International Journal of Serious Games

ISSN: 2384-8766
https://journal.seriousgamessociety.org/

Article

Training Conversational Skills Using a VR Game: Effect of Immersion

Rianne de Wit[1], Hugo Huurdeman[2,3]

[1] Lectoraat Urban Care & Education, Windesheim University of Applied Sciences, Almere, the Netherlands; [2] 4D Research Lab, University of Amsterdam, Amsterdam, the Netherlands ; [3] Faculty of Educational Sciences, Open Universiteit, Heerlen, the Netherlands
r.de.wit@windesheim.nl; h.c.huurdeman@uva.nl

Keywords:

Virtual reality
Immersion
Barrier game
Conversation skills
Special education
Primary education

Received: November 2024
Accepted: May 2025
Published: May 2025
DOI: 10.17083/ijsg.XXXX

Abstract

In Dutch special education, less than half of the students achieve the desired target level for conversational skills. A VR barrier game can contribute to additional practice opportunities: players cannot see each other's actions, making the exchange of information (EoI) necessary to achieve the game’s goal. The purpose of this Educational Design Research study was to design and evaluate a digital barrier game for training conversational skills, in which players practice giving, receiving and clarifying information. The main research question was: "What is the difference between the immersive and the non-immersive version of a VR barrier game in terms of perceived efficiency, user satisfaction, and effectiveness in inducing EoI among special education students aged 9 to 13?” This within-subjects study involved 38 participants. Scores were above average in both the immersive and non-immersive conditions for efficiency (SUS score 72.30 and 69.47, respectively) and satisfaction (IMI score 5.80 and 5.44). The differences were not significant. In terms of effectiveness, both conditions elicited many EoI-oriented utterances (6.33 and 6.47 per minute, respectively), which were analyzed using a smaller sample size (n=26). In this paper, we address the lack of previous studies on training conversational skills and using VR in first language learning. The main takeaway is that immersion was not identified as a key factor influencing engagement in special education students completing a game-based task. Although it was not directly the goal of the study, the results suggest that the evaluated VR barrier game is potentially suitable for creating additional practice opportunities for EoI in special schools.

1. Introduction

Oral language proficiency is crucial for children's development and self-reliance [1]. It is not only essential for functioning in school and society and for building social contacts [2], [3], but oral fluency is also an important predictor of reading ability, as shown by longitudinal research [4]. Moreover, oral and written skills mutually reinforce each other [5].

In the Netherlands, formal targets for oral language proficiency exist for the end of primary school [6]. However, less than half of the students in special education[1] reach the desired target for the conversation component at the end of primary school, according to a recent Dutch research study [1]. An explanation for the differences in performance could lie in the time spent and the number of (targeted) practice opportunities, as frequent sessions make oral language interventions more successful [7] and targeted practice has been shown to be an effective approach [3].

Barrier games may offer the possibility to create more targeted practice opportunities for conversational skills, particularly for the ‘exchange of information’ component. In a barrier game, two players are separated by a physical barrier, preventing them from seeing each other's actions. Verbal information exchange is necessary to achieve the game goal. In speech-language therapy for children with language development disorders, the use of barrier games can enhance information exchange [8], but the individualized guidance provided by a speech therapist is challenging for a classroom teacher to replicate [3]. Integrating instruction and feedback in a digital form of barrier game may require less intensive supervision. This has been demonstrated by Zhang et al. [9] in their study using a collaborative virtual environment (CVE). Pairs of children, with and without autism, worked together to solve puzzles by exchanging information that was available to only one of them. This had a positive effect on information sharing and turn-taking [9]. A disadvantage of this specific CVE is the lack of connection between its design and other school subjects. Integration of language teaching and subject teaching can positively affect both language and subject knowledge [10], [11]. This integration may be feasible via a digital game, which can be more easily adapted to changing subject themes than games utilizing physical materials.

The purpose of the Educational Design Research study [12] described in this paper is to design and evaluate a digital barrier game for training conversational skills, in which players practice giving, receiving and clarifying information. Such a serious game (i.e., a computer game with an educational purpose [13]) may be easily adapted to changing themes and played by children under minimal guidance. This study addresses the lack of research into the use of VR for training conversation skills within special education, while it is precisely for this population that it is difficult to achieve the goals set by governments [1], [14]. This lack of research is addressed by designing and evaluating a digital barrier game for training conversational skills. Within this game, players can practice giving, receiving and clarifying information.

[1] Within the Dutch education system, children with learning problems can attend a special school for primary education (SBO). SBO schools have smaller group sizes, so children receive more guidance. For children who need more support, special education (SO) may be the best option. SO schools are geared to children with disabilities, chronic illnesses or serious behavioral problems [64], [65]. In the Dutch context, there are four types of SO, which cater to different target groups. The cited research [1] was conducted at SBO schools and at SO schools for children with mental disorders and behavioral problems.

1.1 Background

A meta-analysis by Wouters et al. [15] on the effect of serious games (SGs) on learning and motivation suggests that SGs are more effective for learning than conventional instructional methods, especially in the area of language acquisition. They indicate that rich environments have potentially beneficial properties, including graphics that may improve the encoding of meanings and interpretations of words [16], as well as authentic and playful elements. Regarding motivation, Wouters et al. found no significant difference in the effects of SGs. However, several recent studies highlight motivating aspects of the use of SGs. Bjørner et al. reported a positive effect on reading engagement [17], while Aguilar-Cruz and Guayara found that their SG for learning English as a foreign language motivated students in a qualitative case study [18]. Participants in a study by Dinçer and Dinçer emphasized that a SG on aviation was both beneficial and motivating for language learning [19].

Serious games can take place in a digital three-dimensional (3D) environment where players interact using virtual reality (VR). VR refers to computer-simulated environments that can simulate physical presence and participation in real-world places or imaginary worlds [20], [21]. VR applications may differ from each other in the degree of immersion, which refers to two different constructs [22]. A broad interpretation of immersion points to a user's subjective experience while interacting with a virtual environment (e.g. [23]). This is related to concepts such as presence, which is the (subjective) sense of being in a virtual environment [24], and intrinsic motivation [25]. The narrower meaning of immersion concerns the objective technical capability of a system to deliver a surrounding environment [24]. In this context, a distinction is made between *immersive VR*, by means of head-mounted displays and *non-immersive VR*, using e.g. desktop computers [20], [26] (also referred to as "desktop-VR"). The current study further assumes this narrower meaning of immersion.

An advantage of the use of VR is that situations difficult to access in education can still be experienced in a lifelike manner [20], [27]. According to constructivism, such a lifelike experience in an authentic context is important for the transfer of learning [28]. Additional advantages of immersive VR include acquisition of cognitive skills (e.g., understanding visual information), psychomotor skills (e.g., visual scanning) and affective skills (e.g., controlling emotional responses) [21]. A possible disadvantage of VR is that added effects can distract attention from the actual learning content [20]. Cognitive Load Theory (CLT) explains this effect: unnecessary information in the presentation of learning content (extraneous load) reduces the working memory available for the actual learning content [29]. Practical disadvantages of using headsets include technical limitations and motion sickness [21], i.e. short-term nausea, fatigue, dizziness, physical disorientation, and eyestrain. The likelihood of motion sickness is reduced by limiting the time spent within the VR environment and allowing participants to sit while using VR [30].

Research on the effect of degree of immersion on language learning, especially for conversational skills, is limited. Halabi et al. compared different interfaces to improve communication skills. Interviews revealed greater user satisfaction using immersive VR-headsets compared to using non-immersive desktop-VR [31]. In a between-subjects experimental study on reading comprehension, Kaplan-Rakowski and Gruber [32] found higher motivation scores for users of an immersive headset in comparison to those who watched a non-immersive video of the same story. For skills other than language learning, Feng et al. [33] conducted research using identical tasks in immersive (VR-headsets) and non-immersive conditions (desktop-VR). In a wayfinding task, they found no significant differences on the System Usability Scale [34]. A meta-analysis by Chen et al. [26] investigated the effects of VR training on children's language acquisition. Their analysis suggests a significant advantage for VR training, in terms of linguistic gains (e.g., language skills) and affective gains (e.g., learning attitude and motivation), compared to non-VR conditions. The effect on linguistic gains was most significant for non-immersive devices. No significant differences were found between educational levels and language domains. Chen et al. indicate that VR technology holds large potential for both second and first language learning, but that there is a lack of research on the latter. In addition, Parmaxi [27] emphasizes a lack of studies on VR for language learning in primary and secondary education, and a lack of studies using fully immersive VR.

In the field of speech therapy, a review by Bryant et al. describes the potential of using VR to improve communication for people with communication disabilities. Previous research on the effects of non-immersive VR on communication skills has shown positive effects, with initial evidence of transfer to the real world [30]. Another application of VR for training communication skills is practicing interview competencies among police officers. Guimarães et al. [35] found that the assessed game improved the self-perceived competence of its players.

Integration of teaching oral language skills with other language domains and with subject teaching is one of four evidence-based recommendations for interactive oral language teaching formulated by Bruggink and Stoep [2], based on prior research e.g. [10], [36], [37]. Their other three recommendations include: working purposefully on oral language instruction, teaching strategies and providing extensive practice opportunities (e.g. [3], [38], [39]); providing a rich language environment, encouraging interaction, giving feedback and ensuring a safe speaking environment (e.g. [3], [38], [40]); and monitoring and differentiating between students (e.g.[38]).

In summary, while serious games and VR have the potential to create a realistic and motivating environment that invites collaboration and interaction (e.g. [9], [15], [26], [30], [2]), research is limited on using VR for first language learning, especially within primary and secondary education [27]. Furthermore, few studies directly compare immersive and non-immersive VR using identical tasks. No known research exists on the use of VR for training conversational skills in Dutch special education. This Educational Design Research study directly addresses this gap by researching a VR barrier game for practicing interaction, aimed specifically at this target group.

1.2 Research questions

The purpose of this study was to evaluate a newly designed VR barrier game aimed at training the conversational skill "exchange of information" of students aged 9 to 13 years in special education (Figure 1). This is a game which is played in pairs. Players must verbally exchange information to complete the game. In the game, players assume the role of students tasked to help their teacher prepare an exhibition for parents.

Figure 1. Screenshot from informational video about the VR barrier game 'The exhibition'. In this game, two players each have a different view of the same scene. The goal, to place pictures in the correct location in a 3D school environment, is achieved by exchanging information. Involved pictures closely resemble each other (e.g. the “blue ship with containers” depicted in the figure), making accurate descriptions essential.

Exchange of information (EoI)[2] is one of the tasks outlined in the conversation component of oral language skills within the reference framework for Dutch language learning [41]. To this end, the relationship between the degree of immersion and the utility of the VR barrier game to train exchange of information was investigated. Degree of immersion was operationalized using the binary conditions *immersive* (smartphone-VR) and *non-immersive* (laptops) based on the same game. For the immersive condition, the 3D experience is facilitated by inserting a smartphone into a viewer with two lenses (see e.g. [27]). For the purposes of this study, utility was divided into *perceived efficiency*, *effectiveness* and *satisfaction*[3]. Perceived efficiency was operationalized as the score on Brooke's System Usability Scale (SUS) [34]. Effectiveness in inducing exchange of information was operationalized as the frequency scores of three communication variables [9]. Satisfaction was operationalized as the score on the main subscale of the Intrinsic Motivation Inventory (IMI) [42], ‘interest/enjoyment’.

The research question is as follows: "What is the difference between the immersive and the non-immersive version of a VR barrier game in terms of perceived efficiency, user satisfaction, and effectiveness in inducing exchange of information among special education students aged 9 to 13?”

[2] Translated to English, the target level in this task is: “The learner can give and ask for information and listen critically to this information in conversations inside and outside of school. The student can evaluate information and give a response.” [41]
[3] While effectiveness (reaching desired goals), efficiency (accuracy, expenditure of resources) and satisfaction (attitudes towards use) originate from usability studies [66], they are also used for evaluating educational technology; as done in this paper.

It was anticipated that motivation, and consequently user satisfaction, would be higher for the immersive form because of the potential ability of immersion to increase engagement and realism as shown by [23], [31], [32]. We expected effectiveness to be greater with the non-immersive form as the success of VR training depends heavily on context and pedagogical support [20], and a lower degree of immersion in VR-assisted language learning can lead to better outcomes, likely due to reduced cognitive load and fewer physical side effects [26]. For perceived efficiency, no difference was expected because no statistically significant difference on the SUS between immersive and non-immersive forms was found by Feng et al. [33]. These expectations led to the following hypotheses:

H1: The immersive form of the VR barrier game leads to greater user satisfaction than the non-immersive form.

H2: The immersive and non-immersive forms of the VR barrier game do not differ in terms of perceived efficiency.

H3: The non-immersive form of the VR barrier game leads to greater effectiveness in inducing exchange of information than the immersive form.

Because this is a complex problem in educational practice for which no clear guidelines or solutions currently exist, Educational Design Research (EDR) was chosen as the research method [12]. EDR is aimed at increasing theoretical knowledge while simultaneously solving practical educational problems. It typically consists of three phases [43], which may occur during multiple cycles. In phase 1, *analysis and exploration*, the problem is explored. In this study, this phase involved a review of literature and other documentation, leading to an overview of design requirements that a VR barrier game for training exchange of information should meet. During phase 2 of this research, *design and construction*, a draft version of the VR barrier game was developed and tested on a small scale by teaching personnel (6 participants), followed by a focus group session. After modifications, a second test was conducted (4 participants). In addition, the instruments for data collection were finetuned, based on feedback from the teaching personnel. The main focus of the current article is phase 3, *evaluation and reflection*, in which the resulting game was evaluated. Phase 1 and 2 have been described in detail in the Dutch journal Tijdschrift Taal [44].

2. Methods and Material

2.1 Participants

The participants in this study were students in grades 6 to 8 (ages 9 to 13) of a primary school for special education in the southwest of the Netherlands. This school has students with learning and behavioral problems, below average intelligence, serious social-emotional problems and/or psychiatric problems. The students in grade 8 are divided in a practical group (8P) and a theoretical group (8T), i.e. with a higher attainment level. All students in grades 6 through 8 were invited to participate in the study, except those in 8T, since the VR barrier game would probably be too easy for these students. Out of the 62 students approached, 38 participated in the study. Participation was not possible for students known to have motion sickness, due to the risk of motion sickness VR poses.

Participating students were divided into two groups that differed in the order of the study conditions (see Section 2.3). To minimize the influence of student characteristics, the groups were formed through stratified randomization [45] ensuring comparability in terms of grade (6, 7 or 8) and expected exit level (practical education or VMBO), see Table 1. Other characteristics such as autism or behavioral problems were not taken into account. This approach was taken because students with different care profiles did not show a clear difference in results in a study in Dutch special education by the Education Inspectorate [1] concerning the task characteristics of the conversation subdomain. The VR barrier game was played in pairs composed randomly, except for combinations that were organizationally impossible. The pairs could differ in grade and expected exit level. The composition of the pairs was the same in both conditions. Because the same participants were tested two times, a paired t-test was used to compare the results. To achieve a power of .80, a paired t-test required a minimum of 34 participants. This calculation was performed in G*Power [46], with a mean effect size of Cohen's d = 0.5 and α = .05.

Table 1. Composition of groups by grade and expected exit level

Grade	Testgroup 1	Testgroup 2
6	practical education: 3 VMBO: 4	practical education: 3 VMBO: 4
7	practical education: 3 VMBO: 2	practical education: 3 VMBO: 2
8	practical education: 5 VMBO: 3	practical education: 3 VMBO: 3

Note. VMBO (preparatory vocational secondary education) is a four-year programme offering theoretical and practical courses [47]. The VMBO group includes students expected to graduate to one of the levels of VMBO, with or without learning support. Practical education is a form of secondary education which is attended by students who are likely unable to attain a VMBO diploma. They mainly learn practical skills, such as cooking or filling in forms [48].

2.2 Measurement instruments and materials

2.2.1 Questionnaire satisfaction and efficiency

The used questionnaire consisted of the seven items of the interest/enjoyment subscale of the Intrinsic Motivation Inventory (IMI) [42] and the ten items of Brooke's System Usability Scale (SUS) [34]. The items of the IMI and SUS were translated by the researcher and adapted to the language level of the students in consultation with the teachers.

The IMI is based on Ryan and Deci's self-determination theory [25], [42], [49]. The interest/enjoyment subscale measures intrinsic motivation based on self-reporting and is used here as a measure of satisfaction. This subscale (7 items) was found to be sufficiently reliable (using Cronbach’s alpha) in previous studies with α = .78 [50]. Items are rated on a seven-point Likert scale, where 1 represents "not at all true," and 7 represents "completely true."

Perceived efficiency was operationalized with the outcome on the SUS. The SUS provides an overall picture of the subjective usability of a system, regardless of the technology used [34]. The score can range from 0 to 100. The literature uses different thresholds for an acceptable SUS score [51]. The average SUS score is 68 [52], [53], therefore a score of 68 or higher is considered acceptable in this study. The reliability of the SUS (10 items) was found to be high in previous studies with α = .91 [51]. All items are rated on a five-point Likert scale, where 1 represents "totally disagree," and 5 represents "totally agree."

2.2.2 Measurement instruments effectiveness

At the time of the study, no validated measurement tool existed to examine information exchange. Therefore, we selected three (out of nine) communication variables devised by Zhang et al. [9] for evaluating a collaborative virtual environment. These specific variables were selected because they are functions of exchange of information according to Van der Beek et al. [41] (see Table 2 and [54]).

To compare the exchange of information between conditions, the participants' verbal expressions during the game were scored by the researcher using an audio recording, by counting the occurrence of the three communication variables. For each communication variable, the outcome measure was the number of times it occurred divided by number of pure game minutes (i.e. total game time minus in-game instruction time).

To score the inter-rater reliability of the measurement instrument, two recordings (311 utterances) were also scored by a colleague, in the manner described in Measuring instrument 'Exchange of Information' [54]. Using SPSS, Cohen's κ was calculated. According to previous literature, a value of Cohen's κ of .61 or higher counts as substantial [55].

Table 2. Communication variables for coding exchange of information, based on Zhang et al. [9], p. 2784

Communication variable	Description
Frequency of question asking	The number of task-related questions participants ask per minute
Frequency of information sharing – response	How often participants respond to partners per minute
Frequency of information sharing – spontaneous	How often participants spontaneously provide information per minute

2.2.3 VR barrier game 'The Exhibition'

To elicit exchange of information, the VR barrier game 'The Exhibition' was used [56]. This game is played in pairs, where each player has a slightly different view and in which players can solely rely on exchanging verbal information to complete the game. Players assume the role of students tasked to help their teacher prepare an exhibition for parents. Together, the players should hang all the pictures made by students in the right place. The enjoyment of the game lies in collaborating on challenges to achieve a common goal in a familiar environment. See Table 3 for the mechanics of the VR barrier game.

Figure 2. Level 1: "Search the picture": the player must choose the picture described by the fellow player from 8 pictures.

The game starts with a short training task to practice navigating and clicking. In this task, players walk around the virtual school and click on dirty coffee cups to clean them up and make the school tidy. This is followed by three levels (of increasing complexity), each preceded by a brief explanation. Each level consists of 6 tasks; player 1 and player 2 switch roles after each task, so in each level both players get the same practice opportunities. In level 1, players take turns looking for the picture described by the other (Figure 2). In level 2, the pictures must be hung in the right place. Player 1 describes the picture, while player 2 describes its location in the room (Figure 3).

Figure 3. Level 2: “Describe the picture”: Player 1 has to describe the picture and hang it in the right place as told by player 2.

In level 3, students test a photo scavenger hunt for parents. The pictures now have to be described in more detail. Describing the place of a picture in terms of its spatial location is also more difficult, because player 2 sees the room in 3D and player 1 sees the walls of the room in 2D (Figure 4).

Figure 4. Level 3: top image: player 1 sees the walls of the room in 2D ("Tell what you see. Click on the right wall"). Bottom image: player 2 views the room in 3D ("Search for the picture. Tell where it is hanging"). The rectangular blue and red image annotations have been added to the two figures above to indicate the corresponding wall and picture between the two views.

The VR barrier game was designed based on the evidence-based recommendations for interactive oral language teaching formulated by Bruggink and Stoep [2]. The six components of the framework for educational game design by Annetta [23] also served as a guideline. For instance, giving players an *identity* makes them become more invested in the game and *interaction* was facilitated by social communication between players [23]. The exchange of information is considered successful if the players click on the correct answer. In-game feedback that appears upon clicking an answer was formulated according to the guidelines for informative feedback established by Narciss and Huth [57]. The feedback includes knowledge of result or response (KR) and elaborated feedback (EF): each clicked answer is followed by an on-screen message providing feedback on whether the answer is correct or incorrect, along with a hint[4]. A Dutch language method [58] and a Dutch wordfinding workbook [59] were used for formulating the hint. The design of the game is further detailed in [44].

The VR barrier game runs on the CoSpaces[5] platform, which allows for creating educational 3D applications, compatible with desktop computers or laptops, as well as smartphone VR. This enables both non-immersive and immersive experiences [60]. Students must click a button to activate the next task themselves. In case a student pair is no longer at the same point in the game, they can restart at level 1 by clicking a reset button. The maximum duration for the whole game is 15 minutes. When this time finishes, the total number of points, the game duration and the score per minute is displayed. A student manual provides brief instructions on the purpose of the game using written text and pictures, and details how it works (e.g. starting and operating it).

Table 3. Game Mechanics VR Barrier Game

Mechanics	Description
Introduction and training task	The story of the game is introduced by a virtual teacher. Players practice walking around, selecting and clicking. They are also instructed on how to make sure they stay at the same point in the game.
Introduction per level	Each level starts with the virtual teachers explaining and demonstrating the players' objectives.
Level 1 (6 tasks)	A neutral 3D environment with pictures. Player 1 sees one picture and describes it, while player 2 sees all pictures and must find and click on the one that player 1 describes. The pictures are very similar, so it is a challenge to describe the picture so accurately that the other player chooses the correct picture.
Level 2 (6 tasks)	A 3D school environment with picture frames. Some frames display pictures (2D images appropriate to the theme), while other pictures still need to be placed in empty frames. Players must communicate where to place each picture. Player 1 has to describe the picture and hang it in the right place as indicated by the fellow player 2. The pictures are very similar, so it is a challenge to describe the picture in such accurate detail that the other player chooses the correct picture.
Level 3 (6 tasks)	Player 2 sees the same scenario as in level 2, with all the pictures available. Player 1 views a clipboard with a detail image of one of the pictures and, as answer options, the walls of the school (showing only empty picture frames, cabinets and doorways). Player 1 describes the detail image, while player 2 looks for the corresponding picture. Player 2 then

[4] Translated to English, an example of the given feedback is: "Sorry, this is not the right place. Tip: Ask to describe the place more precisely. For example, using below, above, left, right, next to."
[5] Since March 2025, CoSpaces is called Delightex.

Mechanics	Description
	describes on which wall this picture is located, enabling player 1 to click on it.
Clicking an answer	The player clicks on a picture (level 1), a picture frame (level 2) or on an answer option on the clipboard.
Feedback	Upon clicking an answer, the participant receives feedback indicating whether their choice was correct or incorrect. If incorrect, a language method-based tip appears, offering strategies, such as to continue asking questions. A new attempt can be made, with a maximum of three attempts.
Points	The player is allowed 3 attempts for each task. Attempt 1 correct: 3 points awarded Attempt 2 correct: 2 points awarded Attempt 3 correct: 1 point awarded
Score	Points per minute
Reset-button	Players can click a reset button from any level, if they have lost track of each other. The game restarts at Level 1.
End of game	The game ends 15 minutes after starting the game, or as soon as all tasks have been completed. The virtual teachers thank the players for their assistance. Subsequently, the final score screen is displayed, which includes: - Awarded points, as calculated above - Time in seconds (excluding the duration of introduction videos) - Score in points per minute (rounded to two decimals).

Note. Player 1 and player 2 switch roles after each task.

2.3 Procedure

The study has been approved by the ethics committee (cETO) at Open Universiteit in the Netherlands. Students and their parents were contacted through the school and asked to sign a consent form for participation. As described in 2.1, the game is played in (randomly composed) pairs of students. The pairs in this within-subjects study played the barrier game in immersive (using smartphone VR) and non-immersive conditions (using laptops), with an average of 12 days between the two sessions. The conditions were counterbalanced to limit the learning effect. As visualized in Figure 5, student pairs in group 1 experienced the non-immersive condition in the first session. After a break of (on average) 12 days, the same student pair experienced the immersive condition in the second session. For group 2, this sequence was reversed. The pictures in the game were different in the first and second round of play, but chosen to be as similar as possible in the context of the theme of the round (see Figure 6). The within-subjects design was mainly chosen because of the diverse student population in special education. Repeated testing of the same pairs neutralizes this variation as much as possible.

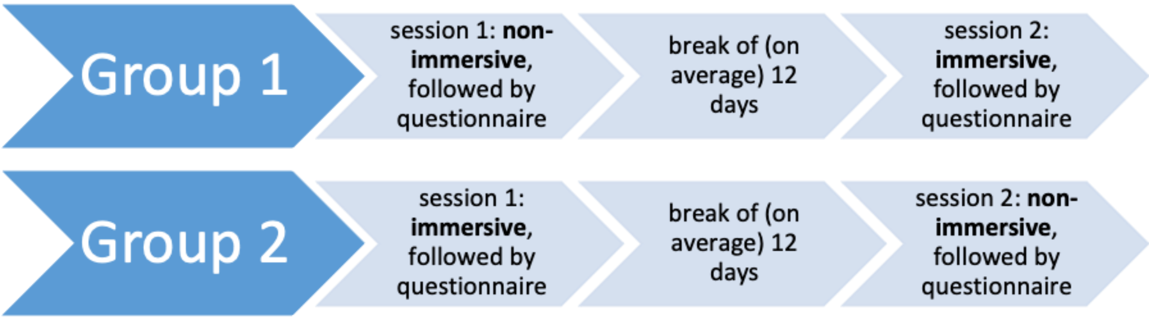

Figure 5. Study sequence for each pair of students playing the barrier game, with counterbalanced order of conditions.

Figure 6. Example of pictures in the VR barrier game round 1 (Water theme, top) and round 2 (land theme, below). *Note:* All pictures in the VR barrier game are from Pixabay and may be used without attribution [61].

To avoid disruptions, the students played the game in a room outside the classroom under the supervision of the researcher. To reduce the chance of socially desirable answers, the students were not told that the game was developed by the researcher.

Each study session lasted about 30 minutes, including playing the VR barrier game (max. 15 minutes). Students received the written manual, and the instructions in the manual were also given verbally by the researcher. Students who regularly used a read-aloud feature in class due to reading difficulties also could use this in the game. Appropriate to the target group, it was sometimes necessary to repeat part of the instruction during the game so students could continue playing. In that case, the researcher only repeated (parts of) the instruction already given in the manual or during the explanation in the game, but did not add any new information. After playing the game, students completed a questionnaire in both sessions (see Section 2.2.1.). The researcher explained the Likert scale to the student and verified that both pages of the questionnaire were completed.

2.4 Data analysis

The data was analyzed in SPSS. The independent variable was the degree of immersion, (conditions: immersive and non-immersive). Dependent variables were satisfaction, perceived efficiency, and effectiveness. Satisfaction was represented by the mean of the scores on the interest/enjoyment subscale of the IMI [42]. Perceived efficiency was represented by the SUS score [34]. Effectiveness in inducing exchange of information was represented by the frequency scores of the communication variables. For each item, the mean was calculated for each condition. A paired t-test (when scores were normally distributed) or a Wilcoxon signed-rank test (when scores were not normally distributed) was used to analyze for each item whether the difference between conditions was significant ($p < .05$). Subsequently, the effect size was calculated for each item (Cohen's d for the paired t test, r for the Wilcoxon signed-rank test).

3. Results

In this section, we report our findings, while in the next section we answer the research question and interpret the meaning of the results. First, we assessed the reliability (Cronbach's alpha) of the measurement instruments. For SUS non-immersive (10 items): $\alpha = .72$ and for SUS immersive (10 items) it was $\alpha = .70$. For IMI, the Cronbach's alpha value for non-immersive (7 items) was $\alpha = .90$ and for IMI immersive (7 items) $\alpha = .87$. Hence, in the context of this study, the reliability of SUS can be considered acceptable ($\alpha \geq .70$) and for IMI good to excellent ($\alpha \geq .80$). For the variable "effectiveness", the inter-rater reliability of scoring communication variables was assessed, there was substantial agreement between the two assessors, $\kappa = .69$, $p < .001$ on 311 utterances.

Next, the summary statistics, effect size and significance of the difference between the immersive and non-immersive versions of the VR barrier games were calculated as shown in Table 4. For the variables "perceived efficiency" and "effectiveness in inducing exchange of information," a paired t-test was conducted. For the variable "satisfaction," the Wilcoxon signed-rank test was used because the scores were not normally distributed[6].

The first variable, perceived efficiency, is represented by the total SUS score (N = 38). The SUS score is higher for the immersive form (M = 72.30) than for the non-immersive form (M = 69.47). This is a trivial effect, and it is not significant.

The second variable, user satisfaction is represented by the total score on the interest/enjoyment subscale of the IMI (N = 38). It is higher for the immersive form (M=5.80) than for the non-immersive form (M=5.44). The effect is small, and it is not significant.

The third variable, effectiveness, is represented by the scores on three communication variables. Anecdotally, no significant differences can be observed between immersive and non-immersive conditions; this concerns a smaller sample size (N=26) due to accidental clicks on the in-game reset button[7].

---

[6] A boxplot was used to examine this variable more closely; there was one outlier. The scores of this participant were checked, in this there were no irregularities.

[7] For this variable, an issue was that some players clicked on the reset level button (ten times this happened accidentally, in six cases it was intentional because the players "lost" each other in the game). Therefore, pure game time (i.e. total game time minus in-game instruction time) was used to calculate the EoI-scores. For six pairs, the pure game time could not be calculated because the game ended during the playback of an instruction, resulting in a usable sample size of N = 26 for the variable effectiveness.

Table 4. Comparison communication variable scores "effectiveness in inducing exchange of information" (EoI).

Variable	*N*	non-immersive	immersive	Effect	Significance
Perceived efficiency (SUS score)	38	*M* = 69.47 *SD* = 16.54	*M* = 72.30 *SD* = 15.15	*d* = -0.17 [-0.49, 0.15]	*p* = .293
User satisfaction (Interest/enjoyment IMI score)	38	*M* = 5.44 *SD* = 1.54	*M* = 5.80 *SD* = 1.37	*z* = 1.12 *r* = 0.13	*p* =.264
Frequency of question asking	26	*M* = 1.49; *SD* = 0.77	*M* = 1.53; *SD* = 0.63	*d* = -0.05 [-0.44, 0.33]	*p* = .790
Frequency of information sharing-response	26	*M* = 1.24; *SD* = 0.62	*M* = 1.24; *SD* = 0.42	*d* = -0.00 [-0.38, 0.39]	*p* = .999
Frequency of information sharing-spontaneous	26	*M* = 3.73; *SD* = 1.24	*M* = 3.56; *SD* = 1.31	*d* = 0.14 [-0.24, 0.53]	*p* = .469
Frequency of EoI focused expressions - total	26	*M* = 6.47; *SD* = 1.75	*M* = 6.33; *SD* = 1.60	*d* = 0.09 [-0.29, 0.46]	*p* = .653

4. Discussion

4.1 Main Results

The research question of this study was "What is the difference between the immersive and the non-immersive form of the VR barrier game in terms of efficiency, user satisfaction and effectiveness in inducing exchange of information among students aged between 9 and 13 years old in special education?" No significant differences were found in this study between the immersive and the non-immersive version of the VR barrier game in terms of perceived efficiency and user satisfaction. While the usable dataset for effectiveness in inducing exchange of information was relatively limited, no difference across conditions could be observed either. Below, we discuss the research results for each of this study's three hypotheses.

4.1.1 Satisfaction

The first hypothesis stated that the immersive form of the VR barrier game leads to greater user satisfaction than the non-immersive form, measured with the interest/enjoyment subscale of the Intrinsic Motivation Inventory (IMI) [42]. A difference was found, but the effect was weak and not significant. Therefore, the hypothesis is rejected.

However, satisfaction was relatively high in both conditions with M = 5.44 in the non-immersive condition and M = 5.80 in the immersive condition, on a scale of 1 (most negative rating) to 7 (most positive rating). After the research study finished, students frequently asked the researcher when they could play the game again. While this is only anecdotal evidence, it does suggest that students enjoyed playing the game.

In this study, we distinguished between immersive and non-immersive VR (the narrow interpretation of immersion), similar to the study by Kaplan-Rakowski and Gruber [32]. In contrast to our results, they did find a difference in the IMI-score between conditions. This difference might be explained due to the task used in their study, as they used a reading comprehension task as opposed to the information exchange task in our study. Also, it might be related to the broader meaning of immersion, in which the subjective sense of presence determines the intrinsic motivation of the player and thus satisfaction [21], [22]. This was not measured in the study by Kaplan-Rakowski and Gruber [32] and our study. Sense of presence is not necessarily related to the used technology, as the study by Feng et al. [33] shows, who found no difference in sense of presence using the same wayfinding task in desktop and headset conditions.

4.1.2 Perceived efficiency

The second hypothesis stated that the immersive and the non-immersive form of the VR barrier game do not differ in terms of perceived efficiency, measured using the score on Brooke's System Usability Scale (SUS) [34]. This hypothesis was confirmed. Although the immersive version of the barrier game scored slightly higher (average SUS score: 72.30) than the non-immersive version on perceived efficiency (69.47), the difference was trivial. Similar to our paper, Feng et al. [33] compared SUS-scores between conditions, and did not find significant differences.

4.1.3 Effectiveness

The third hypothesis stated that the non-immersive form of the VR barrier game leads to a greater effectiveness in inducing exchange of information than the immersive form, i.e. providing and asking for information as well as listening and responding to given information. This was measured using the frequency scores of the communication variables. However, the usable dataset was too small to confirm or reject this hypothesis, since 12 of the 38 scores for effectiveness could not be used because the reset button was clicked during the game. Focusing solely on the usable dataset, no significant differences were found in the frequencies of the communication variables between the immersive and non-immersive condition. Both Halabi et al. [31] and Feng et al. [33] used similar conditions and found significantly greater effectiveness with desktop-VR than with headsets, but these studies focused on different skills than the current study. According to Feng et al., a possible explanation for the difference in effectiveness is the participants' higher familiarity with desktop than headsets. While this was likely also present in the current study, this was not reflected in the effectiveness scores related to the exchange of information.

In general, the total number of information-focused utterances per minute (M = 6.47 in the non-immersive condition and M = 6.33 in the immersive condition, see Table 4) shows that the students were actively exchanging information while playing the game. This suggests that the game can be used to create extra practice opportunities in the classroom.

4.2 Implications

Virtual Reality technology offers great potential for language teaching, but in this context relatively little research has been done [26]. No previous research has been conducted about the use of VR for training conversation skills within special education, while it is precisely for this population that it is difficult to achieve the goals set by governments [1], [14]. Our findings contribute to knowledge on the use of immersive and non-immersive VR for improving conversation skills among students in special education.

This study has various implications. First, the fact that no significant differences were found between modalities suggests that teachers can use existing hardware in their school (e.g. laptops or VR headsets) for creating extra practice opportunities for exchange of information, without the need for acquiring additional hardware. This improves the cost-effectiveness of VR training, which is an important consideration in scalability potential [18], [20]. Keeping in mind the risk of cybersickness which headset-based VR poses, the opportunity of using desktop-based VR also has advantages in terms of accessibility. Second, the active exchange of information between students in special education suggests that VR barrier games, such as ‘The Exhibition,’ could be a promising way to help improving conversational skills. Thus, they may be used for additional practice opportunities for exchange of information in the classroom. These practice opportunities are challenging to organize in a traditional classroom setting, in particular with regards to pinpointed feedback to students. Due to built-in feedback VR barrier games can be played under minimal guidance, also contrary to traditional physical barrier games and games for practicing conversational skills created in the context of previous research studies. This means that these games can be relatively easily integrated into classroom use, although further integration into the learning process and curricula ought to be researched. Third, the studied barrier game involved a limited number of game mechanics, which may have contributed to the engagement of participants (for instance collaboration, feedback and levels [62]). The relation between game mechanics and student enjoyment should be investigated further, and more game mechanics could be integrated in future VR barrier games. Fourth, the literature review conducted for this study underlined that tools to monitor conversation skills are only available to a limited extent [63], making the tracking of the development of students' skills challenging. In this study, a measurement instrument for the frequency of communication variables [54] was created. This instrument can be developed and validated further and serve as a basis for future research into the effectiveness of interventions in inducing exchange of information.

4.3 Limitations and Future Work

The presented research has some limitations. Firstly, the group of participants was relatively diverse (students had learning and behavioral problems, below average intelligence, serious social-emotional problems and/or psychiatric problems), so the results may have been influenced by the differences between the participants. Future research could involve more participants from different schools, so results can be generalized. Future work can also build on the results of the present study by including other user groups than students in special education. For example, students with language development disorders, or students learning Dutch as a second language. It is relatively straightforward to expand the game with other themes, because the pictures in the game can be customized without changing the game structure.

A second limitation was caused by the inadvertent use of the reset button. Therefore, the usable dataset for the effectiveness variable was smaller than envisioned. In addition, the recorded game scores were affected, which could have been an additional measure for effectiveness in initiating exchange of information. It is unclear if there was an influence of the accidental use of the reset-button in 10 of the 38 sessions on the slightly lower SUS-score (relating to efficiency) of the non-immersive condition. In a future version of the game, a feature will be added to cancel a reset in case of accidental use of the reset-button.

A third limitation is that immersion in smartphone-VR may not offer the same experience as dedicated VR hardware (such as an Oculus Quest 3). Although this was a deliberate choice because of costs, it might have affected the results of the study. As Parmaxi [27] indicates, within the context of language learning, more research on fully immersive VR systems is needed.

Finally, although the current research shows that the VR barrier game can be used to create extra practice opportunities for exchange of information in the classroom, it is not known whether transfer occurs to other situations and what the effect will be on students' conversation skills in the longer term [27]. Previous research on the use of VR to improve communication in speech-language problems does provide a positive outlook of transfer to the real world [30]. For instance, the barrier game described in this paper could be replayed at different times in grades 6, 7 and 8, within different thematic contexts by changing the pictures.

5. Conclusions

Listening, speaking and having conversations are important skills for children's development and self-reliance [1], [2], [3], [4], [5], but research in a Dutch context has shown that too few special education students achieve the desired level for the conversation component [1], [6]. An explanation can be found in the limited number of (targeted) practice opportunities and limited amount of time spent practicing these skills [3], [7]. To address these issues, the VR barrier game ‘The Exhibition’ was created, in which players practice giving, receiving and clarifying information. The main research question was: "What is the difference between the immersive and the non-immersive version of a VR barrier game in terms of perceived efficiency, user satisfaction, and effectiveness in inducing exchange of information among special education students aged 9 to 13?” The findings of our Educational Design Research study show that immersion was not a significant factor influencing engagement in special education students completing a game-based task, in terms of perceived efficiency and user satisfaction.

Although it was not directly the goal of the study, we observed students actively exchanging information while playing the serious game in both conditions. This suggests that the VR barrier game can be used by teachers to create extra practice opportunities for exchange of information in special education, with minimized guidance and the opportunity to adapt the game to different subject themes. Since no significant differences were found between the immersive and the non-immersive form of the VR barrier game, teachers can use existing hardware already present at school (e.g. laptops or VR headsets) without a need for additional expenses.

There has been limited research on training conversational skills and using virtual reality in first language learning. This study provides insights into the potential use of the evaluated VR barrier game for additional practice opportunities for exchange of information in special education. To fully understand its impact, further research is recommended to explore the long-term effects on learning outcomes. For this purpose, the measurement instrument for the frequency of communication variables created for this study can be further developed and validated.

Acknowledgments

We would like to thank the school staff and students for their cooperation in the survey. Also, we would like to thank Rob Nadolski (Open Universiteit), for his valuable feedback on the research methodology. The research this article is based on was conducted in the context of the Master’s program Educational Sciences at the Open Universiteit.

Conflicts of interest

No conflicts of interest are reported.

References

[1] Inspectie van het Onderwijs, “Peil.Mondelinge taalvaardigheid einde speciaal (basis)onderwijs 2017-2018.” [Online]. Available: https://www.onderwijsinspectie.nl/onderwerpen/peil-onderwijs/documenten/rapporten/2019/11/29/peil.mondelinge-taalvaardigheid-einde-speciaal-basisonderwijs

[2] M. Bruggink and J. Stoep, “Mondeling taalonderwijs: Vier aanbevelingen voor de leraar,” Onderwijskennis.nl (NRO). [Online]. Available: https://www.onderwijskennis.nl/artikelen/mondeling-taalonderwijs-vier-aanbevelingen-voor-de-

leraar
[3] E. B. Hadley, K. M. Newman, and J. Mock, "Setting the Stage for TALK: Strategies for Encouraging Language-Building Conversations," *Read. Teach.*, vol. 74, no. 1, pp. 39–48, Jul. 2020, doi: 10.1002/TRTR.1900.
[4] S. Suggate, E. Schaughency, H. McAnally, and E. Reese, "From infancy to adolescence: The longitudinal links between vocabulary, early literacy skills, oral narrative, and reading comprehension," *Cogn. Dev.*, vol. 47, pp. 82–95, Jul. 2018, doi: 10.1016/J.COGDEV.2018.04.005.
[5] D. K. Dickinson, K. T. Nesbitt, and K. G. Hofer, "Effects of language on initial reading: Direct and indirect associations between code and language from preschool to first grade," *Early Child. Res. Q.*, vol. 49, pp. 122–137, Oct. 2019, doi: 10.1016/J.ECRESQ.2019.04.005.
[6] Rijksoverheid, "Referentieniveaus taal en rekenen." Accessed: Jun. 16, 2020. [Online]. Available: https://www.rijksoverheid.nl/onderwerpen/taal-en-rekenen/referentiekader-taal-en-rekenen
[7] Education Endowment Foundation, "Oral language interventions." Accessed: May 25, 2024. [Online]. Available: https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/oral-language-interventions
[8] A. Bekkenutte, F. Van Tilburg, and A. Scheper, "Doeltreffendheid interventie 'Communicatie via Scherm': De verbetering van metalinguïstische en taalvaardigheden van kinderen met een TOS," *Stem-, Spraak- en Taalpathologie*, vol. 23, pp. 9–38, Sep. 2018, doi: 10.21827/5bac8dee9fd1b.
[9] L. Zhang, Z. Warren, A. Swanson, A. Weitlauf, and N. Sarkar, "Understanding performance and verbal-communication of children with ASD in a collaborative virtual environment," *J. Autism Dev. Disord.*, vol. 48, no. 8, pp. 2779–2789, Aug. 2018, doi: 10.1007/S10803-018-3544-7.
[10] L. U. Bradbury, "Linking science and language arts: A review of the literature which compares integrated versus non-integrated approaches," *J. Sci. Teacher Educ.*, vol. 25, no. 4, pp. 465–488, Oct. 2014, doi: 10.1007/S10972-013-9368-6.
[11] A. Wilschut and M. Pijls, *Effecten van vakkenintegratie: Een literatuurstudie*. Kenniscentrum Onderwijs en Opvoeding; Hogeschool van Amsterdam, 2018. Accessed: Jan. 06, 2023. [Online]. Available: https://hdl.handle.net/20.500.11884/7f01c39d-5d80-4d84-8c60-e6cd1081900e
[12] T. Plomp, "Educational design research: An introduction," in *An introduction to educational design research*, 3rd ed., T. Plomp and N. Nieveen, Eds., Enschede: SLO, 2007, ch. Educationa, pp. 9–35.
[13] K. Becker, "What's the difference between gamification, serious games, educational games, and game-based learning?," *Acad. Lett.*, no. 209, pp. 1–4, Jan. 2021, doi: 10.20935/AL209.
[14] M. Langberg, S. Jansma, C. Fortgens, and V. Van der Schie, *Passende perspectieven taal: Profielschetsen*. Enschede, Nederland: SLO, 2012.
[15] P. Wouters, C. van Nimwegen, H. van Oostendorp, and E. D. van Der Spek, "A meta-analysis of the cognitive and motivational effects of serious games," *J. Educ. Psychol.*, vol. 105, no. 2, pp. 249–265, 2013, doi: 10.1037/A0031311.
[16] J. M. Clark and A. Paivio, "Dual coding theory and education," *Educ. Psychol. Rev.*, vol. 3, no. 3, pp. 149–210, 1991, doi: 10.1007/BF01320076.
[17] T. Bjørner, M. S. Petersen, G. S. Jakobsen, D. B. Hendriksen, and N. L. S. Hansen, "How can a foundation be outlined for a successful serious game to increase reading engagement," *Int. J. Serious Games*, vol. 10, no. 1, pp. 81–95, Apr. 2023, doi: 10.17083/IJSG.V10I1.578.
[18] P. J. Aguilar-Cruz and H. A. Álvarez Guayara, "A Serious Game to learn English: The case of Bethe1Challenge," *Int. J. Serious Games*, vol. 8, no. 4, pp. 65–80, Dec. 2021, doi: 10.17083/IJSG.V8I4.448.
[19] N. Dinçer and R. Dinçer, "The effect of a serious game on aviation vocabulary acquisition," *Int. J. Serious Games*, vol. 8, no. 4, pp. 49–63, Dec. 2021, doi: 10.17083/IJSG.V8I4.464.
[20] D. F. Angel-Urdinola, C. Castillo-Castro, and A. Hoyos, "Meta-analysis assessing the effects of virtual reality training on student learning and skills development," *Policy Res. Work. Pap.*, vol. 9587, pp. 1–50, Mar. 2021, doi: 10.1596/1813-9450-9587.
[21] L. Jensen and F. Konradsen, "A review of the use of virtual reality head-mounted displays in education and training," *Educ. Inf. Technol.*, vol. 23, no. 4, pp. 1515–1529, Jul. 2018, doi: 10.1007/s10639-017-9676-0.
[22] P. A. Barclay and C. Bowers, "Associations of subjective immersion, immersion subfactors, and learning outcomes in the revised game engagement model," *Int. J. Game-Based Learn.*, vol. 8, no. 1, pp. 41–51, 2018, doi: 10.4018/IJGBL.2018010103.
[23] L. A. Annetta, "The 'I's' have it: A framework for serious educational game design," *Rev. Gen. Psychol.*, vol. 14, no. 2, pp. 105–112, 2010, doi: 10.1037/a0018985.
[24] M. V. Sanchez-Vives and M. Slater, "From presence to consciousness through virtual reality," *Nat. Rev. Neurosci.*, vol. 6, no. 4, pp. 332–339, 2005, doi: https://doi.org/10.1038/nrn1651.
[25] R. M. Ryan and E. L. Deci, "Self-Determination Theory and the facilitation of intrinsic motivation, social development, and well-being," *Am. Psychol.*, vol. 55, no. 1, pp. 68–78, 2000, doi: 10.1037110003-066X.55.1.68.

[26] B. Chen, Y. Wang, and L. Wang, "The effects of virtual reality-assisted language learning: A meta-analysis," *Sustainability*, vol. 14, no. 3147, p. 3147, Mar. 2022, doi: 10.3390/SU14063147.
[27] A. Parmaxi, "Virtual reality in language learning: a systematic review and implications for research and practice," *Interact. Learn. Environ.*, vol. 31, no. 1, pp. 172–184, 2023, doi: 10.1080/10494820.2020.1765392.
[28] P. A. Ertmer and T. J. Newby, "Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective," *Perform. Improv. Q.*, vol. 26, no. 2, pp. 43–71, 2013, doi: 10.1002/PIQ.21143.
[29] F. Paas, A. Renkl, and J. Sweller, "Cognitive load theory and instructional design: Recent developments," *Educ. Psychol.*, vol. 38, no. 1, pp. 1–4, 2003, doi: 10.1207/S15326985EP3801_1.
[30] L. Bryant, M. Brunner, and B. Hemsley, "A review of virtual reality technologies in the field of communication disability: Implications for practice and research," *Disabil. Rehabil. Assist. Technol.*, vol. 15, no. 4, pp. 365–372, May 2020, doi: 10.1080/17483107.2018.1549276.
[31] O. Halabi, S. El-Seoud, … J. A.-I. J., and U. 2017, "Design of immersive virtual reality system to improve communication skills in individuals with autism.," *Int. J. Interact. Mob. Technol.*, 2017, doi: 10.3991/ijim.v11i2.6555.
[32] R. Kaplan-Rakowski and A. Gruber, "An experimental study on reading in high-immersion virtual reality," *Br. J. Educ. Technol.*, vol. 55, no. 2, pp. 541–559, Mar. 2024, doi: 10.1111/BJET.13392.
[33] Y. Feng, D. C. Duives, and S. P. Hoogendoorn, "Wayfinding behaviour in a multi-level building: A comparative study of HMD VR and desktop VR," *Adv. Eng. Informatics*, vol. 51, p. 101475, 2022, doi: 10.1016/j.aei.2021.101475.
[34] J. Brooke, "SUS: A 'quick and dirty' usability scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and A. L. McClelland, Eds., Londen: Taylor and Francis, 1996.
[35] M. Guimarães *et al.*, "ISPO: A serious game to train the interview skills of police officers," *Int. J. Serious Games*, vol. 9, no. 4, pp. 43–61, Nov. 2022, doi: 10.17083/IJSG.V9I4.514.
[36] M. B. Evans, "The Integration of Reading, Writing, Speaking, and Listening Skills in the Middle School Social Studies Classroom," *All Grad. Theses Diss. Spring 1920 to Summer 2023*, p. 7157, Aug. 2018, doi: https://doi.org/10.26076/6e59-0067.
[37] Education Endowment Foundation, "Improving secondary science. Guidance report." Accessed: Jun. 11, 2024. [Online]. Available: https://educationendowmentfoundation.org.uk/guidance-for-teachers/science
[38] A. Wurth, D. Tigelaar, H. Hulshof, J. De Jong, and W. Admiraal, "Key elements of L1-oral language teaching and learning in secondary education. A literature review.," *L1-Educational Stud. Lang. Lit.*, vol. 19, pp. 1–23, 2019, doi: 10.17239/L1ESLL-2019.19.01.15.
[39] M. Van der Veen, "Dialogic classroom talk in early childhood education," Dissertation, Vrije Universiteit Amsterdam, 2017. [Online]. Available: https://hdl.handle.net/1871/55445
[40] A. Willemsen, M. N. Gosen, T. Koole, and K. de Glopper, "Teachers' pass-on practices in whole-class discussions: how teachers return the floor to their students," *Classr. Discourse*, vol. 11, no. 4, pp. 297–315, Oct. 2020, doi: 10.1080/19463014.2019.1585890.
[41] A. Van der Beek, M. Hoogeveen, and J. Prenger, *Leerstoflijnen mondelinge taalvaardigheid beschreven: Uitwerking van het referentiekader Nederlandse taal voor het onderwijs in mondelinge taalvaardigheid op de basisschool*. Enschede: SLO, 2015. Accessed: Oct. 08, 2022. [Online]. Available: https://www.slo.nl/@4312/leerstoflijnen-0/
[42] Center for Self-Determination Theory, "Intrinsic Motivation Inventory (IMI)." Accessed: Mar. 03, 2023. [Online]. Available: https://selfdeterminationtheory.org/intrinsic-motivation-inventory/
[43] S. McKenney and T. C. Reeves, *Conducting educational design research*, 2nd ed. Routledge, 2019.
[44] R. de Wit, "Gesprekken trainen in virtual reality," *Tijdschr. Taal*, vol. 24, pp. 4–8, 2024, Accessed: Apr. 09, 2024. [Online]. Available: https://www.tijdschrift-taal.nl/art/50-7949_Gesprekken-trainen-in-virtual-reality
[45] J. W. Creswell, *Educational research: Planning, conducting and evaluating quantitative and qualitative research*, 4th ed. Pearson Education Limited, 2014.
[46] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner, "G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences," *Behav. Res. Methods*, vol. 39, pp. 175–191, 2007, doi: 10.3758/BF03193146.
[47] Government of the Netherlands, "Pre-vocational secondary education (VMBO)." Accessed: Apr. 09, 2024. [Online]. Available: https://www.government.nl/topics/secondary-education/different-types-of-secondary-education/pre-vocational-secondary-education-vmbo
[48] Government of the Netherlands, "Obtaining approval for practical education for your child." Accessed: Apr. 09, 2024. [Online]. Available: https://www.government.nl/topics/appropriate-education/approval-for-practical-education-for-your-child
[49] E. L. Deci and R. M. Ryan, *Intrinsic motivation and self- determination in human behavior*. New

York: Plenum, 1985.
[50] E. McAuley, T. Duncan, and V. V. Tammen, “Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis.,” *Res. Q. Exerc. Sport*, vol. 60, no. 1, pp. 48–58, 1989, doi: 10.1080/02701367.1989.10607413.
[51] A. Bangor, P. T. Kortum, and J. T. Miller, “An empirical evaluation of the System Usability Scale,” *Int. J. Hum. Comput. Interact.*, vol. 24, no. 6, pp. 1–44, 2008, doi: 10.1080/10447310802205776.
[52] J. Brooke, “SUS: A Retrospective,” *J. Usability Stud.*, vol. 8, no. 2, pp. 29–40, 2013, Accessed: Feb. 13, 2023. [Online]. Available: https://www.researchgate.net/publication/285811057
[53] J. Sauro, “5 Ways to interpret a SUS score,” MeasuringU. [Online]. Available: https://measuringu.com/interpret-sus-score/
[54] R. de Wit and H. C. Huurdeman, “Measuring instrument ‘Exchange of Information,’” 2025, *Zenodo*. doi: 10.5281/ZENODO.15402912.
[55] J. R. Landis and G. G. Koch, “The measurement of observer agreement for categorical data,” *Biometrics*, vol. 33, no. 1, pp. 159–174, Mar. 1977, doi: 10.2307/2529310.
[56] R. de Wit, “VR-schermspel.” Accessed: Apr. 09, 2024. [Online]. Available: https://taaldeur.nl/vr-schermspel/
[57] S. Narciss and K. Huth, “How to design informative tutoring feedback for multi-media learning,” in *Instructional design for multimedia learning*, H. M. Niegemann, D. Leutner, and R. Brünken, Eds., Münster, New York: Waxmann, 2004, pp. 181–195.
[58] H. Van de Ven and M. Vreman, *Taal op Maat*. Noordhoff, 2012.
[59] Koninklijke Kentalis, Ed., *Woordvinding. Werkboek logopedie.* Koninklijke Kentalis. [Online]. Available: https://shop.kentalis.nl/products/cm-16-047
[60] CoSpaces Edu, “CoSpaces Edu for kid-friendly 3D creation and coding.” Accessed: Jan. 21, 2023. [Online]. Available: https://www.cospaces.io/edu/
[61] Pixabay, “Pixabay terms of service,” Apr. 04, 2023. Accessed: Jul. 16, 2023. [Online]. Available: https://pixabay.com/nl/service/terms/
[62] S. Arnab *et al.*, “Mapping learning and game mechanics for serious games analysis,” *Br. J. Educ. Technol.*, vol. 46, no. 2, pp. 391–411, Mar. 2015, doi: 10.1111/bjet.12113.
[63] E. De Bree, “Taal in uitvoering: Passend onderwijs voor taal en TOS,” 2022. [Online]. Available: https://vhz-online.nl/oratie-prof-dr-elise-de-bree
[64] Government of the Netherlands, “Primary education.” Accessed: May 26, 2024. [Online]. Available: https://www.government.nl/topics/primary-education
[65] Government of the Netherlands, “Types of primary school.” Accessed: May 26, 2024. [Online]. Available: https://www.government.nl/topics/primary-education/types-of-primary-school
[66] ISO, “International standard ISO 9241-11:2018,” 2018. Accessed: Apr. 06, 2025. [Online]. Available: https://www.iso.org/standard/63500.html