The Role of Virtual Characters in Scaffolding Inquiry-Based Learning in Genetics

Laura Marie Newald

Psychology Programme, University of Twente

Faculty of Behavioural, Management and Social Sciences

202000384: BSc Thesis Psychology

Bachelor Thesis

Supervisor: Hannie Gijlers

Second Supervisor: Chandan Dasgupta

July 1, 2025

Word Count: 7214

Abstract

As virtual characters become more common in educational tools, it is worth exploring whether they support students' self-reported perceived learning and satisfaction. This study investigated whether adding a simple, cartoon-style character to an online, inquiry-based genetics lesson would make a difference for secondary school students. This study took place in a real classroom setting, involving 52 students aged 11 to 13, and compared two conditions of the same lesson – one with a static virtual character and one with only pure text instruction.

Students filled out surveys, and their activity in the lesson was tracked. The results showed no significant differences in students' self-reported perceived learning and satisfaction with the genetics course. Interestingly, those who had the character version answered more questions correctly but also spent less time in the lesson overall. In the end, the character did not seem to improve outcomes in any clear way.

These results suggest that adding a static virtual character is not enough to make digital learning better. For these agents to help, their design needs to be more thoughtful, especially when it comes to how they interact with students and support the learning goals. By testing this in real classrooms, the study offers some useful takeaways for educators and designers working to build better digital learning tools.

Introduction

In modern science education, the concept of inquiry learning – an instructional approach emphasising self-directed exploration and guidance-based learning – plays a central role. In Germany, this approach can be found in the curriculum for subjects such as chemistry and biology (Kultusministerkonferenz, 2024). In both subjects, students often conduct experiments to make predictions, test their hypotheses and afterwards analyse the results. Inquiry learning emphasises scientific inquiry, where students build knowledge through applying methods and strategies that are comparable to those used by experienced scientists (Pedaste et al., 2015). To understand how effectively inquiry learning supports students, researchers often focus on key outcomes like self-efficacy, perceived learning and satisfaction. Self-efficacy refers to students' confidence in their ability to succeed in online learning tasks (Tsai et al., 2020). Perceived learning is how much students feel they have understood and gained from the course, while satisfaction measures their overall positive experience with the learning environment and activities (Gray & DiLoreto, 2016). These factors influence motivation and engagement, which are critical for successful inquiry learning. However, students often find it challenging to engage with genetics topics (Mussard & Reiss, 2022). Students find processes in genetics difficult to grasp due to the use of scientific terminology. For instance, many genetic processes involve terminology, such as "DNA polymerase" or "homozygous", which can hamper the understanding and learning processes (Mussard & Reiss, 2022). Additionally, genetic processes entail abstract processes that are not directly observable and therefore inaccessible (Topçu & Şahin-pekmez, 2009). Examples include genes and DNA that operate at the molecular level, which makes it difficult to clearly understand the processes. Computer-supported inquiry learning, particularly when enhanced with simulations, can make abstract phenomena visible and accelerate complex processes, thereby facilitating the exploration of multifaceted topics (Agyei & Agyei, 2021).

Therefore, inquiry learning is suitable for engaging with complex topics, since students are invited to explore and actively deal with the content. Nonetheless, there is a need for guidance in inquiry learning because students might not always make the correct choices or know exactly what to do. Guidance helps students remain on track and address difficulties in complex subjects like genetics and raises the question of how teachers can support students in navigating complex scientific inquiry.

One solution that might help students navigate complex scientific inquiry, while also getting a balanced amount of support, is scaffolding. Scaffolding is a teaching approach that gives students temporary support so they can complete a task or acquire abilities they would not be able to do on their own (Van de Pol et al., 2018). The role of scaffolding is to help students address their learning difficulties by offering assistance in establishing suitable learning objectives, developing necessary skills needed to achieve these objectives, and engaging in self-reflection (Lin et al., 2011). However, as the student becomes more proficient, this help is gradually reduced (Van de Pol et al., 2018). There are various forms of scaffolding. One of them is “Soft Scaffolding”, which entails assistance provided by a teacher or a peer in the moment that is tailored to meet the students’ present needs to overcome a learning obstacle (Renninger & List, 2012). “Contextual Scaffolding” refers to cues or suggestions that offer situational support to assist students in connecting ideas or overcoming certain learning challenges (Renninger & List, 2012). Related to this study, these types of scaffolding are directly relevant as they offer strategies to support students in inquiry-based learning. For instance, soft scaffolding could be used within a digital learning environment where digital agents provide support to the needs of students. On the other hand, contextual scaffolding could be used to provide hints or suggestions within the learning environment to guide students to make connections. These examples or scaffolding strategies can provide students with different levels of support to facilitate their understanding. Several studies have highlighted the importance of scaffolding in learning environments, including digital platforms. For instance, a study by Mamun and Lawrie (2023) has shown that scaffolding enhances cognitive presence and helps students to solve problems they are unacquainted with. Therefore, scaffolding in inquiry learning, especially for challenging topics like genetics, can be a potential solution.

In this regard, virtual agents can serve as a form of scaffolding in order to provide students with personalised support in inquiry learning, as well as enhance their engagement (Zhang & Whu, 2024). Virtual agents are characters presented on a computer screen that guide students through learning materials like videos, games, or interactive lessons (Clarebout & Heidig, 2012). Some of these agents are animated and respond to the actions of students. Others are just still images, illustrations that don't move, but still deliver instructions through text (Clarebout & Heidig, 2012). For this study, the focus lies on static virtual agents, which don't move or talk but appear on the screen to support and guide learners with speech bubbles as they work themselves through the inquiry-based lesson. A virtual character is selected because, according to the social presence theory, virtual characters can increase learners' sense of social connection and reduce feelings of isolation in digital settings (Short et al., 1976). However, previous research has cautioned that highly interactive or animated virtual agents may distract learners or impose excessive cognitive load, thereby reducing learning effectiveness (Clarebout & Heidig, 2012). To balance these effects, this study employs a static virtual character designed to provide social support and guidance without overwhelming students or diverting attention from the content.

Research has identified several important roles virtual characters can play to scaffold learners. Virtual characters can act as expert tutors by teaching complex concepts and offering guidance (Gulz, 2004). They can also act as peer-like companions, providing encouragement and social support (Gulz, 2004), and as engagement tools that make the learning experience more enjoyable and motivating (Zhang & Whu, 2024). By fulfilling these roles, virtual characters help reduce cognitive load (Brachten et al., 2020) and offer timely support. This makes them particularly effective for supporting learning in complex and abstract subjects such as genetics. These findings highlight the possible helpfulness of virtual characters to support the inquiry learning process, as well as keep students engaged and motivated to learn about new topics. However, a key consideration is also how feedback provided by these static characters is perceived and how this perception might influence effectiveness, an aspect that has received limited attention in existing research. This question helps to investigate the impact of virtual characters in educational settings, particularly in terms of the impact on students' perceived learning and satisfaction.

Even though existing studies support the notion that digital agents have a positive impact in digital learning environments (e.g., Brachten et al., 2020; Clarebout & Heidig, 2012; Gulz, 2004; Zhang & Whu, 2024), there remain gaps in understanding how these agents affect student perception and satisfaction, especially in complex and abstract subjects like genetics. While prior research has explored overall learning gains and engagement, fewer studies have examined how static virtual characters specifically influence students' confidence in completing tasks, comprehension of difficult content, and feelings of support. According to Ergül and Koç (2013), most research focuses on animated or interactive agents, leaving the effectiveness of non-animated, static agents less well understood. Addressing these gaps is essential to optimise the design and implementation of static virtual characters in inquiry-based science education, particularly for enhancing problem-solving skills in secondary genetics education.

Taking into account the theoretical potential and positive findings from previous research on the benefits of virtual characters as scaffolding tools in inquiry-based learning, the purpose of this study is to investigate the impact of static virtual characters on student perception of learning and satisfaction in online inquiry-based learning, more specifically in the context of genetics. The relationship between the presence of a static virtual character and students' perception of learning and satisfaction is assessed by conducting a group comparison to investigate differences in students' self-reported learning perceptions and satisfaction levels. Therefore, the independent variable is the presence or absence of a static virtual character in the learning environment. Perceptions of student learning and satisfaction in online learning environments, the study's dependent variables, are evaluated using self-reported measures, in which students reflect on their experiences and sense of guidance during the inquiry process. In addition, students' self-efficacy for online learning is measured prior to completing the learning environment to explore its potential influence on perceived outcomes. Although the main purpose is to assess the relationship between the presence of a virtual character and students' perception of learning and satisfaction, it is important to highlight any potential drawbacks beforehand. Specifically, the relatively small sample size (about 50 students) may limit the statistical power and make it challenging to draw firm conclusions. Nonetheless, this study offers insightful information about the potential effect of static virtual characters on the inquiry learning process in complicated science learning.

This study hypothesises that the presence of static characters will enhance both student self-reported perception and satisfaction in inquiry-based learning for genetics. Although static virtual characters are non-animated and do not provide dynamic or personalised feedback, their presence offers a consistent visual guide that can help students feel more supported and confident. By organising information and signalling key points, static characters may reduce cognitive load and make complex content easier to understand, which could enhance students' learning perception and satisfaction. The research question guiding this study is: "Does the presence of a static virtual character improve students' self-reported perception of their learning and satisfaction in an online inquiry-based genetics lesson?”. If this positive relationship is found, their use may be expanded to new age groups or scientific fields, which might modify digital scaffolding in an educational context. To investigate this, a between-subjects design is employed, comparing a group of students using character-based instructions with a group of students using only text-based instructions.

Methods

Participants

The full sample consisted of 52 children, aged 11 to 13 (M = 12.02, SD = 0.73), with 25 boys and 26 girls. One student chose not to disclose whether they identified as a boy or a girl. This age group corresponds to a stage where German students begin to seriously engage with computers in their schooling, and biology topics such as genetics become increasingly relevant and interesting in their curriculum. The participants were recruited from two classes, 6a and 6c, at a Gymnasium in Germany – a type of academically oriented secondary school that prepares students for university education. Therefore, the sample represents students who are familiar with certain topics in genetics. After data cleaning and processing, the number of participants remained unchanged at 52.

Materials

a) Questionnaires

Pre-Course Questionnaire: Self-Efficacy in Online Learning. Before the start of the course, a pre-course questionnaire was administered to gather data on students' self-efficacy in online learning environments. This questionnaire, consisting of eight items, was distributed prior to the two-session online genetics course in order to assess each student's perceived ability to navigate and successfully complete an online learning experience. The self-efficacy items were adapted from the Self-Efficacy Questionnaire for Online Learning (SeQoL; Tsai et al., 2020) and were tailored to fit the short-format, task-based structure of the course. In addition, the language was tailored to better help the students understand the items. The 5 items of the SeQoL were rated using a 5-point Likert scale, ranging from 1 = "Not confident at all" to 5 = "Very confident". The questionnaire comprised three subscales: student engagement, perceived learning, and course structure/organisation. The subscale student engagement comprises 1 item, while the subscale perceived learning entails 2 items. The subscale course structure/organisation comprises 2 items. An example item from the perceived learning subscale is: “I believe I can understand the main ideas in the lesson”.

A self-efficacy score for each student was calculated by averaging the responses within each subscale, followed by computing the overall mean score across all subscales. Higher scores indicated greater self-efficacy in the online learning context. These scores served as quantitative indicators of the construct of online learning self-efficacy. The internal consistency of the pre-course questionnaire subscales was evaluated by using Cronbach’s alpha. The overall scale demonstrated excellent reliability ($\alpha = 0.95$). The distinct subscales also showed strong reliability: perceived learning with $\alpha = 0.88$ and course structure/organisation with $\alpha = 0.88$. The student engagement subscale consisted of a single item and, thus, was not subject to internal consistency analysis. The first part of the questionnaire collects demographic information, while the second part focuses on students’ perceived self-efficacy in the context of online learning. A copy of the questionnaire with its items can be found in Appendix A.

Post-Course Questionnaire: Student Learning and Satisfaction in Online Learning Environments. Following the completion of the course, a post-course questionnaire was administered to assess the students’ perceptions of the online learning experience, the effectiveness of the course materials, and their overall satisfaction with the course. This questionnaire entails selected items from the Student Learning and Satisfaction in Online Learning Environments (SLS-OLE) survey (Gray & DiLoreto, 2016) and was adapted to suit the context of the short-format genetics course. Similar to the pre-course questionnaire, the post-course questionnaire consisted of two parts: Part 1 collected demographic data, including name, age and gender, while Part 2 focused on students' course perceptions and learning outcomes. All items in Part 2 were rated on a 5-point Likert scale ranging from 1 = "I don't agree at all" to 5 = "I completely agree". The final item was reverse scored to control for response bias.

The questionnaire focused on two key subscales: student satisfaction and perceived learning. The subscale student satisfaction comprises 3 items, while the subscale perceived learning entails 6 items. An example item from the students' satisfaction subscale is: "I am happy with my experience in this course". Student perception and satisfaction of learning for each student were calculated by calculating mean scores for each subscale, as well as calculating an overall average score across all items. Higher scores indicated greater perceived learning and satisfaction, thus linking the questionnaire scores directly to the constructs of student satisfaction and perceived learning outlined in the introduction.

The internal consistency of the post-course questionnaire was also evaluated using Cronbach's alpha. The overall scale demonstrated excellent reliability ($\alpha = 0.91$), confirming the questionnaire's reliability in assessing students' perceptions of the online learning experience. Subscale reliabilities were strong, with student satisfaction showing $\alpha = 0.85$ and perceived learning $\alpha = 0.86$. All responses were collected anonymously and used solely for research and evaluation purposes. A copy of the questionnaire can be found in Appendix B.

b) Learning Environment

Go-lab Platform. Another key material used for this study is the Go-Lab platform, which enables the students to get familiar with the content through the use of interactive tasks. Two already existing Inquiry Learning Spaces (ILS) from previous studies were translated into German, and both lessons focused on the topic of inheritance. Students accessed this platform via a provided link. The Go-Lab platform was selected for its proven ability to support inquiry-based learning through interactive and scaffolded tasks. To maintain focus on the content and avoid distractions, the chat function was excluded due to previous findings showing it could confuse students unfamiliar with its use. Concepts of the ILS included dominant and recessive traits, homozygous and heterozygous genotypes, the scientific method, and the Punnet Square, etc. Each student had to complete two courses. For both classes, a double lesson was planned, consisting of 2 x 45 minutes.

Lesson 1: Introduction to Genetic Principles. The first lesson supported students in developing a foundational understanding of key concepts in genetics. Using interactive elements, such as multiple-choice questions, labelling tasks, and short answer questions (see Appendix E for full lesson details), students were guided through phases of inquiry. For instance, a multiple-choice quiz challenged their knowledge of traits and phenotypes (see Figure 1), while a labelling activity required them to identify components of a cell, such as genes and chromosomes. Open-ended questions, like “How can an organism with a recessive allele still show a dominant trait?” encouraged the students to apply their understanding to explain real-world scenarios. Throughout the lesson, learning was reinforced through a combination of videos, interactive quizzes, and open-ended response tasks designed to assess conceptual understanding.

Lesson 2: Application and Hypothesis Testing. The second lesson built directly upon the foundational knowledge developed in Lesson 1 and focused on applying concepts through guided inquiry, hypothesis testing, and data interpretation. Using a drag-and-drop activity, students formulated hypotheses about potential outcomes of crosses between two rabbits (see

Figure 2) and then tested these hypotheses against simulated data. They filled in observation tables and answered questions reflecting on their confidence in their hypotheses. This inquiry approach led to a Punnet Square activity, where students predicted genotypic ratios and percentages of their crosses' offspring. By the end of the second course, students had not only deepened their understanding of inheritance but also practised applying scientific reasoning and data interpretation in a structured, inquiry-based setting.

Figure 1

Multiple Choice Quiz about Traits and Phenotypes

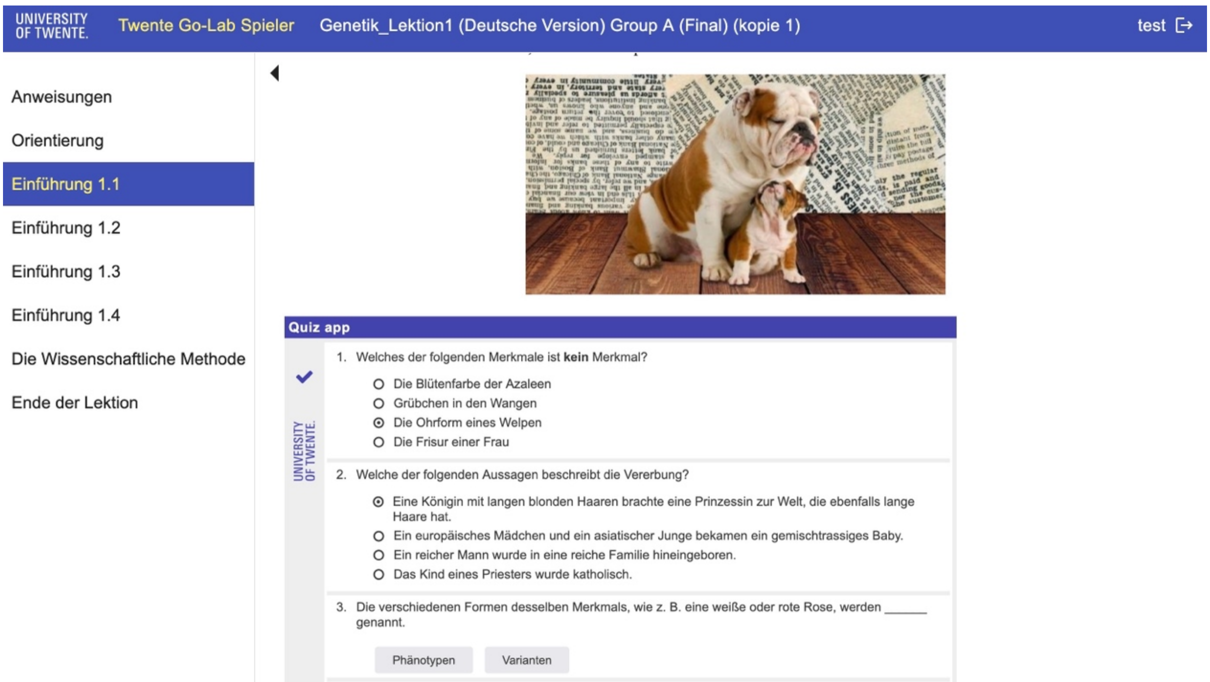
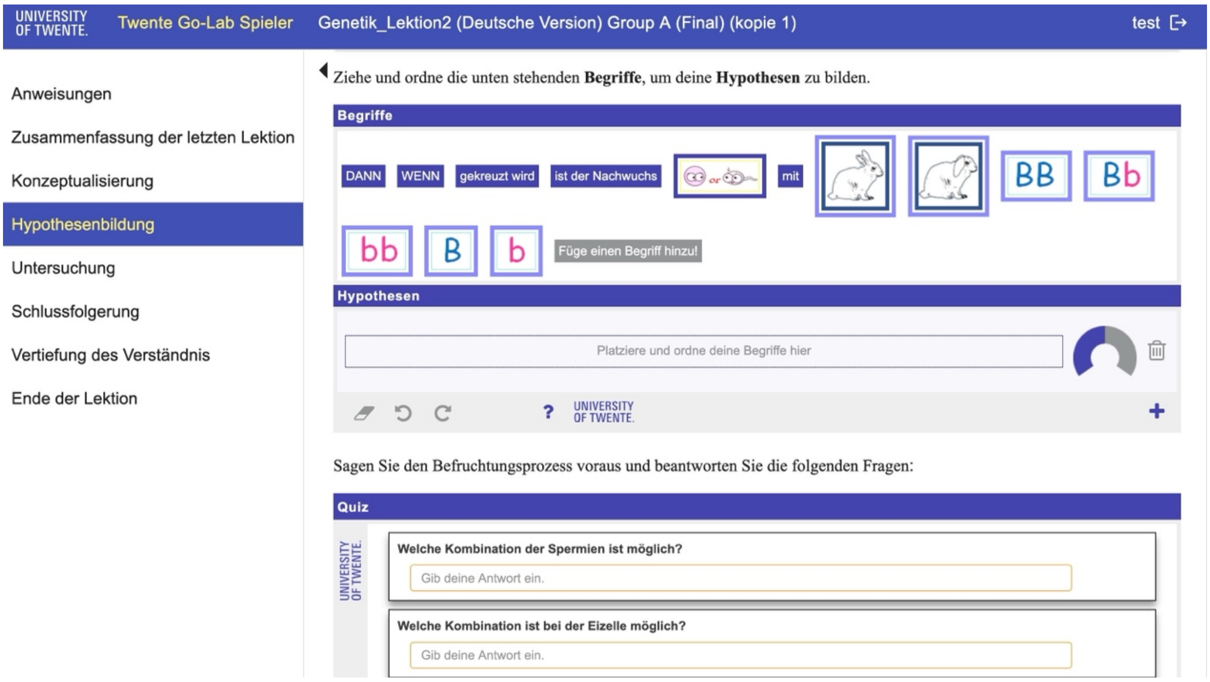


Figure 2

Drag-and-Drop Activity to Build a Hypothesis

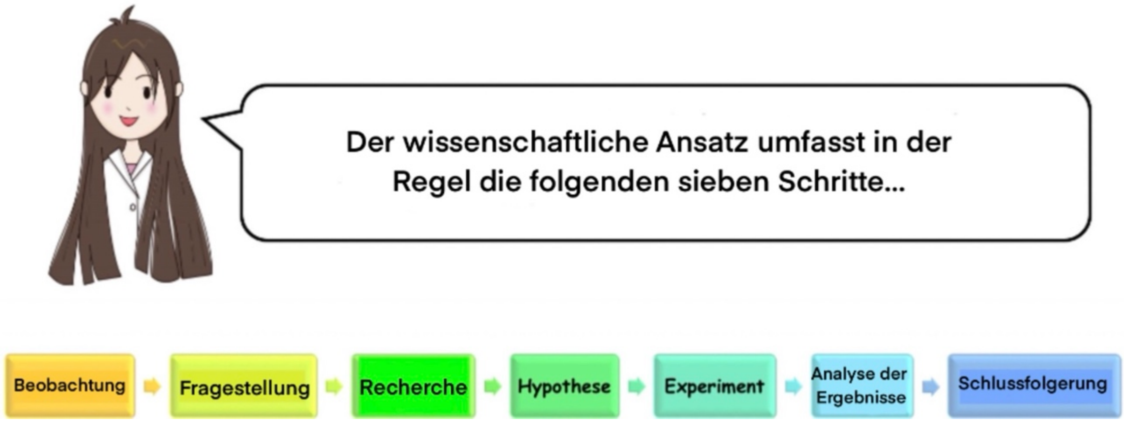


c) Virtual Character Implementation

As already outlined in the introduction, the presence of a virtual character served as the independent variable. A static virtual character was chosen to provide a sense of social presence, fostering learner engagement without adding the cognitive load associated with highly animated or interactive agents (Clarebout & Heidig, 2012). This design balances learner support and focus by offering guidance without distracting from the learning material. In the experimental condition, students interacted with this static virtual character embedded within the Go-Lab interface. This character appeared throughout the lessons (see Figure 3). In the control condition, the same learning content was presented without any virtual character. All other instructional elements and materials were identical across conditions. A more detailed description of this implementation can be found in the Procedure section.

Figure 3

Static Virtual Character Used in the Experimental Condition



Procedure

The study follows a four-step procedure. A few days prior to the experiment, informed consent is obtained from both the students and their parents or guardians. Only students with full consent were permitted to participate in the study. On the day of the experiment, the participating students were informed about the study's procedure, and afterwards they started with the study by completing the pre-course questionnaire (about 10-15 minutes) that included demographic questions and items measuring their perceived self-efficacy in online learning environments.

Once the pre-course questionnaire was completed, students engaged with two Inquiry Learning Space lessons on the topic of inheritance, presented through the Go-lab platform. The two participating classes were each assigned to a different condition: one class completed a learning environment featuring character-based instructions, while the other class worked through a version using only pure-text instructions. This class-based grouping allows for controlled group comparisons central to the study's research question. To illustrate the differences between these two instructional designs, screenshots of the respective learning environment for the distinct groups are provided below (see Figures 4 and 5). During the activity, students were engaged and comfortable using the laptops. However, in the participating school, technology is already integrated extensively (e.g., smart boards, a computer room, and mobile MacBook carts).

The Go-Lab platform monitored the students' interaction with the tasks, including task completion and time spent in each phase. At the end of the session, students filled out the post-course questionnaire (about 10-15 minutes), indicating their experiences with the learning environment and how supported and involved they felt with the content. The whole procedure was completed within a single class period of about 90 minutes. After the data collection stage, both questionnaires and the log files undergo preprocessing and cleaning, which includes handling missing data and eliminating sessions that are not complete. To ascertain whether there are notable variations between the two groups in terms of self-efficacy, completing the learning environment and their overall experiences with the platform, comparative analyses (such as t-tests) will be carried out.

Figure 4

Learning Environment with Balloon-Based Virtual Characters

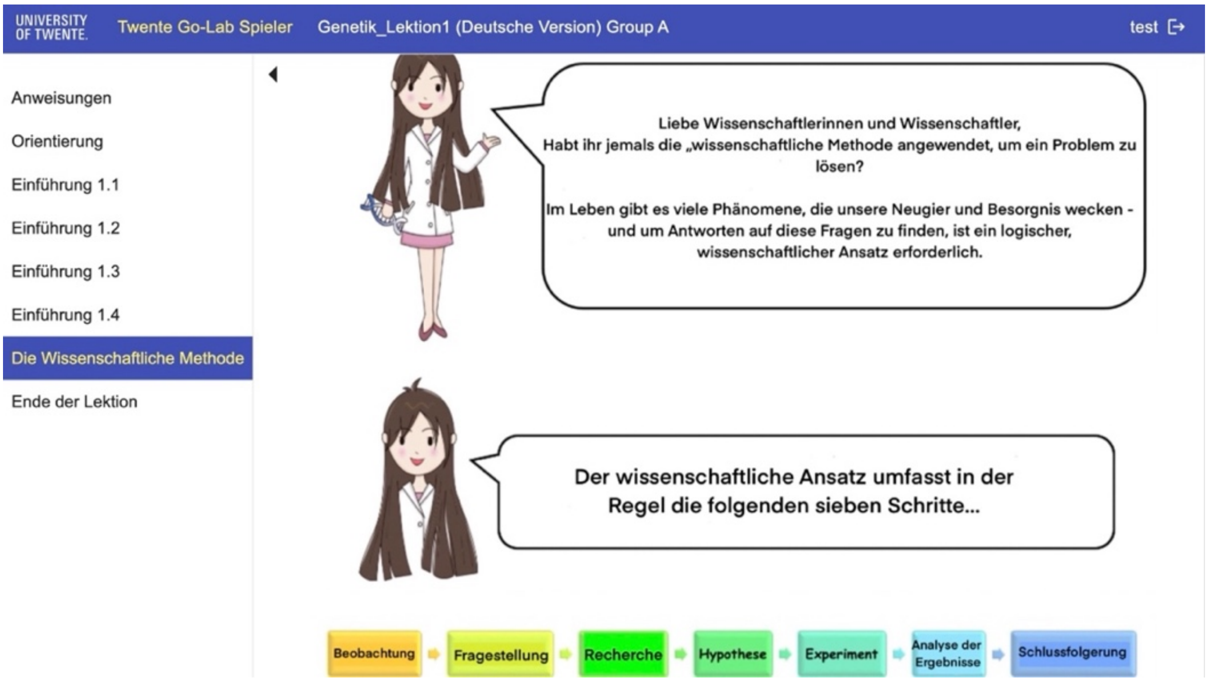


Figure 5

Learning Environment with Pure-Text Instructions

Data Analysis Plan

Scoring of Quiz Assignments. Throughout the lessons, students completed several quizzes that assessed their understanding of core genetics concepts. These quizzes included both open-ended and multiple-choice questions. Open-ended items were not automatically scored, and despite detailed and accurate student answers, the recorded average score was 0. In contrast, the subsequent quizzes primarily consisted of automatically scored multiple-choice items. In these cases, each multiple-choice question was worth 1 point for a correct answer and 0 points for an incorrect answer. The total score for each student was then calculated by adding up all points. These scores reflect the number of correct responses per student and were used to gauge students' grasp of key content areas in the learning sequence. The scores were derived from log files generated by the Go Lab environment, which recorded students' answers and interactions in real time.

Processing and Cleaning Engagement Data from the Learning Environment. Engagement time was defined as the total amount of time students were actively interacting with the tasks within the Go-Lab environment during each course. The Go-Lab platform automatically measured this by recording when a student was clicking or navigating between activities, in other words, when the student directly engaged with the instructional content. These intervals were summed to produce a total engagement time in each phase for each student. To identify unusually low engagement, the mean and standard deviation of engagement times were calculated for each course. Students with engagement times below two standard deviations from the mean were flagged as potential outliers. However, no student fell below this cutoff in the dataset, so no exclusions were necessary.

A cutoff for “engaged” students was established using the median of engagement time within the data. Students with engagement times equal to or exceeding the median were classified as engaged. For course 1, this cutoff was 1386 seconds (23.1 minutes), and for course 2, it was 1352 seconds (22.53 minutes). This classification allowed focused analysis on students demonstrating higher levels of engagement within the learning environment.

Data Analysis

RStudio (Version 2025.05.0) was used to perform the analysis. The CSV files of the two questionnaires, as well as the CSV files of all four courses, were downloaded and imported. All files were merged and formatted to produce the final dataset.

First, descriptive statistics were calculated, including means, standard deviations and ranges. This was done for all important demographic variables, such as gender and age, to provide an overview of the sample. Before conducting any inferential tests, an assumption check was made. Normality of the data was assessed both statistically using the Shapiro-Wilk test and visually through Q-Q plots to support the interpretation of the distribution patterns. Homogeneity of variance between the instructional groups was tested using Levene's test, while independence of observations was ensured through the study's between-subjects design, as students completed the course individually and were assigned to conditions based on their class. The assumption of linearity was not assessed, as it is only relevant for correlation or regression analyses involving continuous variables. Since this study compared group means using t-tests and non-parametric alternatives, linearity is not applicable.

After confirming that the data did not meet the normality assumption, the nonparametric Mann-Whitney U test was performed to compare group means across all subscales. These comparisons aimed to detect statistically significant differences between the group that received character-based instructions and the group that worked with pure-text instructions. At the end of the analysis, box plots were created to visually see the difference between the character-based condition and the text-based condition in all subscales.

Results

Descriptive Statistics

Table 1 summarises the descriptive statistics for each instructional group across seven subscales related to pre-course expectations, engagement with the learning environment, and post-course experience. Overall, the groups were comparable across most dimensions. In terms of pre-course variables, the text-based condition demonstrated slightly higher pre-engagement, pre-perceived learning, and course structure clarity scores. For the learning environment measures, the text-based condition spent more time in the environment (M = 136 sec, SD = 63.40) than the character-based condition (M = 109 sec, SD = 51.80), but the character-based condition rated the accuracy of the environment slightly higher. Additionally, the character-based condition reported slightly lower post-course satisfaction (M = 3.81, SD = 0.98) compared to the text-based condition (M = 4.17, SD = 0.56). Similarly, post-perceived learning scores were comparable across groups, with marginally higher scores in the text-based condition. These descriptive results indicate small but potentially meaningful differences in how each group engaged with the instructional material and perceived their learning outcomes.

Table 1

Descriptive Statistics by Subscale (*N* = 52)

Variable	Character-based (*M*)	Character-based (*SD*)	Text-based (*M*)	Text-based (*SD*)
Pre Engagement	3.52	1.25	3.88	0.88
Pre Perceived Learning	3.43	1.12	3.68	0.97
Pre Course Structure	3.70	1.12	4.06	0.81
LE Accuracy	1.21	0.33	1.03	0.33
LE Time Spent (sec)	109.00	51.80	136.00	63.40
Post Satisfaction	3.81	0.98	4.17	0.56
Post Perceived Learning	3.65	0.83	4.13	0.49

Note. *LE* = Learning Environment.

Engagement time in the learning environment (LE Time Spent) showed substantial variability within groups (see Table 2). For the character-based group, time spent ranged from 25.9 seconds to 206.0 seconds (M = 109.0, SD = 51.8), while for the text-based group, the range was 29.1 to 214.0 seconds (M = 136.0, SD = 63.4). These minimum and maximum values highlight the spread of engagement durations among the students.

Table 2

Engagement Time in Learning Environment by Group: Means, Standard Deviations, and Range (*N* = 52)

Group	*M* (sec)	*SD* (sec)	Min (sec)	Max (sec)
Character-based	109.00	51.80	25.90	206.00
Text-based	136.00	63.40	29.10	214.00

Note. Time spent (in seconds) in the learning environment.

Assumptions Check

Normality

Before conducting group comparisons, the assumption of normality was assessed using the Shapiro-Wilk test for each subscale score. Results indicated that all subscales significantly deviated from a normal distribution (all $p < .05$), suggesting violations of the normality assumption (see Table 3). Q-Q plots were also generated to visually support these findings. The representative Q-Q plots for these scales are included in Appendix C (see Figures C1–C7). Given these findings, non-parametric tests were chosen for subsequent group comparisons.

Table 3

Shapiro-Wilk Normality Test Results by Subscale (N = 52)

Subscale	***W***	***p*-value**
Pre Engagement	.87	<.001*
Pre Perceived Learning	.93	.003*
Pre Course Structure	.91	<.001*
LE Accuracy	.93	.003*
LE Time Spent	.92	.001*
Post Satisfaction	.92	.002*
Post Perceived Learning	.95	.029*

Note. *p < .05. All subscales significantly deviated from normality.

Homogeneity of Variance

To assess the assumption of equal variances, Levene's test was conducted for each subscale. The results are summarised in Table 4. The results indicated a significant violation for the subscales Post Satisfaction, $F(1, 50) = 6.30$, $p = .015$, and Post Perceived Learning, $F(1, 50) = 8.76$, p = .005, suggesting unequal variances. All other subscales did not show significant differences in variance ($p > .05$), indicating that the homogeneity of variance assumption was met for these subscales.

Table 4

Levene's Test for Homogeneity of Variance by Subscale ($N = 52$)

Subscale	***F*-value (df)**	***p*-value**
Pre Engagement	3.14 (1, 50)	.083
Pre Perceived Learning	0.28 (1, 50)	.600
Pre Course Structure	1.84 (1, 50)	.182
LE Accuracy	.02 (1, 50)	.889
LE Time Spent	.39 (1, 50)	.535
Post Satisfaction	6.30 (1, 50)	.015*
Post Perceived Learning	8.76 (1, 50)	.005*

Note. $p < .05$ indicates a significant violation of the homogeneity of variance assumptions.

Group Comparisons

Since the Shapiro-Wilk tests revealed violations of normality for all subscales, and homogeneity of variance was violated for one subscale, nonparametric Mann-Whitney U tests were conducted for all subscales to compare the character-based condition and text-based condition. Results indicated significant differences for the subscale LE Accuracy, $U = 447.0$, $p = .045$, $r = .28$, and LE Time Spent, $U = 223.0$, $p = .037$, $r = -.29$, with the character-based condition scoring higher on accuracy and spending less time on the learning environment compared to the text-based condition (see Table 5). No significant group differences were found for the subscales Post Satisfaction, Post Perceived Learning, Pre Engagement, Pre Perceived Learning, or Pre Course Structure (all $p > .05$). Effect sizes ranged from small to medium.

To visualise the group differences found in LE Accuracy and LE Time Spent, box plots were created, providing a clear depiction of the distribution and central tendency for each group. The box plots can be found in Appendix D (see Figures D1 and D2).

For full details, see Table 5, which summarises the Mann-Whitney U test statistic, p-value, and effect sizes for all subscales.

Table 5

Mann-Whitney U Test Results by Group ($N = 52$)

Subscale	***W***	***p***	***r***
Pre Engagement	285	.321	-0.13
Pre Perceived Learning	296	.447	-0.10
Pre Course Structure	281	.296	-0.14
LE Accuracy	447	.045*	0.28
LE Time Spent	223	.037*	-0.29
Post Satisfaction	286	.345	-0.13
Post Perceived Learning	235	.060	-0.26

Note. *W* = Mann-Whitney U test statistic; *p* = p-value; *r* = effect size (rank-biserial correlation). Effect size interpretation: small = 0.1, medium = 0.3, large = 0.5.

Discussion

Recap of Main Findings

This research explored whether adding a static, cartoon-style female character in an inquiry-based learning environment would positively influence students' perceived learning and satisfaction. Surprisingly, adding the character did not lead to any noticeable improvements in students' satisfaction or perceived learning after the course. This suggests that a static virtual agent may not meaningfully influence students' subjective learning experience.

The only statistically significant differences observed were related to students' interaction with the learning environment itself. The character-based condition answered more tasks correctly, while the text-based condition spent a longer time in the environment. However, these differences did not lead to higher satisfaction or perceived learning outcomes. This points to a potential mismatch between performance and satisfaction, where students in the character-based condition performed better in terms of accuracy, yet did not report increased satisfaction. This suggests that doing well does not always translate to feeling more satisfied. This disconnect may be partially explained by a lack of emotional engagement, despite behavioural engagement being present (Fredricks et al., 2004). However, it is also important to consider that the character's specific design, such as its static nature, cartoon style, and gender representation, could have influenced how students interacted with it and perceived its usefulness. This raises questions about the age appropriateness of the design, while cartoon characters may appeal to younger learners, middle school students might perceive them as over simplistic.

Another reason for these results may be that the static virtual character lacked interactivity and adaptive scaffolding, which prior research has shown to be important for engagement and meaningful learning support in complex subjects (Clarebout & Heidig, 2012; Renninger et al., 2020; Topçu & Şahin-pekmez, 2009). The lack of motivational support and autonomy-enhancing features (Deci & Ryan, 2000) may have limited students' sense of ownership over the task, reducing intrinsic motivation. Still, it is worth mentioning that these results need to be interpreted with caution, since it is not clear whether one group already started with more knowledge about the topic than the other. Prior knowledge can significantly affect students' perceptions of learning, satisfaction, and actual performance. As discussed in the introduction, complex genetic concepts involve abstract and often inaccessible ideas such as molecular interactions (Topçu & Şahin-pekmez, 2009), which students with more background knowledge may already grasp more easily. These students may feel more confident navigating the content and report higher satisfaction or perceived learning, not necessarily due to the virtual character, but because of their existing understanding. Thus, unequal prior knowledge could have influenced both the subjective and objective outcomes observed.

An important point to consider here is that while performance outcomes were different, the subjective experiences of the learners were not aligned, suggesting that the static virtual character did not meaningfully influence students' affective or motivational states. This misalignment may indicate that the character failed to stimulate deeper levels of engagement, particularly emotional and cognitive engagement, which are known to be important for sustained learning. The current study suggests that middle school students engaged in complex inquiry learning may require more sophisticated virtual agent designs than simple static characters can provide.

Overall, these results do not support the initial stated hypothesis that the presence of a static character would enhance student perceptions of learning and satisfaction. Instead, findings suggest that a static, non-interactive character, particularly one lacking personalisation, interactivity, and responsiveness, may be insufficient as a form of scaffolding in complex science learning concepts like genetics. While the character may have to some extent affected learners' accuracy in task completion, it did not significantly impact the core learning outcomes, perceived learning and satisfaction targeted in this study. So, the original hypothesis of the study could not be supported.

Interpretation of (Low) Engagement

Building on these considerations, it is also important to reflect on how students engaged with the learning environment, as patterns of engagement may help explain why the virtual character had a limited impact on perceived learning and satisfaction. Although engagement was not part of the research question, behavioural indicators such as time spent in the learning environment and task accuracy offer indirect insights into how engaged the students were. Linking back to the findings, students in the character-based condition were more accurate but not more satisfied, indicating that their behavioural engagement may not have been accompanied by emotional involvement. Students in the character-based condition did not spend much time in the learning environment or show clear learning improvements, which suggests they may not have been fully engaged with the task. Possible reasons for this outcome can be grouped into three themes: learner-related factors, the design of the virtual character, and the classroom environment.

One of the learner-related factors is that students in this age group often struggle with motivation and get distracted easily, especially with school tasks that feel forced or unrelated to their lives (Ryan & Deci, 2000). Since the topic was genetics, a subject that can seem abstract and hard to relate to (Topçu & Şahin-pekmez, 2009), students, regardless of the condition, might not have felt connected to it, which lowered their interest and concentration. This highlights the importance of aligning content with students' developmental stage and intrinsic goals to promote self-determined motivation. This reduced engagement could explain why perceived learning and satisfaction scores were low, as students may not have felt that the learning experience was meaningful or rewarding.

Another possible factor is the design of the virtual agent itself. It was a simple, cartoon-style character with little interactivity. For older students, this kind of design might come across as unhelpful or childish. This notion arises from research suggesting that teenagers tend to connect more with digital characters who look or act like them (Steinke et al., 2012). So, using a female character may not have worked for all students, especially boys or those who didn't relate to the virtual character's look or tone. This lack of personal relevance and interactivity could have contributed to the absence of positive effects on perceived learning and satisfaction, as students may have viewed the character as decorative rather than supportive. Personalisation and identification are critical here. Without these, students may not emotionally connect to the agent. Future studies should thus assess whether gender representation and personalisation of the agent play a role in influencing how different student groups perceive and interact with virtual agents.

Finally, the classroom setting could also play a part. Although students only used the tool once, they were generally excited to work on laptops and welcomed the opportunity to do something different from their regular classroom routine. The sessions were supervised, which helped maintain order and kept students from engaging in off-task behaviours like chatting or digital multitasking. From an engagement theory perspective, however, this suggests that students may have been behaviourally engaged, since they followed the instructions and stayed on task, but were not necessarily emotionally or cognitively engaged (Fredricks et al., 2004). This type of engagement is often shallow and may not reflect real learning or sustained interest. This distinction might also help explain the intriguing gap observed between the two groups. Despite higher task accuracy in the character-based condition, these students did not report greater satisfaction. As mentioned at the beginning of the discussion, this reflects a disconnect between objective performance and subjective experiences, where behavioural indicators suggest engagement, but subjective reports do not reflect positive experiences. This suggests that while they were behaviourally engaged enough to perform well, the lack of emotional and cognitive engagement possibly limited their positive perception of the learning experience. In other words, doing well on tasks, in this case, did not equate to feeling satisfied. While the novelty of the activity and the use of technology sparked initial interest, that excitement may not have translated into deeper cognitive engagement. For instance, students could have played with the simulation, doing trial and error without a clear plan, simply testing different options to see what would happen. This pattern reflects what is often called the “novelty effect”, where new tools or formats briefly capture attention but do not lead to lasting involvement (Miguel-Alonso et al., 2024). So, while the virtual agent may have added some visual appeal, it likely was not enough to sustain meaningful learning engagement on its own.

Meaningful Interpretation of Inconclusive Results

Taken together, these engagement-related insights help contextualise the study’s largely inconclusive results, pointing toward deeper design and alignment issues that may explain why the virtual character had limited impact. The small differences between the character-based and text-based groups suggest that adding a static, cartoon-style persona did not really change how students felt about their learning or how satisfied they were. Its lack of interactivity likely limited engagement, as it could not offer the kind of adaptive support necessary for complex learning tasks. It is possible that students saw the character as decorative rather than helpful, which limited its effect. From a scaffolding perspective, this suggests that the virtual character may not have provided the adaptive or contextual support students needed. Unlike more responsive forms of scaffolding that adjust guidance based on learners' immediate needs (Renninger et al., 2020), this static agent lacked interactivity and feedback necessary to offer meaningful assistance, limiting its effectiveness as a learning support. Given the abstract nature of genetics (Topçu & Şahin-pekmez, 2009), students likely needed more tailored support than a static character could provide. From a cognitive load theory perspective (Sweller, 1988), the static character may not have matched students' expectations for virtual agents, possibly increasing cognitive load. If students expected interactive support but encountered only a simple cartoon figure, this mismatch could have distracted or confused them, detracting from learning (Fredricks et al., 2004). These findings align with previous research indicating that the positive effects of virtual agents on learning (Clarebout & Heidig, 2012; Gulz, 2004; Zhang & Whu, 2024) and engagement are contingent on specific design features such as adaptivity, interactivity, and contextual responsiveness, features that the static character did not possess. This underscores the importance of designing virtual agents not just as visual additions but as dynamic scaffolds tailored to learners' evolving needs.

Even though the results weren't statistically significant, this study still adds valuable insights to the research on digital agents in education, especially when it comes to understanding when and for whom these tools work. While earlier studies have shown that virtual characters can support learning and engagement (e.g., Gulz, 2004; Zhang & Whu, 2024), the current findings suggest that this may not always apply in every context. In this case, the results show that simply adding a non-interactive, cartoon-style female character to an online inquiry-based learning environment does not automatically ensure higher satisfaction or perceived learning in secondary school students. This highlights the importance of considering not just whether to use virtual agents, but how they are designed and integrated into the learning process. Additionally, when considered in combination with the behavioural data (e.g., task accuracy and time on task), the self-reported data adds depth to this interpretation, revealing a complex picture of performance versus satisfaction. This interaction between log data and subjective reports offers a meaningful contribution despite a lack of significance. Future research should build on this by systematically combining behavioural log data, subjective self-reports, and objective performance measures to offer a good understanding of learning with digital agents.

Limitations and strengths

While these findings offer important implications for the design and implementation of virtual agents, they should be interpreted in light of several limitations that may have influenced the outcomes. One notable limitation was the absence of a pre- or post-test to directly measure learning progress. Without that kind of data, it is hard to say for sure whether students improved or if one group simply started out with more background knowledge than the other, which could have affected the results. The study also did not take into account whether students had previous experience with digital agents, which might have influenced how they reacted to the character. However, this factor was not prioritised, as current research provides limited evidence that prior exposure plays a major role.

From a data perspective, the relatively small sample size (52 students) limits how widely the findings can be applied. And since the study was carried out during a single session, it wasn't possible to look at any long-term effects, like whether engagement or perceptions changed over time. These limitations mostly reflect the practical realities of working within the constraints of a school schedule, where extended testing isn't always feasible. In this way, the real classroom setting, while valuable for authenticity, also introduced constraints that shaped the study's design.

That said, the lesson topic was reviewed and approved by the classroom teacher, which helped ensure that students saw the activity as relevant and meaningful within the context of their curriculum.

Despite some limitations, this study also has several notable strengths. First, it took place in a real classroom setting, which makes the findings more relevant to everyday learning environments. Instead of an artificial lab setting, this research shows how students actually interact with educational technology during normal school activities. Second, the study used well-established self-report questionnaires to measure students' self-efficacy, perceptions, and satisfaction, combined with careful tracking of engagement. This mix of subjective and behavioural data makes the results more trustworthy. Finally, the study offers valuable insights into a design feature – static virtual agents – that is commonly used in educational technology but hasn't been studied enough. By exploring how these agents work in a real school context, this research helps us better understand their benefits and limits.

Implications for Research and Practice

Despite these limitations, the findings offer several practical and research-oriented implications for educators and designers. From a design perspective, this study reinforces that virtual characters must be thoughtfully designed to offer adaptive or interactive support rather than serving as passive visual elements. Developers and instructional designers should carefully evaluate whether such elements provide actual support or merely serve as decoration. Time and resources might be better spent on characters that offer adaptive, interactive, or personalised feedback.

Another takeaway is that the self-report questionnaires used to measure self-efficacy, learning perception, and satisfaction worked well in a real classroom setting with middle school students aged 11 to 13. The tools provided consistent and meaningful data, even in a small-scale study, reinforcing their value for future research and classroom-based evaluations.

Future Directions

Recognising these implications invites further exploration into how digital agents can be better designed and integrated. Future studies could build on these findings by taking a closer look at how different features of agent design impact student engagement and learning. For example, future research might compare different types of virtual agents – realistic vs. cartoon-style, male vs. female, static vs. animated, or with and without voice – to find out which combinations work best for different types of learners. It would also be helpful to examine how factors like timing, tone, or how personalised the agent's feedback is might affect students' reactions.

Another interesting direction would be to study younger students, who might respond differently to digital agents. It would also be useful to explore how repeated exposure to the same agent over multiple sessions influences engagement, helping to tell the difference between short-term novelty effects and more lasting impact. Using richer outcome measures, like physiological responses or emotional connection with the agent, could also give deeper insights into the learning process.

To be clear, this study doesn't claim that cartoon-style agents are ineffective in general. Rather, it shows that their success depends on how they are designed, what role they play in the learning experience, and who the target learners are.

Conclusion

This study investigated whether adding a static, cartoon-style virtual agent to an inquiry-based learning environment could positively influence secondary students' perceived learning and satisfaction. The results showed no significant differences in the text-based and character-based conditions, suggesting that the mere presence of a virtual character, without interactivity or personalisation, may not be sufficient to support student learning.

While the real classroom setting ensured authenticity and practical relevance, it also introduced limitations such as time constraints, lack of pre-/post-testing, and a relatively small sample size. These factors should be considered when interpreting the results. However, the study still offers valuable insights, especially by demonstrating that the design and alignment of virtual agents with learner needs is valuable.

Despite non-significant findings, this study contributes to a growing body of research that should further investigate when, how, and for whom virtual agents are effective. Future work should explore more dynamic, interactive characters and investigate their impact across different age groups, subject areas, and time frames.

AI Statement

During the preparation of this work, the author used Chat GPT in order to interpret the results of the outputs of the data analysis. Microsoft Word and Grammarly were used to check grammar and spelling. After using this tool/service, the author reviewed and edited the content as needed and takes the full responsibility for the content of the work.

Appendix A

Pre-Course Questionnaire Self-Efficacy in Online Learning

This 8-item questionnaire was administered before the two-session online genetics course to evaluate each student's perceived self-efficacy in navigating and completing an online learning experience. The self-efficacy questions were adapted from the Self-Efficacy Questionnaire for Online Learning (SeQoL; Tsai et al., 2020) and tailored to the short-format, task-based structure of the course.

Part 1: Demographic Information

1. What is your name?
2. What is your age?
3. What is your gender?
   ☐ Male ☐ Female ☐ Prefer not to say ☐ Other: _______

Part 2: Online Learning Self-Efficacy

1 = Not confident at all, 2 = A little confident, 3 = Kind of confident, 4 = Mostly confident, 5 = Very confident

#	Statement	1 2 3 4 5
1	I feel confident that I can finish an online course.	
2	I believe I can understand the main ideas in the lesson.	
3	I feel confident I can do the quizzes after the lessons.	
4	I feel confident using the online platform to move through the lessons.	
5	I feel confident reading and following instructions on the screen during the lesson.	

Appendix B

Post-Course Questionnaire Student Course Perception

The following questionnaire was distributed at the end of the online course to collect students' measurement of general satisfaction, the effectiveness of course material and activities, and their perceived learning results. This instrument includes selected items from the validated **Student Learning and Satisfaction in Online Learning Environments (SLS-OLE)** survey (Gray & DiLoreto, 2016). The results will help evaluate the success of the course in meeting its learning objectives and supporting student development in an online environment.
Responses are anonymous and will only be used for research and evaluation purposes.

Part 1: Demographic Information

1. What is your name?
2. What is your age?
3. What is your gender?
   ☐ Male ☐ Female ☐ Prefer not to say ☐ Other: _______

Part 2: Post-Course Perceptions

Rating Scale:
1 = 1 = I don't agree at all, 2 = I don't agree much, 3 = I'm not sure, 4 = I mostly agree, 5 = I completely agree

#	Statement	1 2 3 4 5
1	I am happy with my experience in this course.	
2	I am happy with what I learned in this course.	
3	I am happy with the topics and lessons in the course.	
4	I am happy with how much I learned in the course.	
5	The tasks and activities helped me understand better.	
6	I learned things that will help me in the future.	
7	The activities helped me reach the learning goals.	
8	The course helped me grow and improve.	
9	I learned less than I thought I would. *(Reverse scored)*	

Appendix C

Figure C1

Q-Q plot for Pre-Engagement showing deviation from normality.

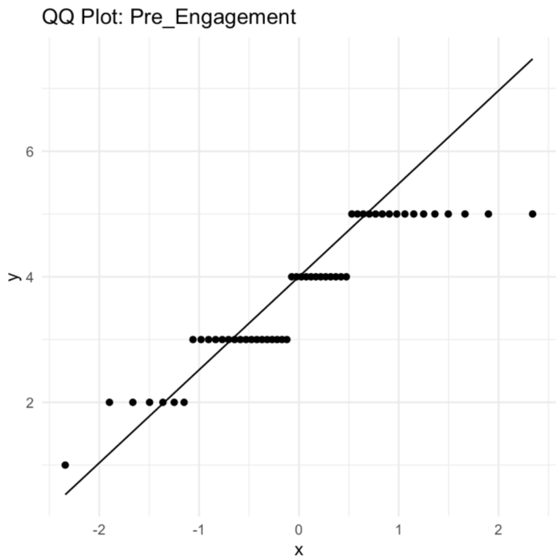


Figure C2

Q-Q plot for Pre-Perceived Learning showing deviation from normality.

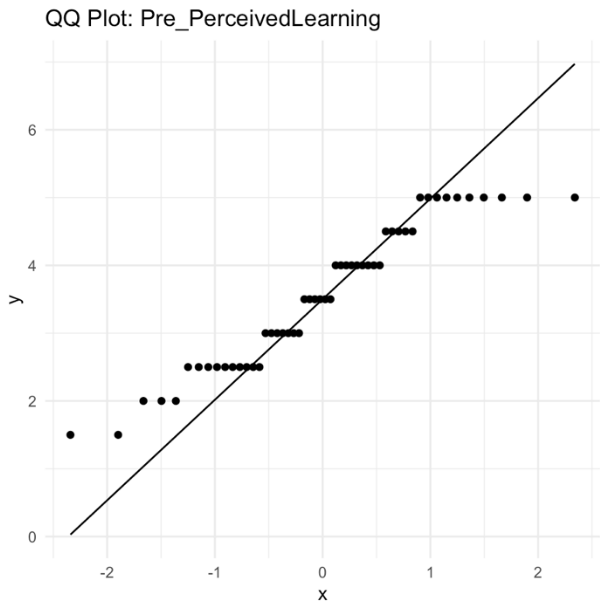


Figure C3

Q-Q plot for Pre Course Structure/Organisation showing deviation from normality.

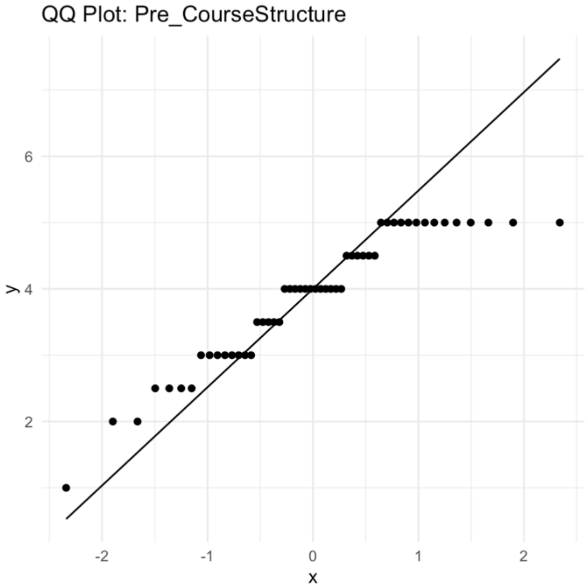


Figure C4

Q-Q plot for LE Accuracy showing deviation from normality.

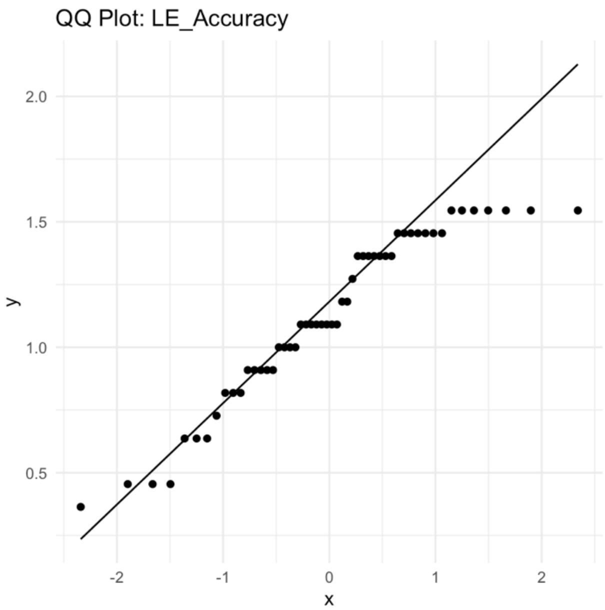


Figure C5

Q-Q plot for LE Time Spent showing deviation from normality.

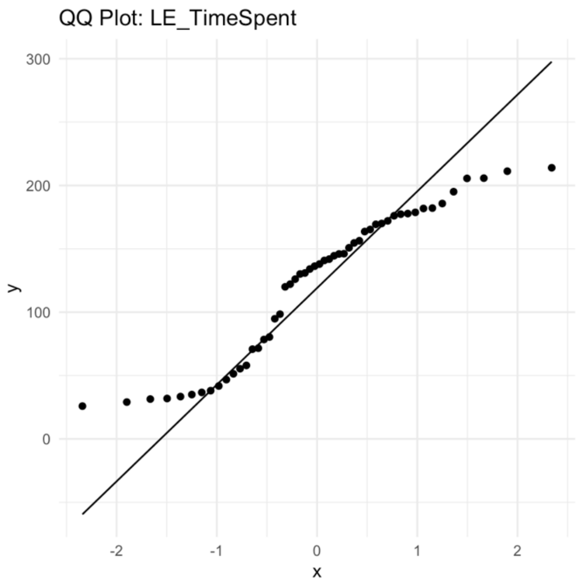


Figure C6

Q-Q plot for Post Satisfaction showing deviation from normality.

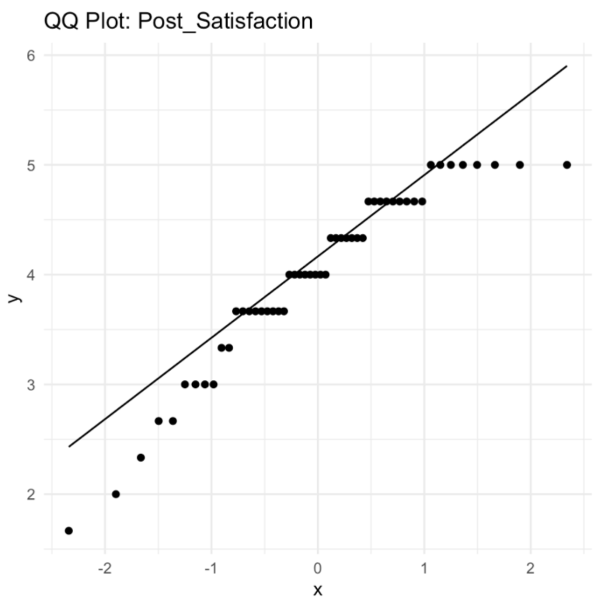
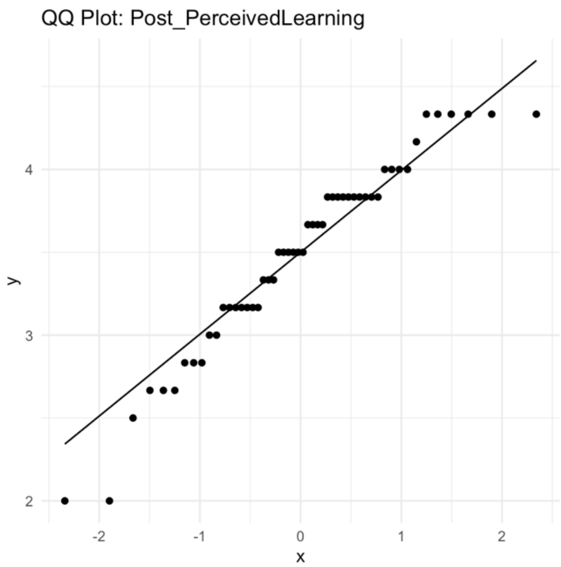


Figure C7

Q-Q plot for Post Perceived Learning showing deviation from normality.



Appendix D

Figure D1

LE Accuracy Score by Group

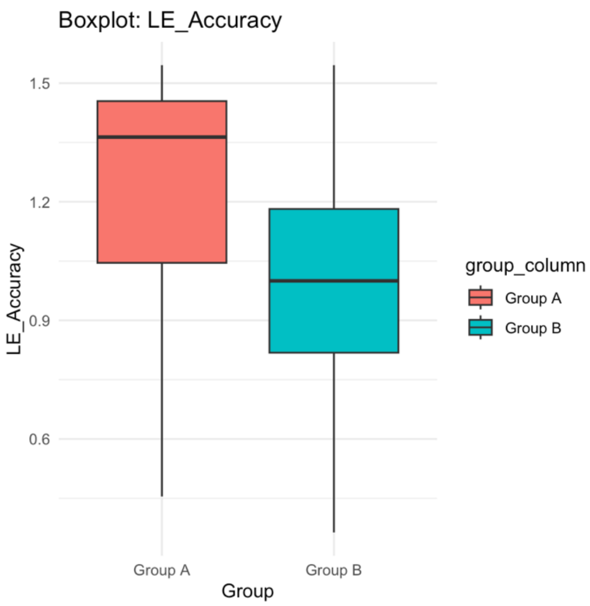
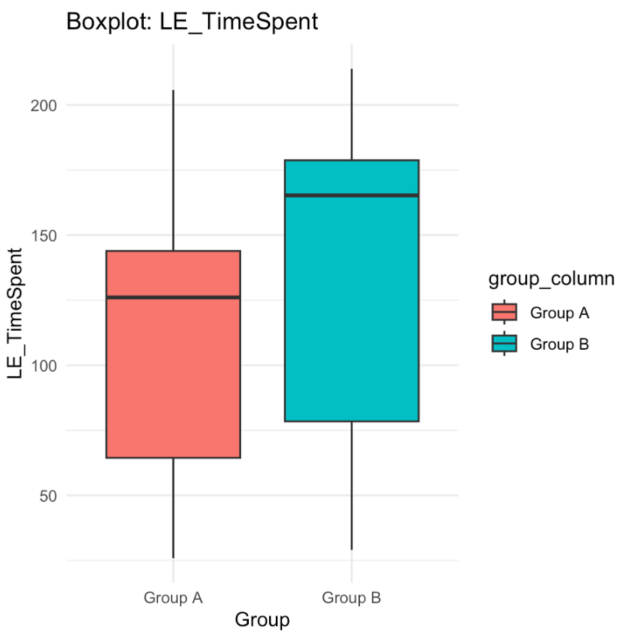


Figure D2

LE Time Spent Score by Group



Appendix E

Appendix E.1 – Lesson 1: Tasks and Quizzes

Orientation

In the observation station, numerous rabbits are kept for research and observation. The photo below shows some of these animals.

Task:
Carefully observe the rabbits. Which different traits can you identify?
Write your answers in the space provided below the picture.

Introduction 1.1

Which of the following is not a trait?

a) Flower colour of azaleas
b) Dimples in cheeks
c) A puppy's ear shape
d) A woman's hairstyle

Which of the following statements describes heredity?

a) A queen with long blonde hair gave birth to a princess who also has long hair.
b) A European girl and an Asian boy have a mixed-race baby.
c) A rich man was born into a rich family.
d) The child of a priest is a Catholic.

The different forms of the same trait, for example, a white or red rose, are called:

a) Phenotypes
b) Variants

The observable traits of a cat, for example, a spotted fur, are called:

a) Variants
b) Phenotypes

Introduction 1.2

The photo shows the flower colour of peas and their corresponding allele combinations.
Answer the following questions:

"AA, Aa, aa" are … of the flowers:

a) Genotypes
b) Phenotypes

The purple and white colours are … of the flower:

a) Genotypes
b) Phenotypes

Using the picture above:

The label A refers to:

a) Chromosome
b) Allele
c) Gene

The label B refers to:

a) Chromosome
b) Allele
c) Gene

The label C refers to:

a) Chromosome
b) Allele
c) Gene

Introduction 1.3

Question 1:
How can an organism with a recessive allele still show a dominant trait?
Explain with an example.
Indicate both the genotype and the phenotype in your answer.

Question 2:
How is a disease, such as Cystic fibrosis, which is recessive, inherited?

Question 3:
Explain why a recessive trait can remain hidden for generations but suddenly appear in a descendant.

Introduction 1.4

Question 1:
What does "purebred" mean?
How does it differ from a hybrid?

Question 2:
What are gametes, and why do they have only half the number of chromosomes?

Question 3:
Suppose you have a purebred tall flower (BB).
What would happen if you crossed it with a small flower (bb)?

The scientific method

To investigate and analyse questions related to "life phenomena", we use the scientific method.
The first four steps of the scientific method are:

A) Hypothesis formulation
B) Observation
C) Literature review
D) Experiment planning

What is the correct order?

a) B – C – A – D
b) A – B – C – D
c) C – B – A – D
d) D – C – B – A

When Tien-Tien travelled to Spain, she noticed that bullfighters often use red cloths to provoke the bulls. It seemed to her that the red cloth made the bulls more aggressive.
She asked herself:
"Why does the red cloth make the bulls aggressive?"
This question corresponds to which step of the scientific method?

a) Observation
b) Formulating a research question
c) Developing a testable hypothesis
d) Designing a procedure for investigation

The platypus is currently hard to find in the wild. This may be due to industrial water pollution. Scientists are analysing platypus blood samples to determine whether high lead concentrations are responsible for their decline.
The sentence in quotes corresponds to which step of the scientific method?

a) Searching for sources
b) Observation
c) Hypothesis formulation
d) Experiment

What should scientists do if the results of an experiment do not match their original hypothesis?

a) Adjust the results to match the hypothesis
b) Change the procedure to align with the hypothesis
c) Draw conclusions based on the results
d) Revise the hypothesis, plan a new experiment, and observe again

Appendix E.2 – Lesson 2: Tasks and Quizzes

Conceptualisation

Using the given scenario of rabbit ear shape (BB = upright ears, bb = lop ears), answer the following questions:

1. What are the genetic factors in the sperm and egg cells regarding ear shape?
2. What is the genotype of the offspring?
3. What is the phenotype of the offspring?

Forming Hypotheses

Using the Hypothesis Tool, form hypotheses by arranging the provided terms:

a) What are the genetic factors in the sperm and egg cells?

b) What is the genotype of the offspring?

c) What is the phenotype of the offspring?

Predict the Fertilisation Process

Answer the following questions based on your knowledge of gametes:

a) What is the possible combination of the sperm?

b) What is the possible combination of the egg?

c) What is the genotype of the offspring?

d) What is the phenotype of the offspring?

Investigation

Using the Collaborative Lab tool, observe breeding rabbits and complete the following:

a) Record observable traits of parent rabbits and their offspring in the provided table.

b) Determine which traits (ear shape, coat colour) are dominant or recessive.

Conclusion

Using the Conclusion Tool, answer the following questions:

a) What is the ear shape of the offspring?

b) What is the dominant trait for ear shape?

c) Justify your answers based on results from the Hypothesis Tool and Collaborative Lab.

The offspring have ______ ears.

a) Upright

b) Lop (hanging)

Explain your reasoning (open response).

What is the dominant trait for ear shape?

a) Upright ears

b) Lop (hanging) ears

Explain your reasoning (open response).

Deepening Understanding: Punnett Squares

Complete the Punnett Square table to determine genotype and phenotype probabilities of guinea pig offspring (Black = B dominant; Brown = b recessive):

a) Fill in the genotypes and corresponding phenotypes of the offspring.

b) Determine the ratio of genotypes among the offspring (BB : Bb : bb).

c) Determine the ratio of phenotypes among the offspring (Black : Brown).

d) Convert these ratios into percentages.

Which of the following statements is correct?

a) The probability of genotype BB in the offspring is 1/4.

b) There is no possibility of black guinea pigs among the offspring.

c) The offspring will have an equal number of black and white guinea pigs.

d) Genotype Bb will not appear among the offspring

References

Agyei, E. D. & Agyei, D. D. (2021). Enhancing Students' Learning of Physics Concepts with Simulation as an Instructional ICT Tool. *European Journal Of Interactive Multimedia And Education*, *2*(2), e02111. https://doi.org/10.30935/ejimed/11259

Brachten, F., Brünker, F., Frick, N. R. J., Ross, B. & Stieglitz, S. (2020). On the ability of virtual agents to decrease cognitive load: an experimental study. *Information Systems And e-Business Management*, *18*(2), 187–207. https://doi.org/10.1007/s10257-020-00471-7

Clarebout, G. & Heidig, S. (2012). Pedagogical agents. In *Springer eBooks* (S. 2567–2571). https://doi.org/10.1007/978-1-4419-1428-6_942

Ergül, E. & Koç, M. (2013). The Role of Animated Agents in Web-based Distance Education. *Procedia - Social And Behavioral Sciences*, *83*, 1016–1022. https://doi.org/10.1016/j.sbspro.2013.06.189

Fredricks, J. A., Blumenfeld, P. C. & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review Of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Gray, J. A. & DiLoreto, M. (2016). The Effects of Student Engagement, Student Satisfaction, and Perceived Learning in Online Learning Environments. In National Council of Professors of Educational Administration, *NCPEA International Journal Of Educational Leadership Preparation* (Bde. 1–1) [Journal-article]. https://files.eric.ed.gov/fulltext/EJ1103654.pdf

Gulz, A. (2004). Benefits of Virtual Characters in Computer Based Learning Environment: Claims and Evidence. *International Journal of Artificial Intelligence in Education*, *14*(3), 313-334

Kultusministerkonferenz. (2024). *Bildungsstandards für den Fachbereich Naturwissenschaften: Biologie, Chemie, Physik und ihre didaktische Bedeutung* [Brochure]. Kultusministerkonferenz. https://www.kmk.org/fileadmin/Dateien/pdf/Bildung/Qualitaet/ImplBroschuere_BiSta_NATURWISSENSCHAFTEN_2024-06-06.pdf

Lin, T., Hsu, Y., Lin, S., Changlai, M., Yang, K. & Lai, T. (2011). A REVIEW OF EMPIRICAL EVIDENCE ON SCAFFOLDING FOR SCIENCE EDUCATION. *International Journal Of Science And Mathematics Education*, *10*(2), 437–455. https://doi.org/10.1007/s10763-011-9322-z

Mamun, M. A. A. & Lawrie, G. (2023). Cognitive presence in learner–content interaction process: The role of scaffolding in online self-regulated learning environments. *Journal Of Computers in Education*, *11*(3), 791–821. https://doi.org/10.1007/s40692-023-00279-7

Miguel-Alonso, I., Checa, D., Guillen-Sanz, H. & Bustillo, A. (2024). Evaluation of the novelty effect in immersive Virtual Reality learning experiences. *Virtual Reality*, *28*(1). https://doi.org/10.1007/s10055-023-00926-5

Mussard, J., & Reiss, M. J. (2022). Why is genetics so hard to learn? Insights from examiner reports for 16-to 18-year-old in England. *The School science review*, *103*(384), 32-40.

Pedaste, M., Mäeots, M., Siiman, L. A., De Jong, T., Van Riesen, S. A., Kamp, E. T., Manoli, C. C., Zacharia, Z. C. & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14*, 47–61. https://doi.org/10.1016/j.edurev.2015.02.003

Renninger, K. A. & List, A. (2012). Scaffolding for Learning. *Encyclopedia of the Sciences* of Learning, 2922–2926. https://doi.org/10.1007/978-1-4419-1428-6_569

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Short, J., Williams, E., & Christie, B. (1976). The social psychology of telecommunications. Retrieved June 26, 2025, from https://cornerstone.lib.mnsu.edu/cgi/viewcontent.cgi?article=1001&context=isalt_resources

Steinke, J., Applegate, B., Lapinski, M. K., Ryan, L. & Long, M. (2012). Gender Differences in Adolescents' Wishful Identification With Scientist Characters on Television. *Science Communication*, *34*(2), 163–199. https://doi.org/10.1177/1075547011410250

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Topçu, M. S., & Şahin-pekmez, E. (2009). Turkish Middle School Students' Difficulties in Learning Genetics Concepts. *Journal of Turkish Science Education*, *6*(2).

Tsai, C., Cho, M., Marra, R. & Shen, D. (2020). The Self-Efficacy Questionnaire for Online Learning (SeQoL). *Distance Education*, *41*(4), 472–489. https://doi.org/10.1080/01587919.2020.1821604

Van De Pol, J., Mercer, N. & Volman, M. (2018). Scaffolding Student Understanding in Small-Group Work: Students' Uptake of Teacher Support in Subsequent Small-Group Interaction. *Journal Of The Learning Sciences*, *28*(2), 206–239. https://doi.org/10.1080/10508406.2018.1522258

Zhang, R. & Wu, Q. (2024). Impact of using virtual avatars in educational videos on user experience. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-56716-9