [5,6].

In addition to its general applications, VR-based training has been widely used by people with intellectual disabilities. In this paper, we define the term intellectual disability as a condition in which a person has mild limitations in intellectual functions. According to the Korean Disability Welfare Act, these are people with 3rd degree intellectual disability, and they are able to participate in the usual classes of normal schools and receive an education adequate to get a job and lead a social life [7]. With increased integration of people with intellectual disabilities into society, vocational education for the disabled has become increasingly important [8]. Due to their limitations in intellectual ability, it is probable that they often require investments of time and repetitive training to master a certain task [9]. This may prolong the learning curve of people with intellectual disabilities and become the main obstacles to employers hiring them [10].

To broaden their employment opportunities, various attempts, e.g., enacting laws or incentives to promote increased employment, have been conducted worldwide. In some countries, governments have established job training centers for people with intellectual disabilities [11,12]. These training centers help them become socially self-reliant by providing opportunities to practice vocational skills in the same environment as that of a real workplace. The job experience facilities offer experience and training in vocational areas, such as being a barista, steam car washer, librarian, nurse's aide, and packager which offers them a better chance of success. Currently, job training conducted with the help of teachers within limited time and space does not allow trainees to independently practice through repetitive actions. In addition, acquiring the skills necessary in an actual workplace poses a safety hazard during training.

VR has proven an effective means of learning for individuals with intellectual disabilities. Studies have shown that VR has a positive impact on the learning attitude and academic achievement of students with intellectual disabilities who may experience difficulty learning [13–16]. A possible impediment to VR training is concern that people with intellectual disabilities may have difficulty manipulating VR devices; however, it has been demonstrated that most people with intellectual disabilities can fully control VR-based systems if they receive prior instruction about the VR devices [17–19]. One study reported that HMDs were enjoyable, easy to use, and even exciting for most children with intellectual disabilities [20].

Despite being able to readily use of VR-based systems, individuals have required supplemental assistance to successfully complete VR- based job training. Accordingly, to increase the effectiveness of virtual job training for people with intellectual disabilities in training situations in which they may experience difficulty and become unable to proceed further, the contents need to automatically identify such moments and provide support so that they may correctly perform the task. In this paper, this kind of assistance is defined as an 'intervention'. The type of intervention can include direct help from a person or voice/video provided by the VR-based job training contents. Although several research studies have discussed the effectiveness of VR-based job training for people with intellectual disabilities, very few have proposed specific strategies to automatically identify the moment needing intervention.

This paper offers a method of using eye tracking information to identify the moment needing intervention for people with intellectual disabilities while performing VR-based job training. According to research results studying eye tracking, an individual's interest at a specific moment is highly likely to coincide with the object being gazed at [21,22]. Our study utilizes these findings to identity the moment people with intellectual disabilities may encounter difficulties during training and are unable to proceed any further. We focus on the occupation of barista, as it is one of the most preferred occupations for people with intellectual disabilities. According to a recent study, being a barista is not only the top priority job to which students with intellectual disabilities aspire but also the job most preferred by their parents [23].

2. Related Work

Eye tracking is a method for recording and assessing people's eye movements while they perform a task. The assumption underlying the use of eye-tracking technology is that there is a connection between where an individual gazes and what the individual is focusing on or thinking about at that point. Cognitive processes are considered to be complicated, and it is possible that an individual may be gazing at one thing but thinking about something else. Nevertheless, a significant number of studies have acknowledged that eye movement gives insight into understanding an individual's cognitive processes [24–26].

Eye tracking research commonly uses visual scanning behavior, named eye fixation. Eye fixation occurs when the eye is focused on a particular point for at least 100–200 ms. Eye transition, eye movements between two different areas of interest, is also a frequently used concept in eye tracking research. The number of eye transitions implies how frequently an individual switches his or her eye fixations from one point to another, with higher numbers indicating more transitions. Figure 1 illustrates the definition of eye transition.
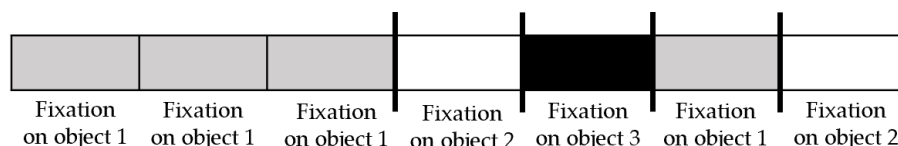
Figure 1. Example of eye transition: 4 transitions.

Eye-tracking research is widely used in various fields. In the medical field, an error that often occurs during the process of verifying patient identification was detected by analyzing nurses' eye tracking data. That is, it was determined which information in the patient's chart, identification band, and medication label was being fixated on [27,28]. In our study, we measure eye fixation and transition to capture the objects in the virtual environment of a barista that people with intellectual disabilities focused on. These analysis results could be used to identity the moment the trainees are unable to proceed any further and encounter difficulties during barista training.

3. Experimental Design

3.1. Participants

A total of 21 male and female participants in the training center for the intellectual disabilities were selected as subjects for the study. A recruitment notice was posted on the bulletin board of the training center, and the notice was explained during class by the teacher in the center. The recruitment was conducted regardless of gender, and the age range was limited to adults aged 18 to 50 years old. Those who might cause sudden disruptions during the experiment due to severe autism or mental illness were excluded. When a student wished to participate in the experiment, the recruitment notice and consent form were fully explained in person or by phone call. His or her parents after hearing a sufficient explanation of the purpose and methods of the experiment decided whether to participate. Only after the parents' permission was obtained was the student selected as a subject of the study.

3.2. Ethical Considerations

Issues about protecting personal information were addressed as follows. To safeguard personal information, a unique identifier was assigned to each subject and was used instead of a name throughout the entire research process. Mapping between the name and the unique identifier as well as records or data that could identify a subject were restricted only to those researchers directly participating in the study. Personal information and research data will be stored for three years after the end of the study, and afterwards, all electronic and paper-type information will be discarded.

3.3. Settings

The experiments were conducted in the training center for intellectual disabilities where the subjects were registered, which is an environment that was easy to visit and familiar to people with intellectual disabilities. Visual and auditory stimuli generated in the VR-based barista training contents during the experiment may cause dizziness in subjects. In anticipation of such a situation, a teacher at the center who has specialized knowledge about the characteristics of people with intellectual disabilities was close at hand in the experimental setting in order to handle an unexpected situation. If the subject were to

complain of eye fatigue or dizziness, preparations had been made to transfer the subject to a hospital, but this situation did not occur during the entire experiment.

3.4. Instruments

Three instruments were used in this study: An HMD, controllers, and a VR-based barista training contents. The HMD used in this study was one of the commercial HMDs developed by Looxid Labs [29]. This headset was equipped with eye-tracking cameras and so could collect eye-related data, ranging from pupil dilation and movements to blinking patterns. VR-based barista training was performed with left and right controllers. In the VR environment, they are shown as left and right hands. Figure 2 shows the HMD and controllers used in this study.

| (**a**) | (**b**) |

Figure 2. A screenshot of the head mounted display and controllers: (**a**) HMD; (**b**) controllers.

The VR-based barista training contents consisted of 'terminology education' and 'barista education'. For the initial step of the training, terminology education was conducted to perform cognitive training on a single barista object (e.g., "Find an espresso machine and touch it with your hands"). Educating the students about terminology consisted of 13 steps. At each step, the barista objects which students were trained to recognize were: An espresso machine (step 1), grinder (step 2), group head (step 3), portafilter (step 4), tamper (step 5), tamper pedestal (step 6), shot glasses (step 7), knock box (step 8), steam milk pitcher (step 9), steam valve (step 10), steam nozzle (step 11), dry linen cloth (step 12), and wet linen cloth (step 13). Each barista object used in terminology education training is shown in Figure 3.

Figure 3. Barista objects in virtual reality-based barista training contents.

When the terminology training began, the training contents provided a voice guide for the first step: "Find an espresso machine and touch it with your hands". When the subject touched the espresso machine, the word "espresso machine" was shown near the espresso machine in the virtual space, followed by a voice guide for the next step: "Find the grinder and touch it with your hands". When a barista object named by the training contents by voice was touched, the voice guidance for the next barista object was continued and so on. The education phase of the study was completed when the wet linen cloth, the last barista object, was touched. Figure 4 shows a screen capture immediately after the subject touched the grinder in accordance with the voice guidance.

Figure 4. A screenshot of the terminology education.

The barista education consisted of 21 steps, as shown in Table 1. Barista education was also conducted similarly to the terminology education. At the start, voice guidance for the first step announced: "Remove the portafilter from the espresso machine". As soon as the subject removed the portafilter from the espresso machine, voice guidance was provided for the next step: "Wipe the portafilter with a dry linen cloth". This manner was repeated until all 21 steps were successfully completed.

**Table 1.** Steps comprising the barista education.

Steps	Tasks
1	Remove the portafilter from the espresso machine
2	Wipe the portafilter with a dry linen cloth
3	Insert the portafilter into the coffee grinder
4	Turn on the coffee grinder
5	Grind the coffee beans by pulling the coffee grinder lever
6	Turn off the coffee grinder
7	Remove the portafilter from the coffee grinder
8	Tap the coffee powder in the portafilter evenly
9	Place the portafilter on the tamper base
10	Press the coffee powder in the portafilter to make it hard by using tamper
11	Remove the coffee powder on the surface of the portafilter by hand
12	Remove any remaining debris by pouring water on the group head
13	Attach the portafilter to the coffee machine
14	Place the shot glasses under the portafilter
15	Press the espresso brew button to extract espresso
16	Move the shot glasses onto the tray
17	Remove the portafilter from the espresso machine
18	Throw away the residue left on the portafilter in the knock box
19	Rinse the portafilter with water
20	Wipe the portafilter with a dry linen cloth
21	Attach the portafilter to the coffee machine

3.5. Procedure

The experimental procedure approved by the Institutional Review Board at Pusan National University, South Korea was as follows. When the subject arrived at the experimental setting, a researcher first allowed him to take a rest and then performed the experimental procedures and associated precautions. With the help of the researcher, the subject wore an HMD and practiced holding the controller in his or her hands and operating it. After becoming familiar with the HMD and controllers, the subject began training starting with education about terminology. If the subject was unable to proceed further while performing the training, intervention was requested in accordance with a previously agreed upon signal method, i.e., the subject raising a hand or speaking. When a request for intervention was made, the researcher recorded the point to identify it as a 'ground truth' for judging the timing of the intervention and then provided help in an audio format. Even though voice intervention was provided, if the training was not continuously performed, the subject was asked to watch the helper video after which the training was resumed. After the experiment had concluded, the subject was given sufficient time to rest and left after a brief interview which had been prepared in advance.

3.6. Research Hypotheses

The goal of this study was to identify the point of intervention required when people with intellectual disabilities undergo barista training using VR-based contents. To achieve this goal, we established the following hypotheses and verified them by analyzing subjects' eye-tracking data. Here we use the term 'barista objects' to indicate objects used for barista work (e.g., espresso machine) and 'environmental objects' to mean objects not used for barista work (e.g., table). Figure 5 shows the relationship among objects appearing in VR-based barista training contents. Objects are the virtual space which can be classified as either barista objects or environmental objects. Barista objects can be classified into objects associated with a given task or not associated with a given task, depending on whether the objects are essential to perform the task or not.
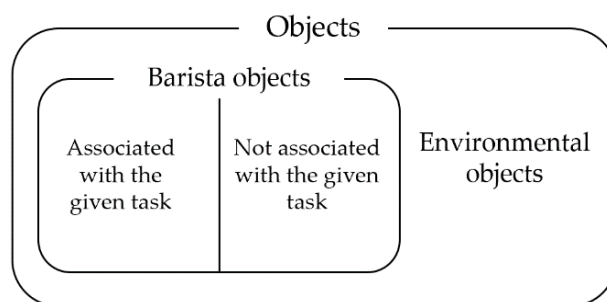
Figure 5. Relationship among objects appearing in VR-based barista training contents.

Hypothesis 1. *People with intellectual disabilities who complete a given step independently will have greater numbers of eye fixations on barista objects. On the other hand, those who request intervention will have greater numbers of eye fixations on environmental objects.*

Hypothesis 1 is motivated by the presumption that subjects who completed a step without asking for any assistance may focus their attention on barista objects such as espresso machines, grinders, and portafilters. Subjects in need of intervention, however, may be frustrated at the moment they are unable to proceed any further perhaps because they do not know what to do. This may cause their attention to be distracted to environmental objects such as windows, tables, and chairs.

Hypothesis 2. *People with intellectual disabilities who complete a given step independently will have greater numbers of eye fixations on barista objects associated with a given task (e.g., grinder in step 2). On the other hand, those who request intervention will have greater numbers of eye fixations on other objects, i.e., barista objects not associated with the given task and environmental objects.*

Hypothesis 2 is motivated by the presumption that subjects who completed a step without asking for any intervention may focus their attention on barista objects, but only those necessary to complete the given task in the step. For example, subjects who have passed the first step during terminology education may focus mostly on the espresso machine. Subjects in need of intervention, however, may be distracted by other barista objects such as grinders or portafilters, which are not related to the first step, along with environmental objects.

Hypothesis 3. *People with intellectual disabilities who complete a task will have fewer numbers of eye transitions than those who request intervention.*

Hypothesis 3 is motivated by the presumption that subjects who completed a step without asking for any intervention may focus on a fewer number of specific objects, while those who request intervention may scatter their attention to various objects inconsistently. Especially, this hypothesis differs from H1 and H2 in that it is possible to prove whether this hypothesis is correct, without any detailed information what object the eye fixated on.

4. Analysis Approach

4.1. Data Collection

The data collected in this study were the name of the object the subject's eyes were focusing on and the time of that moment during the entire training process. These were collected on an eye tracker attached to the HMD. Figure 6 shows part of the collected raw eye tracking data.

00:01:13:265	Grinder
00:01:13:279	Grinder
00:01:13:311	Table
00:01:13:323	Table
00:01:13:339	Table
00:01:13:367	Table
00:01:13:375	Table
00:01:13:412	Table
00:01:13:419	Table

Figure 6. Example of the collected raw eye tracking data.

In addition to the text form of the eye tracking data, video data was also collected. The entire experimental process was recorded to observe the subject's movements. Figure 7 shows the virtual space (top left), the subject's behavior (bottom left), and eye-tracking information (yellow arrow). These three types of videos were integrated to be played simultaneously so that they could be observed altogether throughout the entire training process.
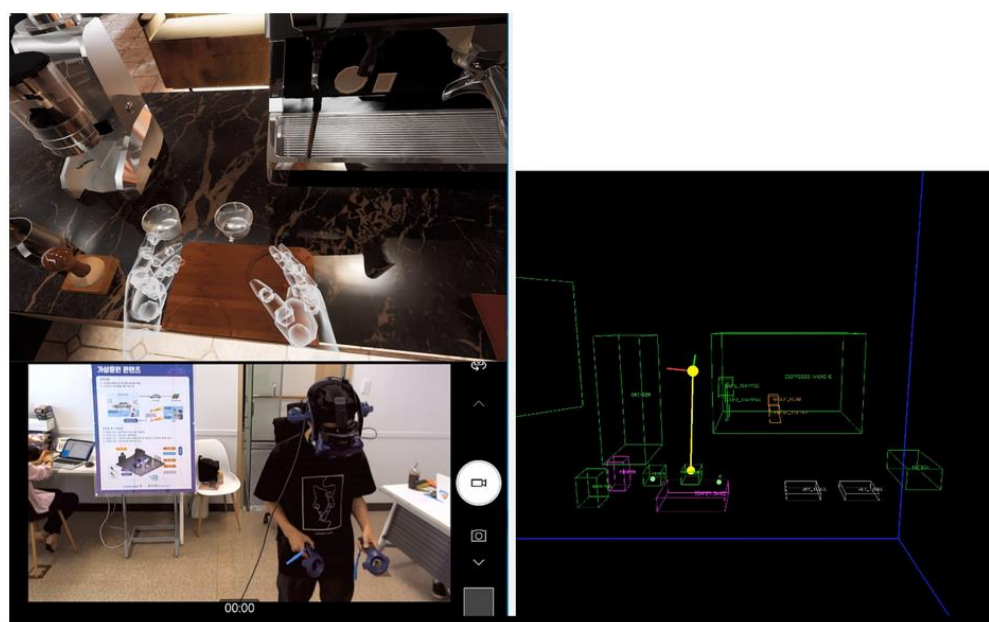
Figure 7. A screenshot of the video data displaying the virtual space, subject's behavior, and eye-tracking data simultaneously.

4.2. Data Processing

Collected eye-tracking data were processed for analysis as follows. First, the single data file generated during the 13 steps of terminology education for each subject was separated into 13 files. Similarly, the single data file collected during barista education of each subject was divided into 21 steps, thereby creating 21 separate files. Thus, the total amount of data to be analyzed was: 273 (=21 subjects × 13 steps) for terminology education; 441 (=21 subjects × 21 steps) for barista education.

Analysis of the data at each step determined whether the subject passed the step without intervention, or requested intervention. Each step is labeled as 'Pass' for the former and 'Fail' for the latter. Finally, each subject's data are composed of multiple steps classified either Pass or Fail, as shown in Figure 8.
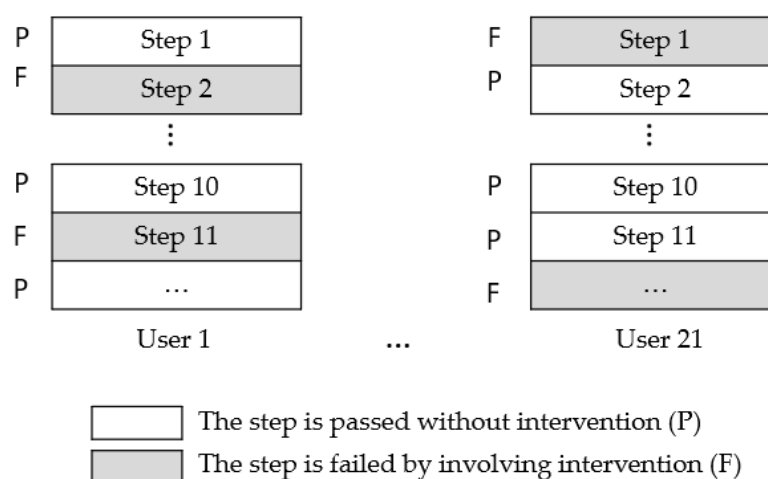
Figure 8. Step labeled with pass (P) and fail (F).

For this analysis, in the case of a step classified as Pass, all data from the beginning to the end of the step were included in the analysis. In the case of a step classified as Fail, however, only the data from the start of the step to just before the request for intervention were included in the analysis, as shown in Figure 9. This is because during the time between when the intervention is requested and then granted, the subject is not acting using his or her own free will, so the data are contaminated and cannot be used to determine the timing of the intervention.
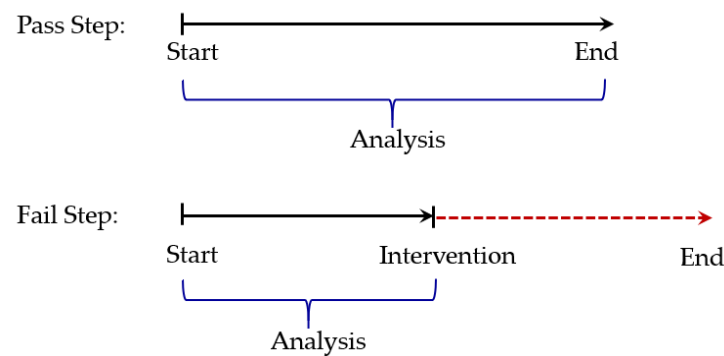
Figure 9. Data included in the analysis from each of the two groups.

4.3. Data Analysis (H1)

To analyze H1, we used the following analysis methods. We reasoned that steps belonging to the Pass group have a greater number of fixations on barista objects; steps belonging to the Fail group have a greater number of fixations on other objects, i.e., environmental objects.

1. We measured the number of fixations on barista objects and environmental objects, respectively.
2. For each step belonging to the Pass group, we evaluated whether the number of fixations on barista objects (BO) is greater or less than the number of fixations on any remaining objects (RO), i.e., environmental objects, by calculating the fixation ratio (FR), as in (1).

$$FR\ (H1) = \#\ of\ fixations\ on\ BO / \#\ of\ fixations\ on\ RO \tag{1}$$

A value greater than 1 means that the number of fixations on barista objects is greater than the number of fixations on environmental objects. On the other hand, a value less than 1 means that the number of fixations on barista objects is less than the number of fixations on environmental objects.

3. For each step belonging to the Pass group, we determined whether the value of FR is greater or less than 1. In the same way, for each step belonging to the Fail group, we determined whether the value of FR is greater or less than 1.
4. Using two variables, categories of groups (i.e., Pass, Fail) and categories of FR (i.e., FR > 1, FR < 1), we determined whether any significant relationship exists between these two variables, using the Pearson Chi-Square test based on the 2×2 contingency table. The test was conducted using a web-based Chi-Square calculator [30].

4.4. Data Analysis (H2)

To analyze H2, we used the following analysis methods, similar to H1. We reasoned that steps belonging to the Pass group have a greater number of fixations on barista objects but are limited to only objects required to complete the task (e.g., grinder in step 2); steps belonging to the Fail group have a greater number of fixations on other objects, i.e., barista objects not required to complete the task as well as environmental objects.

1. We measured the number of fixations on barista objects but counted only those objects required to complete the task. In addition, we also counted the number of fixations on barista objects not required to complete the task as well as environmental objects.
2. For each step belonging to the Pass group, we evaluated whether the number of fixations on barista objects required to complete the task (BOR) is greater or less than the number of fixations on the remaining objects (RO), by calculating FR, as (2).

$$FR\ (H2) = \text{\# of fixations on BOR}/\text{\# of fixations on RO} \tag{2}$$

In H2, the remaining objects include barista objects not required to complete the step as well as environmental objects.

3. For each step belonging to the Pass group, we determined whether the value of FR is greater or less than 1. In the same way, for each step belonging to the Fail group, we determined whether the value of FR is greater or less than 1.
4. Using two variables, categories of groups (i.e., Pass, Fail) and categories of FR (i.e., FR > 1, FR < 1), we determined whether any significant relationship exists between these two variables, using the Pearson Chi-Square test based on the 2×2 contingency table. The test was conducted using a web-based Chi-Square calculator [30].

4.5. Data Analysis (H3)

To analyze H3, we used the following analysis methods. For this hypothesis, we reasoned that steps belonging to the Pass group have more consistent eye fixations than steps belonging to the Fail group.

1. We measured the number of transitions. This is to determine how frequently the subjects shifted their fixations between objects.
2. To address individual differences for eye fixation patterns of each subject, we normalized the number of transitions.
3. Using an unpaired *t*-test, we determined whether any significant relationship exists between the Pass group and Fail group with regards to the number of eye transitions. The test was conducted using a web-based *t*-test calculator for 2 independent means [31].

5. Results

Of the 273 steps (21 subjects performing 13 steps) for terminology education, 3 (1%) did not have sufficient quality to judge what objects the participant was looking at due to eye tracking failures. Of the 441 steps (21 subjects performing 21 steps) for barista education, 63 (14%) were excluded because of 3 subjects' eye tracking failures through all steps of the training. Thus, we included 270 steps for terminology education and 378 steps for barista education in our analysis.

Figure 10 shows the proportion of passed and failed subjects at each step of terminology education. Step 6 ("Please find the tamper pedestal and touch it with your hands") was one that most subjects (95%) passed by themselves without requesting intervention. It was followed by step 5 ("Please find the tamper and touch it with your hands") and step 13 ("Please find the wet cloth and touch it with your hands") with 90% of subjects passing. On the other hand, Step 1 (espresso machine) was one that the majority of subjects (90%) failed by requesting intervention. Figure 11 shows the proportion of passed and failed subjects at each step of barista education. Compared to the terminology education, a considerable number of subjects passed the steps, but in the case of step 12 ("Remove remaining debris by pouring water on the group head"), only 33% of the subjects passed.
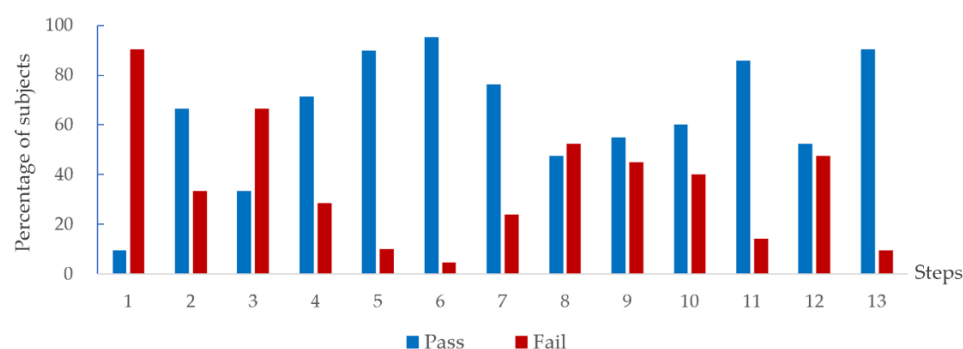
Figure 10. Percentage of passed and failed subjects at each step of terminology education.
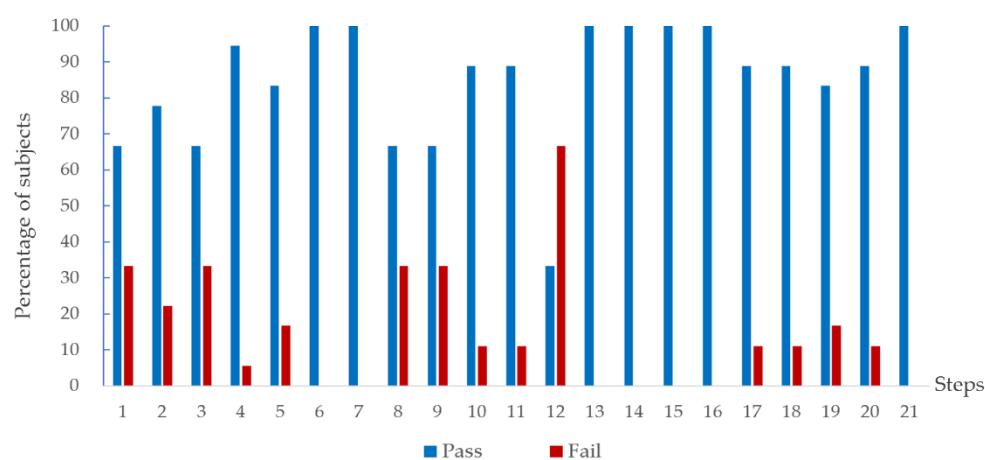
Figure 11. Percentage of passed and failed subjects at each step of barista education.

With regard to H1, we found that there is no significant relationship between the two variables, categories of groups (i.e., pass, fail) and categories of FR (i.e., FR > 0, FR < 0), and so rejected the hypothesis ($p > 0.05$). We hypothesized that subjects who passed steps would fixate more on barista objects; on the other hand, subjects who failed steps would fixate more on environmental objects. However, they tended not to shift their eyes to environmental objects at the moment of failure, but still carefully searched while fixating their eyes among the barista objects.

With regard to H2, we found that there is no significant relationship between the two variables, categories of groups (i.e., Pass, Fail) and categories of FR (i.e., FR > 0, FR < 0), and so rejected the hypothesis ($p > 0.05$). We hypothesized that subjects who passed steps would fixate more on barista objects, but this would be limited to only objects required to complete the task; on the other hand, subjects who failed steps would fixate more on other objects (i.e., barista objects not required to complete the task and environmental objects). However, most of the subjects belonging to the Pass group actually tended to focus frequently on specific barista objects (e.g., grinder), which were not relevant to the step, but could pass the step by focusing on the objects relevant to the step just for a short period of time, but not most of the time as we had assumed.

With regard to H3, we found that there is a significant relationship between subjects' eye transitions belonging to the Pass group versus subjects' eye transitions belonging to the Fail group and so accepted the hypothesis ($p < 0.001$). Tables 2 and 3 show the analysis results of the unpaired *t*-tests for terminology education and barista education, respectively. As we hypothesized, the subjects belonging to the Pass group tended to stare at specific objects for a certain period of time, while the subjects belonging to the Fail group tended to switch their gaze randomly between several objects.

Table 2. Results of unpaired *t*-test for hypothesis 3 for terminology education.

Terminology Education	M [1]	SD [2]	*t*	*p*
Pass	0.18	0.23	−8.51	$p < 0.001$
Fail	0.48	0.35		

[1] M stands for mean. [2] SD stands for standard deviation.

Table 3. Results of unpaired *t*-test for hypothesis 3 for barista education.

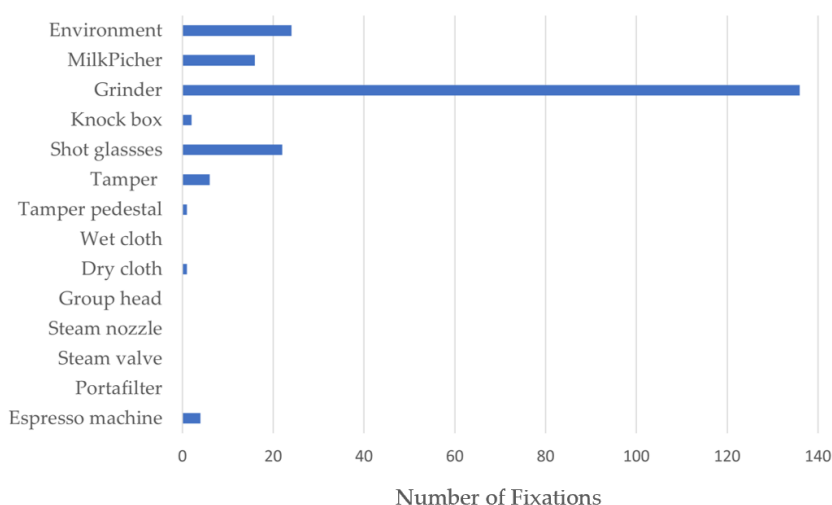| Barista Education	M [1]	SD [2]	*t*	*p*
Pass	0.19	0.24	−6.64	$p < 0.001$
Fail	0.41	0.32		

[1] M stands for mean. [2] SD stands for standard deviation.

6. Discussion

Our analyses revealed important findings for how to automatically identify the moment needing intervention while performing VR-based barista training. With regard to the test results of H1 and H2, the information about the types of fixated objects (e.g., barista objects, environmental objects) may not be useful in identifying the difference between people with intellectual disabilities who complete a given step independently and those who request intervention.

The reasons for rejecting H1 included the following. At the moment when it was difficult to perform the task, people with intellectual disabilities tended not to avoid such situations by shifting their eye fixations on objects not related to the work (e.g., windows, tables) but instead focused their attention more on the given task. Some people with intellectual disabilities did not move their hands at such moments, but their gaze was still inclined to switch back and forth between various barista objects.

For example, Figure 12 displays the number of fixations on each object while one subject conducted step 1 ("Please find the espresso machine and touch it with your hands") of terminology education. This subject failed the step but the subject's fixation remained mostly on "grinder" which is one of the barista objects, but not on the environmental objects as we had assumed. Another example is shown in Figure 13, describing the number of fixations on each object while one subject conducted step 20 ("Wipe the portafilter with a dry linen cloth") of barista education. The subject failed step 20 but looked at the knock box slightly more than the environment.
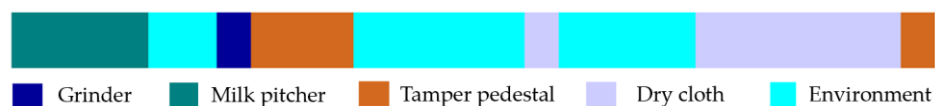

Figure 12. Example of the number of fixations on each object: Terminology education, step 1, fail.
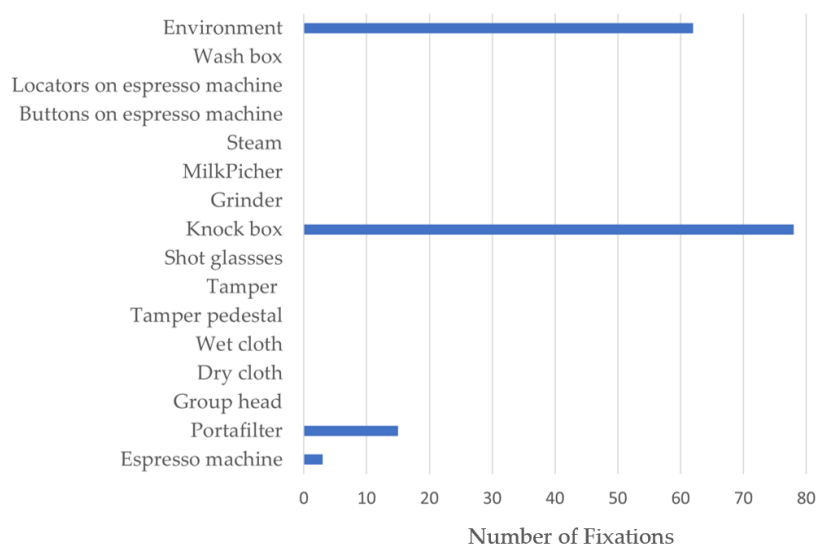
Figure 13. Example of the number of fixations on each object: Barista education, step 20, fail.

The reasons for rejecting H2 included the following. People with intellectual disabilities who independently completed a given task tended not to pay considerable attention to only those barista objects directly related to work from the outset. Rather, at the beginning of training there was a tendency to spend a lot of time on large and conspicuous objects, such as the grinder, and later turn their attention to other objects.

For example, Figure 14 describes the number of fixations on each object while one subject conducted step 6 ("Please find the tamper pedestal and touch it with your hands") of terminology education. This subject tended to focus mostly on environmental objects and other barista objects such as the milk picture and dry cloth but could pass the step by briefly focusing on the tamper pedestal.

Figure 14. Example of visualization based on eye fixations: Terminology education, step 6 ("Find the tamper pedestal and touch it with your hands), pass.

Another example displayed in Figure 15 shows the number of fixations on each object while one subject conducted step 10 ("Remove the coffee powder on the surface of the portafilter by hand"). This task can be completed by removing the coffee powder on the surface of the portafilter by hand, but the subject constantly pressed the coffee powder with the tamper for a long time, as shown in Figure 16.

Figure 15. Example of visualization based on eye fixations: Barista education, step 10 ("Remove the coffee powder on the surface of the portafilter by hand"), pass.

Figure 16. A screenshot of the moment when the subject focused on the tamper shown in Figure 15.

However, it was possible to complete the step in a short time when the subject separated the tamper from the portafilter and touched the portafilter with his hand. Note that the 'tamper stand' in Figure 15 is the object on which to place the portafilter when compressing the coffee powder to make it hard by using the tamper, while the 'tamper pedestal' means the object on which to place the tamper which is located on the left side of the tamper stand in Figure 16.

The reasons for accepting H3 included the following. People with intellectual disabilities who independently completed a given task may have more consistent eye fixation patterns than those who request intervention. One implication of this finding may be that when people with intellectual disabilities encounter difficulties during training and are unable to proceed any further, their eyes rapidly fixate on a large number of random objects highly likely to be related to the given task rather than fixate on a small number of particular objects. Note that in the case of a step classified as Fail, we included in the analysis only the data from the start of the step to just before the request for intervention because during the time between when the intervention is requested and then granted, the subjects were not acting according to their own free will, so the data were contaminated. On the other hand, in the case of a step classified as Pass, we included all the data from the start to the end of accomplishing the given task. Nevertheless, the eye transition ratio of the Fail group was overwhelmingly higher than that of the Pass group.

There are several limitations to this study. First, we conducted the experiment by classifying people with intellectual disabilities into one group. All subjects for this study were people with mild intellectual disabilities. Since all of them have similar degrees of intellectual disabilities, we did not divide them into separate groups for the experiment. However, the results obtained from this study may differ depending on the range of intellectual disabilities. Second, we used eye tracking data, although they are not completely accurate, as a measurement for visual attention. Eye tracking data, however, have been commonly used for this purpose in other studies and provide us with valuable insight into understanding the visual attention patterns of people with intellectual disabilities. Third, the subjects in our study were selected from a single training center for people with intellectual disabilities. We do not know whether the observed subjects' behavior can be generalized to other people with intellectual disabilities in other settings. Finally, the experimental setting may not accurately reflect the actual setting where there is less time pressure. We would expect more subjects to independently complete a task in a real world setting than in an experimental setting with its increased pressure of time.

7. Conclusions

Focusing on the occupation of barista, we explored a method using eye tracking information to detect the moment when people with intellectual disabilities needed intervention while performing VR-based training. We conducted an experiment to measure eye scanning patterns to identify any difference between people with intellectual disabilities

who complete a given step independently and those who request intervention. We found that information about the types of fixated objects was not useful in identifying a difference, but information about eye transition was helpful to identify any difference. These findings suggest a need for further study of visual scanning patterns. For example, similar to eye transitions, eye saccades are the quick movement of the eye from one point of fixation to another point of fixation. The difference is that transition occurs between two different types of objects (e.g., grinder → espresso machine), whereas saccade occurs between two objects regardless of types (e.g., one point of the grinder → another point of the grinder). Measuring the frequency of eye saccades may help provide further information whether the eye fixations of people with intellectual disabilities transit not only between different types of objects but also between different points within the same object. In addition, it would be informative to compare the results of this study with those obtained of people with normal intellectual functioning. Applying the analysis methods used in this paper to other types of occupations could determine whether the findings of this paper can be generalized.

Institutional Review Board Statement: This study was conducted with the approval of the Institutional Review Board at Pusan National University, South Korea.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. De Ponti, R.; Marazzato, J.; Maresca, A.M.; Rovera, F.; Carcano, G.; Ferrario, M.M. Pre-graduation medical training including virtual reality during COVID-19 pandemic: A report on students' perception. *BMC Med. Educ.* **2020**, *20*, 1–7. [CrossRef] [PubMed]
2. Carruth, D.W. Virtual reality for education and workforce training. In Proceedings of the 15th International Conference on Emerging eLearning Technologies and Applications (ICETA), Starý Smokovec, Slovakia, 26–27 October 2017.
3. Marone, E.M.; Rinaldi, L.F. Educational value of virtual reality for medical students: An interactive lecture on carotid stenting. *J. Cardiovasc. Surg.* **2018**, *59*, 650–651. [CrossRef] [PubMed]
4. Němec, M.; Fasuga, R.; Trubač, J.; Kratochvíl, J. Using virtual reality in education. In Proceedings of the 15th International Conference on Emerging eLearning Technologies and Applications (ICETA), Starý Smokovec, Slovakia, 26–27 October 2017.
5. Maresky, H.S.; Oikonomou, A.; Ali, I.; Ditkofsky, N.; Pakkal, M.; Ballyk, B. Virtual reality and cardiac anatomy: Exploring immersive three-dimensional cardiac imaging, a pilot study in undergraduate medical anatomy education. *Clin. Anat.* **2019**, *32*, 238–243. [CrossRef] [PubMed]
6. Moro, C.; Štromberga, Z.; Raikos, A.; Stirling, A. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anat. Sci. Educ.* **2017**, *10*, 549–559. [CrossRef] [PubMed]
7. Kim, G.E.; Chung, S. Elderly mothers of adult children with intellectual disability: An exploration of a stress process model for caregiving satisfaction. *J. Appl. Res. Intellect. Disabil.* **2016**, *29*, 160–171. [CrossRef] [PubMed]
8. Collins, J.C.; Ryan, J.B.; Katsiyannis, A.; Yell, M.; Barrett, D.E. Use of portable electronic assistive technology to improve independent job performance of young adults with intellectual disability. *J. Spec. Educ. Technol.* **2014**, *29*, 15–29. [CrossRef]
9. Kwon, J.; Lee, Y. Serious games for the job training of persons with developmental disabilities. *Comput. Educ.* **2016**, *95*, 328–339. [CrossRef]
10. Houtenville, A.; Kalargyrou, V. People with disabilities: Employers' perspectives on recruitment practices, strategies, and challenges in leisure and hospitality. *Cornell Hosp. Q.* **2012**, *53*, 40–52. [CrossRef]
11. U.S. Department of Labor. Available online: https://www.dol.gov/general/topic/training/disabilitytraining (accessed on 24 May 2021).
12. Ministry of Employment and Labor. Available online: https://www.moel.go.kr/english/poli/poliNewsnews_view.jsp?idx=1345 (accessed on 24 May 2021).

13. McComas, J.; Pivik, P.; Laflamme, M. Current uses of virtual reality for children with disabilities. In *Studies in Health Technology and Informatics*; IOS Press: Amsterdam, The Netherlands, 1998; pp. 161–169.

14. Standen, P.J.; Brown, D.J. Virtual reality in the rehabilitation of people with intellectual disabilities. *Cyberpsychol. Behav.* **2005**, *8*, 272–282. [CrossRef] [PubMed]

15. Rose, D.; Brooks, B.M.; Attree, E.A. Virtual reality in vocational training of people with learning disabilities. In Proceedings of the 3rd international conference on disability, virtual reality, and associated technology (ICDVRAT), Alghero, Sardinia, 23–25 September 2000; pp. 129–136.

16. Cunha, R.; da Silva, R.L.D.S. Virtual reality as an assistive technology to support the cognitive development of people with intellectual and multiple disabilities. In Proceedings of the Brazilian Symposium on Computers in Education (SBIE), Recife, Brazil, 30 October–2 November 2017; pp. 987–996.

17. da Costa, R.M.E.M.; de Carvalho, L.A.V. The acceptance of virtual reality devices for cognitive rehabilitation: A report of positive results with schizophrenia. *Comput. Methods Programs Biomed.* **2004**, *73*, 173–182. [CrossRef]

18. Newbutt, N.; Sung, C.; Kuo, H.J.; Leahy, M.J. The potential of virtual reality technologies to support people with an autism condition: A case study of acceptance, presence and negative effects. *Annu. Rev. Cyber Ther. Telemed.* **2016**, *14*, 149–154.

19. Lee, T.S. The Effect of Virtual Reality based Intervention Program on Communication Skills in Cafe and Class Attitudes of Students with Intellectual Disabilities. *J. Korea Converg. Soc.* **2019**, *10*, 157–165.

20. Newbutt, N.; Bradley, R.; Conley, I. Using virtual reality head-mounted displays in schools with autistic children: Views, experiences, and future directions. *Cyberpsychol. Behav. Soc. Netw.* **2020**, *23*, 23–33. [CrossRef] [PubMed]

21. Just, M.A.; Carpenter, P.A. A theory of reading: From eye fixations to comprehension. *Psychol. Rev.* **1980**, *87*, 329–354. [CrossRef] [PubMed]

22. Deubel, H.; O'Regan, K.; Radach, R. Attention, information processing and eye movement control. In *Reading as a Perceptual Process*; Elsevier: Oxford, UK, 2000; pp. 355–374.

23. Daejeon Welfare Foundation. Available online: https://daejeon.pass.or.kr/board.es?mid=a11001010100&bid=0024&tag=&act=view&list_no=6785 (accessed on 24 May 2021).

24. Jang, J.H.; Sung, J.E. Age-Related Differences in Sentence Processing of Who-Questions: An Eye-Tracking Study. *Commun. Sci. Disord.* **2020**, *25*, 382–398. [CrossRef]

25. Rayner, K. Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* **1998**, *124*, 372–422. [CrossRef] [PubMed]

26. Hyönä, J.; Lorch, R.F., Jr.; Kaakinen, J.K. Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *J. Educ. Psychol.* **2002**, *94*, 44–55. [CrossRef]

27. Marquard, J.L.; Henneman, P.L.; He, Z.; Jo, J.; Fisher, D.L.; Henneman, E.A. Nurses' behaviors and visual scanning patterns may reduce patient identification errors. *J. Exp. Psychol. Appl.* **2011**, *17*, 247–256. [CrossRef] [PubMed]

28. Jo, J.; Marquard, J.L.; Clarke, L.A.; Henneman, P.L. Re-Examining the Requirements for Verification of Patient Identifiers During Medication Administration: No Wonder It Is Error-Prone. *IIE Trans. Healthc. Syst. Eng.* **2013**, *3*, 280–291. [CrossRef]

29. Looxid Link. Available online: https://looxidlabs.com/looxidlink/product/looxid-link-package-for-vive-pro (accessed on 24 May 2021).

30. Chi Square Calculator for 2×2. Available online: https://www.socscistatistics.com/tests/chisquare/default.aspx (accessed on 26 June 2021).

31. T-Test Calculator for 2 Independent Means. Available online: https://www.socscistatistics.com/tests/studenttest/default2.aspx (accessed on 26 June 2021).