




Language teacher education through engagement in an educational escape room

Joy Egbert, Washington State University 

Seyed Abdollah Shahrokni, Texas A&M University 

Abstract

To fill a number of gaps in both the task engagement and language education literature, 25 international and multilingual graduate teacher education students in groups of 3 and 4 participated in a stand-alone, face-to-face educational escape room (EER). The purpose of the experience was to model facilitators of task engagement for in-service English teachers while supporting them in engaging in learning about and practicing comma use. This exploratory study employed pre- and post-test statistics of learning outcomes to complement descriptive data. Results support assertions in the literature that EERs can be engaging across learners and language content. Guidelines and implications for future EER development are provided, including ideas for how the next iteration of the EER might be designed more effectively and how additional rooms might be designed for language teacher education.

Keywords: Teacher Education, Language, Engagement, Escape Room

Language(s) Learned in This Study: English

APA Citation: Egbert, J., & Shahrokni, S. A. (2026). Language teacher education through engagement in an educational escape room. *Second Language Teaching & Practice*, 6(1), 106–122.

<https://hdl.handle.net/10125/69898>

Introduction

The use of educational escape rooms (EERs) and escape games is becoming ever more popular across educational settings (Manojlovic, 2022; Veldkamp et al., 2021). While the empirical literature on EERs to date is somewhat limited, much anecdotal evidence suggests that EER use might help to solve several issues experienced in graduate language teacher education programs, such as: 1) graduate level teachers who are not effectively engaged in their heavily theoretical and research-based classes; 2) graduate students without undergraduate language teaching degrees who may not have much knowledge about engaging their students in language, and; 3) language education coursework that may not provide enough models of engaging tasks to support these future educators. Further, many graduate programs do not include surface-level writing support, which students may need. Therefore, this paper examines the development and use of a stand-alone physical EER that was built to model engaging language teaching while also engaging participants in content learning about commas.

The purpose of this article is to describe the design process and development of the EER and to present participant outcomes for both task engagement and content learning. Through the use of a design-based research (DBR) approach, the article also provides guidelines and implications for both teaching and future research. To meet this purpose, a literature review presents the definition of EER, current research on EERs, and an explanation of the task engagement facilitators that are used as a theoretical foundation for the study. The methodology is described next, then narrative that describes the study process and outcomes is provided. The article concludes with guidelines and implications. This study is significant in several ways:

1) it is one of the few studies of the use of a complete face-to-face EER in language teacher education; 2) it is one of the few studies with a single language skill as the focus of an EER; 3) it is one of the few EER studies that addresses graduate language teacher education; 4) it employs a theoretical framework that can be used across studies of EERs and language learning.

Literature Review

This literature review provides EER research and presents the task engagement facilitators that serve as the theoretical foundation of this study.

EERs in Education

An educational escape room (EER) is a team-based experience in which participants interact with disciplinary content and/or skills to solve different kinds of puzzle tasks. The tasks are typically linear (meaning one leads to another) and have a specific learning goal that users must meet within a time limit, usually 45 minutes to one hour. Tasks and clues are often hidden in boxes with numeric or alphabetic locks on them. Ouariachi and Wim (2020) note that EERs are “experiential and immersive” (p. 1196), require problem-solving, critical thinking, and collaboration, and provide students with a “sense of urgency” (p. 1196) that helps them to complete the tasks.

Teachers can integrate escape games in classrooms in many ways, such as using simple puzzle boxes (Veldkamp et al., 2020), digital “breakouts” from resources such as Breakout EDU (breakoutedu.com), or escapes for small groups or team competition (Queirguga-Dios et al., 2020). Less common, but perhaps more powerful, are EERs that take place in a physical room specifically designated for the escape, with participants interacting face-to-face. Veldkamp et al., 2020) advise EER developers to “aim for a high success rate” (p.3) and try to make sure that the participants achieve the educational goals.

Researchers agree that there is a lack of studies on aspects of EER use, from the affordances to the outcomes (Grande-de-Prado et al., 2020; Taraldsen et al., 2022). EER research focuses increasingly on diverse disciplines (see, e.g., Horat & Grillenberger, 2021; Lopez-Pernas et al., 2019; Rowe, 2020; Santamaria & Alcalde, 2019; Veldkamp et al., 2020; Wargo & Garcia, 2021). The majority of these studies focus on higher education, are commonly conducted with undergraduates and rarely address the social sciences. More specifically, few studies focus on teacher education (see, e.g., Weisberg et al. 2022, for one example of pre-service teachers creating digital breakout games and Voreopoulou et al. 2024, for a study with in-service teachers).

Recent literature on the use of EERs for language learning includes several studies. For example, Cotoc and Mudure-Iacob (2024) created and explored tasks that supported active collaboration among Romanian English as a foreign language (EFL) students from different disciplines around the use of academic vocabulary. The teacher-researchers focused on supporting 15 MA-level students’ cognitive and social skills and creative use of language and content. Using a self-report survey as data, the study found that participants considered the content of the EER relevant and the collaboration useful and fun, but the study did not measure student outcomes. In another study of EERs and language education, Nasri et al. (2024) explored 212 English language undergraduates’ perceptions of a digital escape room focused on reading. In an anonymous survey, respondents stated that they thought the escape was “good” and that they had improved their reading skills. However, there was no other measure to support this perception. One study that did measure student language outcomes was Saban et al. (2024), who examined 55 4th grade EFL learners’ experiences learning graphemes in a set of small digital escapes. As in Cotoc and Mudure-Iacob (2024) and Nasri et al. (2024), their experiment found that the students reported that they were engaged by the game-like tasks and that the intervention group significantly improved their spelling. These and other recent studies (e.g., Mudure-Iacob, 2024; Rad & Alipour, 2024; Voreopoulou et al., 2024) provide evidence that language students and teachers generally perceive escape rooms as useful and engaging. However, all of these studies focus on digital rather than face-to-face EERs, and none explore graduate language teachers’ experiences in the room. In addition, Saban et al. (2024) is one of the very few studies that measure student learning outcomes from EER use in any way.

Benefits for Students and Teachers

The extant research claims many participants benefit from EER use. For example, Grande-de-Prado et al. (2020) and Veldkamp et al. (2020) found that EER participation can help students develop creativity. A number of studies have provided evidence that EER participation can also provide students with a sense of ownership and autonomy over their learning (see, e.g., a systematic review by Veldkamp et al., 2020). Furthermore, studies by Nicholson (2015), Clarke et al. (2017), and Martens and Crawford (2019) claim that EER use can support the involvement of students who might otherwise not participate in class tasks due to shyness or reticence. In the same vein, studies by Lopez-Pernas et al. (2019) and Nicholson (2015) suggest that EER participation is appealing to and can support all genders equally. Similarly, Dilek and Dilek (2018) state that EER use can deeply involve students for whom classrooms are not the most effective learning spaces. Nowacki and Stasiak (2023) found EERs to be “extraordinary” for many students based on their typical instructional experiences. In addition, Manojlovic (2022) shows that EERs can play a role in cognitive development, including “allowing mistakes, providing instant feedback, and developing creativity...” (p. 178).

There are additional benefits to EER use. Ouariachi and Wim (2020), in their exploratory study using qualitative content analysis, note that EERs can support students in “learning from failure, managing risk-taking...and increasing the feeling of enjoyment from learning” (p. 1194). In another study, Mathieson and Duca (2021) asked community-based science EER participants how they felt about the room experience and whether they learned. They found that the EER game element motivated participants to try, while the theme of the room was of less interest to them. In other words, they determined that those with less interest in the topic became engaged in it in the EER and that they felt they had both learned and believed that the EER was a good way to learn.

In addition to benefits for students, Martens and Crawford (2019) suggest that EER use can benefit teachers by helping them think about “instructional decisions that teachers need to make in order to cultivate wonder” (p. 72). In their anecdotal study with teachers, the authors found additional benefits, including:

- identifying opportunities to theorize and make predictions
- taking risks
- responding to students in real ways
- modeling wonder/curiosity
- personalizing and challenging students’ thinking
- exploring content in new ways
- creating “freeing” learning experiences

On the other hand, Grande-de-Prado et al. (2020) included challenges for teachers and researchers in using and exploring EERs. They noted the ineffective evaluation of EER outcomes, much time spent developing and testing the EER, limited resources in some schools, and a lack of balance between participant skills and the challenge of the task. Furthermore, Kanakambaran (2021) added that students can become exhausted doing EER tasks. However, the majority of studies on EER use have not addressed learning outcomes.

EER Design Issues

Veldkamp et al. (2020) conducted a systematic review of the broader literature on EERs. In addition to standard educational task design principles, they note that students prefer not to have staff in the EER and that teachers should not offer hints or help unless they are asked for it. Conversely, the reviewed studies also showed frustration on the part of students when no help was offered. Another design issue that the authors stressed was the importance of debriefing, noting that 85% of the students across EER studies agreed that debriefing helped them to understand the content. They also claim that EER design should not include break times during the experience, because not only do they impinge on the time that participants have, but they may also cause students to lose their train of thought and feeling of immersion. Veldkamp et al. (2020) conclude that assigning grades is probably unnecessary for EERs because they are not needed as

motivators for this type of task. Overall, the current research on EERs suggests that there are compelling reasons to design, use, and study them. It recommends the use of design principles such as a challenge/skills balance, clear focus and goals, solvable puzzles, and opportunities for support and feedback.

Task Engagement (TE)

These design principles and other design concepts, such as novelty and meaningfulness (Queiruga-Dios et al., 2020), are facilitators of task engagement (TE; Reschly & Christenson, 2022). Employing TE principles in both the design and research of EERs can provide “a ‘robust’ foundation” (Ouariachi & Wim, 2020, p. 1194) that is currently lacking. A TE focus is central in language education studies because the literature provides evidence that when teachers can support language students’ intense concentration and absorption in tasks (in other words, task engagement), they also facilitate their students’ achievement (Csikszentmihalyi, 2014). Almost all of the studies conducted to date that address EERs focus on student engagement with the task or on one or more facilitators of task engagement (listed below). Every study finds that students are engaged during EER participation, and that means that EERs can be effective tasks for supporting learner achievement.

Specific, salient TE facilitators can be synthesized both from past research and from numerous more recent studies. These facilitators are:

- authenticity, or a task designed with clear, relevant goals and value for students (MacDonald et al., 2000);
- interest to learners (Reschly & Christenson, 2013);
- social interaction (Vygotsky, 1978);
- a challenge/skills balance (Csikszentmihalyi, 2014);
- an autonomy/structure balance (Authors, 2019); and
- “timely performance feedback” (Shernoff et al., 2014, p. 216) and other learning support.

As noted previously, in the EER research, participants’ responses have reflected many of the task engagement facilitators; however, these facilitators are often not specifically named or explored.

While there may be some challenges to designing effective EERs, overall, the benefits to learning and engagement appear to be worth the attempt. Mathieson and Duca (2021) note that, “More research is needed to validate escape rooms as an engagement format. Similarly, evidence-based research needs to be conducted to understand what features of an escape room might allow for greater engagement, and how they contribute to this” (p. 354). This paper starts to close some of the gaps in the EER/language education literature.

The Current Study

Research Questions

Based on the gaps in the EER research literature and on the TE facilitators noted above, the research questions for this study were:

1. What are graduate language teacher education students’ perceptions of their TE during an EER experience?
2. What are the learning outcomes from their participation in a language-focused EER?

As there is little literature around the complex, integrated set of theoretical components that this study will explore, and on EER use in graduate teacher education, the outcomes of this project will make a contribution to the literatures on task engagement, EER use, and graduate teacher education.

Methodology

This study employed a qualitative methodological framework supported by numeric data and a design-based research approach ([DBR]; Brown, 1992; Collins, 1992). The study uses DBR to explore the complex, real-life relationships between theory, EER design, and EER use. The study took place at a large public university in the U.S. Pacific Northwest.

Participants

The EER research/design team included: the project PI, who was a teacher education faculty member at the university; an adjunct faculty member who, along with being involved in other aspects of the room, focused specifically on the technology used; 2 PhD students who collected data and supported the room design, and; two Master's level students and an hourly worker – all former elementary school teachers in other countries – who worked as room and puzzle designers. All research team members had previously participated in commercial escape rooms. One PhD student was collecting data from the EER for her dissertation research. At some time, each member of the team introduced the room, conducted the pre- and/or post-debriefing, and/or gave hints to the participants in the EER.

After IRB approval, recruitment for participants was through an email to College of Education faculty to share with their students, word of mouth among students (the most effective), and posters in two hallways of the building in which the College of Education and EER were housed. The overall number of participants in this study was 25 (10 male and 15 female) graduate students; however, not all of the 25 participants completed each phase of data collection (25 took the pre-participation survey; 15 took the post-participation survey; 18 took the content pre/post quizzes; and 12 participated in an interview). The participants, 18 of whom were international students, studied in one of three teacher education programs in the College of Education and ranged in age from 22 to over 35. In terms of prior experience, 21 participants had never experienced an escape room before, while 4 participants had participated in a commercial escape room once. The participants completed the experience in 8 groups of 3 or 4, usually as a team that signed up together, but sometimes with unfamiliar team members who had signed up individually through the website.

Context and Process

In DBR, each step may be reiterated. The basic iterative steps before the final implementation of the EER, which generally follow Clarke et al.'s (2017) and others' recommendations for EER design, were:

1. Observe and document participant behavior.
From the research team's experience and with agreement from other faculty, the decision was made to focus on a simple but pervasive need throughout the college's teacher education programs: for students to improve their comma use. Students typically did not study this elsewhere in the education program, and there was general agreement that our teacher education students do not use commas well and so have problems writing and providing good models for their students.
2. Create goals and a narrative with instructors and designers.
The research team reviewed the small body of empirical literature on EERs available in 2019 and brainstormed a process for EER development, the roles of team members, and general puzzle topics.
3. Determine EER tasks and processes.
Participants were expected not just to complete paper quizzes but to have opportunities for critical thinking and to engage in what might be new and different ways than in their courses. Content development of the puzzles included everyone on the research team, exchanging puzzles, testing, and commenting on them. Pre- and post- data collection tasks such as surveys and quizzes were also developed through team collaboration.
4. Design, layout, and test the room.
The college designated a temporary physical space. The EER design team chose a 45-minute time limit so that the participants could also participate in a post-experience interview and potentially finish the whole experience in an hour. As shown in the lower left of [Figure 1](#), a large digital clock

was included in the EER space to help participants easily track their remaining time. Locks and boxes came from diverse sources, including team members' homes and offices, a subscription to breakout.edu, and Walmart and \$1 Store. The rest of the materials were gathered through donations or with funding support from the department/college.



Figure 1. *The EER Setting*

The team developed prototype puzzles and layouts in a sequential structure (Grande-de-Prado et al., 2020). Before the puzzles and the order in which they would be presented were finalized, they were tested several times in preliminary and beta tests with volunteer participants. Because the EER was created in the hope that a wide range of participants could use it, the initial puzzles and room set up were beta-tested by international students, 4th graders, program faculty, and even staff. The “room” setup that made most sense to the research team was a generic school classroom (shown in [Figure 1](#)), first because most of the materials and resources could be obtained in the college, and second, because it was something that all of the participants would recognize that could be used for a number of different EER themes. The final design was based on engagement facilitators, as noted in [Table 1](#), and everything that the participants needed could be found in the room.

5. Conduct experience and debrief/collect and analyze data to redesign as needed. As beta test data came in, including interviews and observations of the room while users were solving the puzzles, the team made changes including correcting mistakes in puzzle outcomes, clarifying puzzle content, changing the order of the puzzles (from simplest to most difficult), and improving the technologies used to record and implement the room. Depending on the outcomes of each beta round of EER use, any part of the room or methodology could be altered before the following round. During the final round of data collection reported here, the EER design team did not implement additional changes.
6. Save digital versions of the room plans and publish research findings.

Research findings are reported below.

Table 1. *Example Integration of TE Facilitators*

Facilitator	Design Feature
Authenticity	Participants needed the focal skill for their graduate writing and language teaching, making the content meaningful. While a comma monster is not authentic, the issues of not knowing commas that the monster presented are very relatable. In addition, the classroom setting was relevant to all participants.

Interest	Novelty of the puzzles, familiarity of the classroom setting, and the incitement of curiosity through variation in puzzles made the room interesting.
Social interaction	Students were told that they should collaborate, in particular to complete the room within the time limit. They also had the opportunity to negotiate meaning with the EER facilitator who was giving hints.
Challenge/skills balance	Doable challenge was addressed by aspects such as using different puzzle formats, allowing peer collaboration, and creating short puzzles.
Autonomy/structure balance	Each puzzle had enough copies or pieces of the handout so that members of the group could each complete one if they wanted to. Participants could also make choices such as asking for hints.
Support and feedback	A “school secretary” was available through Zoom to provide up to three hints during the experience when asked. Participants had a handout with comma rules that they found as part of the first puzzle and could use throughout the experience.

After going through several design iterations, the following 3-phase design for the experience was finalized:

Pre-Experience Phase

Before participants arrived for their scheduled EER time, they were supposed to complete an online survey about their previous experiences with escape rooms and their self-efficacy with comma use. When they arrived at the staging room on the day of their escape, they completed a pencil-based comma pre-test. Then they watched a brief video clip from an Xbox connected to a TV monitor that showed the “Comma Monster” explaining what was going on with the commas. [Figure 2](#) shows a screenshot of the animated Comma Monster video, in which the monster explains that it would eat all the commas and then the students would not be able to communicate effectively if the students did not solve the comma problems in the classroom. Finally, participants were provided orally with the rules for EER behavior and were escorted around the corner to the room.



Figure 2. *Comma Monster Video Introducing the EER*

Experience Phase

The final escape consisted of six comma exercises of different kinds. The answers to each exercise led to a code or clue leading to the next puzzle. Four of the puzzles had different numbers of questions that provided codes as the answers to multiple choice questions; this helped participants figure out what they were looking to unlock next (e.g., a lock with three numbers or four numbers). The six puzzles were:

1. In the opening clip of the Comma Monster, participants were given an oral clue to find a numbered

list of comma usage rules on the teacher's desk that also contained clues to finding the second puzzle.

2. Students were tasked to put together a sentence on blackboard blocks in the correct order and then use the letter code on the back of the blocks to open a locked box.
3. This paper puzzle asked participants to note which comma rule each sentence used. This provided a number code to the next box.
4. This puzzle asked how many commas were missing in each of 4 sentences. The responses, read in order, were the code to the next puzzle box.
5. This puzzle asked the participants to choose the correctly punctuated answer in each of 3 items. This led to the code to open the final puzzle box that held blacklight flashlights.
6. This was a two-part puzzle: 1) Participants had to use the flashlights to find the code to get into the computer (it was written on the chalkboard with blacklight ink); 2) they had to remove any extra commas from a letter to the Comma Monster that was on the computer in the room and send their responses to the Comma Monster's email. Participants received a return email from the Comma Monster (a monitor in a nearby room) that said either that they had been successful or that they made some mistakes that they should use the comma rules to find.

Except for puzzle #2, each puzzle had enough copies of the handout so that members of the group could each see it. These were left in the room by the participants and were collected by the research team.

Participants had 45 minutes to solve the puzzles and "escape." After they completed the final puzzle, certificates were sent to the printer in the EER from the "Comma Monster" either congratulating them for completing the room or for participating if they had not finished.

Post-Experience Phase

After the allotted time was over, the participants completed the post-experience survey and a research team member retrieved the players. They returned to the staging room for the post-test and debrief, as noted below.

Data Sources and Analysis

The study employed seven data sources to provide a complete description of the process and outcomes of the room experience.

Participant Pre-experience Survey

The online pre-experience background survey included questions that the literature indicated might affect their engagement and other variables during the EER experience. Questions included their age, grade level, gender, type of learning risk, if any (bilingual/ESL, low-income, disabled, etc.), a study description and consent to participate, previous experience with escape rooms, previous study in English punctuation, perceived ability to use commas, perceived level of critical thinking skills, and their interest level in learning about comma use. These data were analyzed with descriptive and inferential statistics to provide overall counts and significance findings, noted later in this paper. The pre-experience survey was a researcher-created instrument designed specifically for the goals of this study. While reliability and validity testing were not conducted due to the study's exploratory nature and small sample size, the survey was carefully developed based on existing literature and reviewed to align with the objectives of the EER.

Content Pre-Test

The participants completed a "What do you know about comma usage?" assessment during the staging phase. This pre-test comprised 12 questions of the same sort that the students would need to answer during the experience and were contextualized as part of the experience. Each assessment was graded and checked to determine the number correct. The responses (right=1, wrong= 0) were input into a spreadsheet for descriptive statistical analysis after the post-test.

Video Capture of EER Events

A facilitator watched and took notes of the process that students used and any unexpected actions or events. The videos were recorded on the Zoom platform for double-checking. Observation data were used to explain statistical and other data and to review the EER design.

Task Documents

The puzzle documents were collected to see whether all the participants had completed the puzzles and so that the puzzles could not be shared with future participants.

Participant Post-Experience Survey

Participants completed an online post-experience survey, the average duration of which was four minutes, while still in the EER. This survey included content and skill questions from the pre-experience background survey stated in past tense; it also asked what they thought they learned, how well they learned it, their interest in using commas in their teaching, and their perceptions of their critical thinking and collaboration skills.

Content Post-Test

The content post-test was the same as the pre-test. Although this happened close to the pre-test, students were unaware that there would be a post-test, so they could not specifically prepare for it. The grading was the same as the pre-test and scores were compared to the pre-test as an indication of learning outcomes. A paired-samples t-test, having met the necessary assumptions for normality and homogeneity, was conducted to compare the gain score from the pre- and post-tests on comma use.

Debrief Interviews

The debrief interview consisted of six open-ended questions asked to the participant groups about their experience such as “How was your experience?” and “How might you apply the information you learned?”; interviewers were also encouraged to ask general follow-ups as needed in order to understand participant self-expressed engagement in the experience. All interviews were recorded and transcribed. The data were then read by two members of the research team for overall patterns, and descriptive anecdotes based on the task engagement facilitators were culled from the data.

Methodological Limitations

This study has several limitations that can be addressed in future research. First, the study included a limited number of participants. Although the EER was closed because of COVID after the data noted here were captured, the data appeared to be sufficient to come to some conclusions about the experience and the EER design. Second, not all of the participants completed all of the data sources; this was due to different EER staff conducting the EER and the timing of the groups. This means that some survey data are missing, but all of the participants were videotaped, pre- and post-tested, and debriefed. To address the missing survey data, we adopted an aggregate reporting approach, which allowed us to include the data that were available without discarding entire datasets due to incompleteness. Specifically, for descriptive analyses, missing data were excluded listwise from specific variables rather than the entire dataset. However, cases with incomplete survey responses were not included in the statistical analyses, as these analyses required complete data for all variables. Despite this limitation, the remaining data were sufficient to draw meaningful conclusions and provide insights into the EER experience.

Third, the study relied partly on self-reported data from participants, which is subject to biases such as self-selection bias, social desirability bias, recall bias, and potential inaccuracies in participants' self-assessments. To mitigate these limitations, we incorporated measures such as pre- and post-tests and video recordings to supplement the self-reported data. While self-reported measures provided valuable insight into participants' perspectives and perceptions, future research could further reduce bias by using additional objective measures or triangulating findings with external evaluators. In addition, random assignment to groups and/or a focus on one or more intact courses might provide data that volunteers may not.

Findings and Interpretation

Findings are presented here in the order of the research questions. Interpretations result in the guidelines presented in following section.

Student Perceptions

Before starting their EER experience, the participants were asked to share what they liked about escape rooms. In the open-ended section of the pre-participation survey, participants highlighted aspects such as fun (4 responses), learning (3), excitement (3), challenge (2), inquiry (2), and interest (1); these responses indicate components of task engagement. Because some of the students who responded had not experienced an escape room, it appears that some of these responses might have come from stories that the participants had heard or other sources of information about escape rooms.

In their post-experience interviews, these same aspects of task engagement were also mentioned. For instance, the EER experience was referred to as cool (5), fun (4), engaging (3), exciting (2), an adventure (1) joyful (1), helpful (1), active (1), I love it (1), and it gives you “a feeling of winning and competition.” These words are not often presented in the literature as a reaction to comma study and were offered without specific prompting about the task engagement facilitators. This indicates that the participants did perceive the TE facilitators during the experience. In addition, in their post-participation surveys, 15 of the pre-/in-service educators expressed interest in participating in other academically- themed escape room experiences focusing on a variety of literacy and technical subjects such as grammar (15), mathematics (3), forensics (3), languages (2), politics (1), history (1), critical thinking (1), and sustainability (1).

Interest Rating

The participants were asked to rate their interest in using commas both before and after their participation in the EER on an 11-point slider scale, with zero indicating no interest, five indicating average interest, and 10 indicating high interest. Results are shown in [Table 2](#).

Table 2. *Interest Rating Across Pre- and Post-Participation Surveys*

Slider Level	Pre-Participation	Post-Participation
1	1	0
2	1	0
3	0	0
4	0	0
5	3	1
6	2	1
7	10	1
8	2	0
9	0	6
10	6	6

Note. Participant numbers vary across surveys (25 participants took the pre-participation survey, 15 took the post).

The frequency of responses across both surveys tended to cluster above average, with average pre-test

response 7.04 and post-test 8.8. In the pre-test, two participants indicated below average interest, while in the post, none did.

To further compare participant interest level, we conducted a Wilcoxon test using the data from the participants from whom we had both pre- and post-participation surveys. The Wilcoxon test is a non-parametric test which is used to compare pairs of data that are non-normally distributed, as was the case in our survey results. The results of the Wilcoxon test indicated that the pre-participation scores ($Md = 7$, $n = 15$) were statistically lower than post-participation scores ($Md = 9$, $n = 15$): $z = -2.249$, $p = .024$, with a large effect size, $r = 0.61$, indicating that the participants' interest in using commas was significantly higher after their EER experience.

Learning Outcomes

The participants' punctuation knowledge was diverse at the onset of the study, with more than half of the participants not recollecting when they had last studied about punctuation, 12 having studied punctuation within the last 4 years, and 2 within the last 8 years. However, in their post-participation survey, most respondents said that they found the EER experience useful in developing new learning. For instance, there were 12 references to having acquired grammar skills such as using commas properly; seven to 21st century skills such as critical thinking, collaboration, problem-solving, and group work; five to professional realizations such as observing individual learner differences when designing collaborative tasks; two to metacognitive strategies such as the need to practicing following instructions better; two to real-world applications such as writing street addresses and unlocking a lock; and two to game-related strategies such as examining the room fully for clues. In other words, without prompting to notice the content and skills outcomes of the EER, the participating students were able to note a number of them.

Ability to Use Commas

As in the beta tests of the EER, participants made gains in comma use. The comparison of the gain scores showed a significant difference between the scores from the pre-test ($M=10.44$, $SD=2.03$) and the post-test ($M=11.11$, $SD=1.77$), $t(17)=3.6$, $p < .05$. These results suggest that the students' knowledge of comma use improved after participating in the EER. While, in actual terms, the gain on the 12-item test was small (about one more answer correct) and the pre-test scores on average were already high, it still indicates that the 45-minute intense focus on commas made at least some difference in participants' comma use.

To further compare participant comma use perceptions across the two surveys, we conducted a Wilcoxon test using the data from the participants from whom we had data from both surveys. The results of the Wilcoxon test indicated that the pre-participation scores ($Md = 8$, $n = 15$) were statistically lower than post-participation scores ($Md = 9$, $n = 15$): $z = -2.730$, $p = .006$, with a large effect size, $r = 0.73$, indicating that participants perceived their ability to use commas significantly higher after their EER experience, whether they actually could or not. This may be because they became more aware of the rules from using them throughout the experience.

Critical Thinking Skills

The participants' responses seemed to be very similar across both surveys, most of them indicating that they perceived that they had above-average critical thinking skills. In the post-participation surveys, however, two participants indicated a low rating for their critical thinking skills. This could indicate that the experience helped them to think about their skills in general, or it could imply that they had found the puzzles too challenging. The interview results provided some grounding for this implication. One participant mentioned that some locks between different puzzles were different—3 vs. 4 digits—and this made the puzzles confusing; another participant shared that the first puzzle—the blocks—was the hardest, as there were no clear instructions; and two participants expressed that the last puzzle—emailing the Comma Monster—was the most challenging one, as there were no clear directions or examples as to finding the password using the blacklight. All participants, however, were content with the number of puzzles and their content-related difficulty level.

To further compare the participants' perceived critical thinking skills before and after the EER experience, we conducted a Wilcoxon test using the data from the participants from whom we had both pre- and post-participation surveys. The results indicated that the pre-participation scores ($Md = 8, n = 15$) did not differ significantly from the post-participation scores ($Md = 9, n = 15$): $z = -1.131, p = .258$, with a medium effect size, $r = 0.36$. One of the reasons for this statistic was the exclusion of five data entries (ties) from the analysis. This could indicate that the participants perceived critical thinking skills did not change before and after their EER experience or that this general question was not a good indicator of specific perceptions.

Teamwork/Collaboration

Again, the participants' responses seemed to be very similar across both surveys, with most participants perceiving their teamwork/collaboration skills as above average both before and after the EER experience. This finding was further corroborated through the interviews. When asked how they would feel if they had to do the task alone, all participants highlighted the importance of teamwork and collaboration. One participant thought doing an EER alone would be "boring" and stated, "it's more engaging with a team." Another participant doubted that they would be able to finish the task without teamwork.

The results of the Wilcoxon test of these results indicated that the pre-participation scores ($Md = 8, n = 15$) did not differ significantly from the post-participation scores ($Md = 9, n = 15$): $z = -1.435, p = .151$, with a medium effect size, $r = 0.42$. This indicates that the participants perceived teamwork/collaboration skills did not change before and after their brief EER experience.

Composite Results

In addition to comparing the scores of individual task engagement facilitators in the survey, we conducted a composite analysis, comparing pairs of all perception indicators combined. The results of the Wilcoxon test indicated that the pre-participation scores ($Md = 7, n = 60$) were statistically lower than the post-participation scores ($Md = 9, n = 60$): $z = -3.738, p = .000$, with a large effect size, $r = 0.54$, indicating that the participants perceived the EER as effective in: supporting their ability to use commas; using critical thinking skills; sustaining interest; and requiring teamwork/collaboration skills.

Effectiveness of an EER as a Teaching and Learning Tool

As we hoped, the interest in EERs went beyond participation, as 15 participants shared in their post-participation surveys that they planned to design an EER for their language students in the future. In the debriefs, the reasons for this consideration were made clearer as, drawing from their own experience and learning, participants briefly mentioned how the medium could support students' "having fun and think[sic]," "immediate feedback," "entertaining, interacting, collaborating," and "problem-solving." One of the participants mentioned that comma use was the hardest topic for their 6th graders, and they thought an EER would be a great tool to help. This participant further mentioned that they designed an EER for their undergraduate student teaching that had students solve a mystery using the content of music videos. However, two of the participants expressed concern over the dominant culture of their home countries considering EERs "as play," sharing that they might not be able to design one and would simply "adapt the puzzles" using an online escape room platform. Another participant also expressed concern over the "time and energy" required to design an EER. To improve the EER experience, some participants offered suggestions such as using audio and video effects and clues, background music, and more technology.

EER Design Recommendations

The results of this design-based study indicate both accord with the literature and the need for more specific guidelines in continuing to work with graduate language teacher education in EERs. Student outcomes show us where we have more design work to do. For example, the videos showed that, in one of the groups, one student appeared to want to work by himself. His group members were so engaged in the process together that they seemed not to notice. This suggests that, if social interaction is a goal of EER use, puzzles must be constructed so that participants are required to work together. This may mean incorporating techniques such as jigsaw (in which different students have different pieces of the information to share) and

other cooperative learning strategies. If social interaction is for some reason not a focus, and not all students like to interact to learn (as in Mathieson & Duca, 2021), perhaps individuals could try the room, or students could have the choice not to work together if that seems logical. This, however, defeats a major goal of EER experiences.

Our results also show that puzzles in the escape room should be considered from the participant point of view. For example, the videos and interviews showed that the puzzle on the blocks was the most difficult to figure out where and what the actual puzzle was. The EER team thought it easiest because participants only had to create a sentence from words and punctuation marks on blocks. The difficulty could perhaps be because the puzzle was different than what the student participants expected, or that the instructions within the comma rules were too opaque, or, more likely, that it came too early in the EER when the participants were not yet acclimated to the room. This result suggests that the puzzles should be in the order of easiest to most difficult for the participants (focusing on authenticity) and/or that participants should be provided with general ideas of what the puzzles will be like before they enter the room, or both.

Mathieson and Duca (2021) noted that their participants wanted to talk about their experiences with others who had had them, and we found the same to be true as the participants emerged from the room talking excitedly and trying to explain their experience. Adding a content debrief (Taraldsen et al., 2022) or some other opportunity where the participants lead the discussion (rather than only pre-created interview questions) might be effective in helping students reflect and “cement” their learning, but only if it focuses on both the process and the content.

While task engagement was not specifically measured in this study, participants clearly perceived the experience as interesting, useful or authentic, providing opportunities for social interaction, and providing just-in-time feedback. Engagement indicators (Egbert & Panday-Shukla, 2024) such as negotiating meaning, collaborating, focusing on group goals, exhibiting positive affect, persisting, and making an effort were clearly seen in the videos and expressed in the interviews. This suggests that a focus on and measure of task engagement might be a useful way to understand the outcomes of EER use.

Overall, the study results lead to a number of guidelines that we intend to follow for the next iteration of the physical room. In addition to those mentioned previously in this section, guidelines to improve learner perceptions and outcomes include to:

- Incorporate appropriate technology to make the rooms more authentic to the students and possibly more engaging. Having participants figure out how to use one or more new language technologies could be a content focus of the room or be part of a critical thinking process for which the students had been prepared in class.
- Provide opportunities for critical thinking about the content and make this goal explicit. Our EER did not require students to think critically about the content, per se, but rather about how to find the content and what actions to take. If critical thinking is an explicit goal, participants must be aware of and prepared for the type of thinking that they will be asked to do.
- Consider whether hints will only be provided if asked for or whether learning is best served by also volunteering help based on participant indicators; in this study, the time it took participants to complete the block puzzle could be an indicator that help is warranted.
- Have other faculty or teachers experience the room and provide feedback. If the room is not built for a specific course or instructor, this feedback might help the room design be more appropriate for diverse student participants.

Further, as Bakkum et al. (2021) discuss, we hope that other language teacher education faculty whose students participate in our EERs might support their students in developing their own escape plans, or even classroom puzzles if they cannot access a whole room. Doing so can provide even more practice with the content and help them to think in pedagogical terms about task engagement and about presenting and assessing content and skills.

Conclusions and Implications

Our findings agree with the literature that EERs must be well-designed in order to have the desired outcomes (Grande-de-Prado et al., 2020). As instructors, researchers, and students, the design team found the use of a DBR process and task engagement model to ground EER design a useful way to frame the work that language teacher education students must learn to do. In other words, teachers cannot teach what they do not know, and their students do not learn what they do not have access to. Outcomes from this study can support creating graduate language teacher education programs that focus on both engaging students across language learning tasks and making task engagement facilitators explicit, whether in EERs or in some other form.

Implications for Teaching

There are many ways that graduate language teacher education students can learn about EERs and the theoretical foundations upon which they are built. For example, students, in particular if they have not participated in previous escapes, can start by completing simple online escapes, then they may develop a simple click-and-answer digital room. Next, they may participate in games through breakout.edu, which involve boxes and clues in the classroom. With a pre-brief about both task engagement and content learning, participants in the EERs might have even greater learning in these areas. Instructors can help students to make explicit the task engagement facilitators and how they are embedded in the rooms and puzzles. Having graduates develop and implement rooms for pre-service language teacher education undergraduates might support their understanding of both research and instruction. Used too often, EERs might get boring like any task. However, there are many formats and ways for students to develop their own EERs, and using them with specific goals can both engage and instruct.

Implications for Research

The overarching question of the next phase of EER research might be: What kind of escape room, integrated how and into what kind of course or syllabus, at what level of learning, for what kind of learners, is likely to be effective for what specific learning purpose? (adapted from Garrett, 2009). We expect that researchers will continue to employ different methodologies in their EER research, and we hope that the call for outcomes data from EER use is heeded more often. In addition, using more extensive pre/post content tests might help to explain some of the discrepancies in the current study data.

Future EER research can also shed additional light on aspects of the task engagement model, such as whether participants who are not interested in the room theme will actually be engaged in the process, and which facilitators are most important in EERs designed for language learning. Another important issue is transfer, or the use of new knowledge in other situations. Studies using intact language teacher education classes might be easier to follow up with than random or volunteer groups and could provide information on the extent of transfer of the EER content. In addition, whether the outcomes found in this study were long-term gains is unclear, so it would be useful to test long transfer in the next iteration of the comma EER. It might also be a benefit to ask students what they know about commas that they did not previously know to get an idea of the exact knowledge that they had acquired.

Overall, we are excited about the possibilities that EER design and use offer our graduate teacher education students, and we plan to explore additional opportunities that EERs and other engaging tasks can provide to them.

References

- Bakkum, M., Richir, M., Sultan, R., de la Court, J., Lambooj, A., van Agtmael, M., & Tichelaar, J. (2021). Can students create their own educational escape room? Lessons learned from the Opioid Crisis escape room. *Medical Science Educator, 31*(1), 1-7. <https://doi.org/10.1007/s40670-021-01425-5>

- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of The Learning Sciences*, 2(2), 141–178. https://doi.org/10.1207/s15327809jls0202_2
- Clarke, S., Peel, D., Arnab, S., Morini, L., Keegan, H., & Wood, O. (2017). EscapED: A framework for creating educational escape rooms and interactive games for higher/further education. *International Journal of Serious Games*, 4(3), 73-86. <https://doi.org/10.17083/ijsg.v4i3.180>
- Collins A. (1992) Toward a design science of education. In E. Scanlon & T. O’Shea (eds.). *New directions in educational technology* (pp. 15-22). Springer. https://doi.org/10.1007/978-3-642-77750-9_2
- Cotoc, A., & Mudure-Iacob, I. (2024). A #co-teaching story from the digital pedagogical framework: Escaping the routine language class via a (new media) digital escape room. *LinguaCulture*, 15. <https://doi.org/10.47743/lincu-2024-si-0386>
- Csikszentmihalyi, M. (2014). *Flow and the foundations of positive psychology: The collected works of Mihaly Csikszentmihalyi*. Springer.
- Dilek, S., & Dilek, N. (2018). Real-life escape rooms as a new recreational attraction: The case of Turkey. *Anatolia*, 29(4), 1-12. <http://doi.org/10.1080/13032917.2018.1439760>
- Egbert, J. & Panday-Shukla, P. (Eds.). (2024). *Cross-disciplinary task engagement: Evidence-based research and strategies to increase student achievement*. Routledge.
- Egbert, J., Shahrokni, S., Zhang, X., Abobaker, R., Bantawtook, P., He., H., Bekar, M., Roe, M.F., & Huh, K. (2021, February). Language task engagement: An evidence-based model. *TESL-EJ*, 24(4). <https://www.tesl-ej.org/pdf/ej96/a3.pdf>
- Garrett, N. (2009). Computer-assisted language learning trends and issues revisited: Integrating innovation. *Modern Language Journal*, 93(1), 719-740. <https://doi.org/10.1111/j.1540-4781.2009.00969.x>
- Grande-de-Prado, M., Garcia-Martin, S., Baelo, R., & Abella-Garcia, V. (2020). Edu-escape rooms. *Encyclopedia*, 1, 12-19. <https://doi.org/10.3390/encyclopedia1010004>
- Horat, B., & Grillenberger, M. (2021, October 18-20). *CS learning in educational escape room* [Paper presentation]. WiPSCE ‘21: The 16th Workshop in Primary and Secondary Computing Education, Virtual Event, Germany. <https://doi.org/10.1145/3481312.3481327>
- Kanakambaran, R. (2021). Educational escape rooms for teaching. *Preprints*, 2021030125. <http://doi.org/10.20944/preprints202103.0125.v1>
- Lopez-Pernas, S., Gordillo, A., Barra, E., & Quemada, J. (2019). Examining the use of an educational escape room for teaching programming in a higher education setting. *IEEE Access*, 7, 31723–37. <https://doi.org/10.1109/ACCESS.2019.2902976>
- MacDonald, M., Badger, R., White, G. (2000). The real thing? Authenticity and academic listening. *English for Specific Purposes*, 19(3), 253-267. [https://doi.org/10.1016/S0889-4906\(98\)00028-3](https://doi.org/10.1016/S0889-4906(98)00028-3)
- Manojlovic, H. (2022). Escape room as a teaching method. *Opus et Educatio*, 9(2), 178-188. <https://doi.org/10.3311.ope.504>
- Martens, S., & Crawford, K. (2018). Embracing wonder and curiosity: Transforming teacher practice through escape room design. *Childhood Education*, 95(2), 68–75. <https://doi.org/10.1080/00094056.2019.1593764>
- Mathieson, A., & Duca, E. (2021). STEM escape rooms for public engagement. *Research for All*, 5(2), 347-55. <http://doi.org/10.14324/RFA.05.2.10>

- Mudure-Iacob, I. (2024). Game-mode activated: Building communities of practice through digital escape rooms as ESP teaching and learning tools. *Studia UBB Philologia*, 69(4), 75-96. <https://doi.org/10.24193/subbphil.2024.4.03>
- Nasri, N., Zaidi, N., Kamarudin, S., Hanafi, M., & Nasir, Z. (2024). Digital escape room: Students' perceptions as an educational tool in learning reading skills. *International Journal of Education*, 16(2), 21-37. <https://doi.org/10.5296/ije.v16i2.21752>
- Nicholson, S. (2015). *Peeking behind the locked door: A survey of escape room facilities* [White Paper] <http://scottnicholson.com/pubs/erfacwhite.pdf>
- Nowacki, M., & Stasiak, A. (2023). Identification of memorable game experiences: Analysis of escape rooms visitors' experiences. *Games and Culture*. <https://doi.org/10.1177/15554120231190648>
- Ouariachi, T., & Wim, E. (2020). Escape rooms as tools for climate change education: An exploration of initiatives. *Environmental Education Research*, 26(8), 1193-1206. <http://doi.org/10.1080/13504622.2020.1753659>
- Queiruga-Dios, A., Sanchez, M., Queiruga-Dios, M., Martinez, G., & Encinas, A. (2020). A virus infected your laptop. Let's play an escape game. *Mathematics*, 8(2), 166. <https://doi.org/10.3390/math8020166>
- Rad, H., & Alipour, J. (2024). Investigating the effectiveness of digital escape rooms in enhancing L2 learners' vocabulary achievement, retention, and learning motivation. *Computer Assisted Language Learning*. <https://doi.org/10.1080/09588221.2024.2436448>
- Reschly, A. L., & Christenson, S. L. (2013). Grade retention: Historical perspectives and new research. *Journal of School Psychology*, 51(3), 319-322. <https://doi.org/10.1016/j.jsp.2013.05.002>
- Rowe, D. (2020). Library escape room. In M. Bennett (ed.), *Seventy-third annual conference of ATLA* (pp. 165-173). ATLA Summary of Proceedings. <http://doi.org/10.31046/proceedings.2019.1626>
- Saban, M., Grunke, M., & Khan-Horwitz, J. (2024). Enhancing spelling competence for English as a foreign language young learners through digital escape rooms. *Frontiers in Education*, 9, 1434336. <https://doi.org/10.3389/educ.2024.1434336>
- Santamaria, A., & Alcalde, E. (2019). Escaping from the English classroom: Who will get out first? *Aloma*, 37(2), 83-92. <http://doi.org/10.51698/aloma.2019.37.2.83-92>
- Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2014). Student engagement in high school classrooms from the perspective of flow theory. In M. Csikszentmihalyi (ed.). *Applications of flow in human development and education* (pp. 475-494). Springer.
- Taraldsen, L. H., Haara, F. O., Lysne, M. S., Jensen, P. R., & Jenssen, E. S. (2022). A review on use of escape rooms in education—touching the void. *Education Inquiry*, 13(2), 169-184. <https://doi.org/10.1080/20004508.2020.1860284>
- Veldkamp, A., van de Grint, L., Knippels, M.-C., & van Joolingen, W. (2020). Escape education: A systematic review on escape rooms in education. *Educational Research Review*, 31, 100365. 1-18. <https://doi.org/10.3389/educ.2024.1434336>
- Voreopoulou, A., Mystakidis, S., & Tsinakos, A. (2024). Augmented reality escape classroom game for deep and meaningful English language learning. *Computers*, 13, 24. <https://doi.org/10.3390/computers13010024>
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wargo, J., & Garcia, A. (2021). (Re)reading the room: The literacies of escape rooms. *Journal of Curriculum and Pedagogy*, 1-26. <http://doi.org/10.1080/15505170.2021.1960224>

Weisberg, L., Dawson, K., & Dana, N. (2022). Engaging preservice teachers in the design of digital breakout games in an educational technology course. *Journal of Digital Learning in Teacher Education*, 38(2), 71-88. <https://doi.org/10.1080/21532974.2022.2038738>

About the Authors

Joy Egbert is Regents Professor of Multilingual Learners and Education Technology at Washington State University, USA. Her teaching and research interests include computer-assisted language learning, teacher education, and task engagement. She has published 16 books and numerous articles in these areas.

E-mail: jegbert@wsu.edu

Seyed Abdollah Shahrokni is the Open Publishing Platform Manager at Texas A&M University. His scholarly interests include task engagement, technology-enhanced learning, and open pedagogy.

E-mail: s.shahrokni@tamu.edu