

Gamified approach to teaching introductory astronomy online

Daniel F. Barringer* and Julia D. Plummer

*Department of Curriculum & Instruction, The Pennsylvania State University,
141 Chambers Building, University Park, Pennsylvania 16802, USA*

Julia Kregenow and Christopher Palma

*Department of Astronomy & Astrophysics, The Pennsylvania State University,
525 Davey Lab, University Park, Pennsylvania 16802, USA*



(Received 15 May 2017; published 15 June 2018)

[This paper is part of the Focused Collection on Astronomy Education Research.] An increased interest in using video games in educational contexts has led to many innovations in both formal and informal environments. Educational researchers, instructional designers, and educators have sought and developed ways to incorporate video games or gamelike elements into a curriculum. We present a study of a web-based college-level introductory astronomy class in which content is delivered entirely in a digital world setting through nonplayer characters and built-in minigames. Based on prior research on the potential benefits of video games for use in education, we expect that topics covered in the minigames, the most gamified aspects of the course, should be particularly effective. Using data collected through pre- and post-testing of two sections of ASTRO 001V (the designation given in Penn State’s course catalog) with the Test Of Astronomy Standards, we focused our analysis on six questions that pertain directly to minigame topics. We found that two of these questions showed encouraging gains, while the other four demonstrated that students continued to hold on to common alternate conceptions within those topic areas. This finding suggests that more work is required to understand how to improve the games in ways that will further support student astronomy learning.

DOI: [10.1103/PhysRevPhysEducRes.14.010140](https://doi.org/10.1103/PhysRevPhysEducRes.14.010140)

I. INTRODUCTION

Play has long been recognized for the role it plays in the mental development of children and, as such, it provides a tantalizing possibility for use in educational settings [1,2]. Over the past two decades, researchers have built upon the concept of educating through play by focusing on the potential of video games and digital simulations as educational tools [3]. The promise of video games is twofold. First, video games have become ubiquitous in popular culture; a 2015 report from the Pew Research Center showed that, across all ages, 49% of Americans play video games [4]. Second, video games possess many features that have been shown to positively contribute to learning.

The cognitive benefits of video games are widespread. Commercial action or shooter video games can positively contribute to spatial cognition, an important component in

mathematical reasoning [5]. Squire, DeVane, and Durga [6] used a historical strategy game to improve student interest in and understanding of history and how it is intertwined with geography. College freshman engineering majors studying English vocabulary through web-based games showed significant improvement over a control group [7]. Video games also contribute strongly to language learning across a wide range of contexts [8]. Science classrooms across grade levels have also implemented video game-based modules and lessons to great success [9–11]. Games are also highly motivational, and inherently contain features that promote player engagement [3,12,13].

The potential benefits of games for educational and motivational purposes have also fueled interest in gamification, the process of incorporating gamelike elements in nongaming contexts with the goals of improving user engagement and motivation [14]. Some such game elements include the introduction of well-defined goals, competition or cooperation, reward structures like achievement badging, and mechanics for “leveling up” [15,16]. Gamification has been used in the business world both on the consumer end to promote consumer engagement and loyalty, and in the workplace to motivate employees to be more productive and complete training [17]. Gamification has even been used to promote participation in citizen

*Corresponding author.
danbarringer@gmail.com

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

science projects like FoldIt and those promoted under the Zooniverse umbrella [18,19].

For all the interest in the potential that games and game elements are purported to have in educational contexts, studies on gamification in higher education remain sparse. Dicheva, Dichev, Gennady, and Galia's [20] systematic mapping study identified only 34 papers that empirically studied the implementation of game elements in educational settings, only two of which studied K–12 education. Of these 34, only two studies looked at gamification in online classes. A later study, with less stringent selection criteria, identified 139 articles on gamification in educational settings between 2010 and 2014, but went into little detail about the overall content of those articles or what levels of education they addressed [21].

Thus, the goal of the current study is to expand the existing body of literature on games and gamification in higher education while exploring the effectiveness of a novel approach to teaching introductory undergraduate-level astronomy. In this paper, we explore the research question *Do gamified elements in an online course environment promote effective student learning?* We hypothesize that students should perform well on material that is directly covered by the most gamified elements of ASTRO 001V.

II. CONCEPTUAL FRAMEWORK

Because play has been a topic of interest to educational researchers and cognitive psychologists since at least 1933,¹ the positive aspects of games on learning are well established in the literature. Researchers have more recently identified how video games can likewise positively impact learning. We focus on the features that comprise video games and how these features have been shown to benefit student learning with emphasis on why video games and gamification are seen as such promising tools and methods for education.

A. Video games in education

In order to address how video game elements can positively impact learning, we must start by defining what a video game is and what features are inherently present in video games. Learning Science Through Computer Games and Simulations [3] identifies particular features of games such as providing feedback and responding and changing based on player interactions with the game. However, their definition becomes problematic, particularly when discussing learning games, when they state that games are generally played in informal contexts for fun, which supposedly differentiates video games from simulations.

¹1933 is the original publication year of Vygotsky's *Play and Its Role in the Mental Development of the Child* in Russian. It was translated into English in 1967.

We, therefore, define video games as having the following features:

Digital—Video games exist as programs or applications on digital media, including computers, gaming consoles, cell phones, etc.

Interactivity—Video games require a method for the user to interact with it. Video games receive and respond to inputs from players, and these inputs determine how the game evolves.

Simulations—Video games all run on internally consistent mechanics that determine how the player interacts with the game and how the game responds to those interactions.

Feedback—Games inherently test how well someone is able to play that game in a way that lets players know how well or poorly they are doing. Similarly, a player's ability to progress depends on successfully incorporating the feedback they receive to become better at the game. Feedback in games can be in the form of progress toward some ultimate goal, a score, completion time, achievement badges, a boss fight, or victory in a player vs player environment.

Here we will first discuss how three of the features that define a video game in terms of the roles those features can fulfill, particularly in learning games. The only feature that will not be discussed is the requirement that games be digital, as this is primarily a matter of video games' broad accessibility, as seen in Ref. [4]. To address how these features can be beneficial in educational settings, we turn to Steinkuehler and Squire's [13] discussion of the roles that video games can take on: how video games can be used as content, bait, and assessment (p. 378). Content refers to the substance of the game, the information it conveys, and what the game is about. In a learning game, the game's content will generally be the material the game was designed to teach. We use the term content delivery to describe the means by which the game conveys this information. Bait describes a game's ability to engage players and learners for entirely noneducational reasons. Assessment describes the forms of player evaluation inherent in game design and the ways in which these methods of providing feedback to a player can fit into educational settings.

B. Simulation

The simulation aspect of video games lends itself well to the role of content delivery, as this feature of video games fundamentally describes the content, internal mechanics, and rules of the game. In learning games, these worlds contain the means to deliver educational content to the player, who will ideally learn the intended content directly from interactions with this world. An early example of this comes from the game *Supercharged!*, in which students navigated mazes designed around introductory electrostatics problems. Students who played *Supercharged!* developed deeper conceptual understandings of the material than students who received traditional instruction [11].

Video games can also simulate situations where players can solve real-world problems in digital environments. Barab, Thomas, Dodge, Carteaux, and Tuzun's Taiga Park module within the larger Quest Atlantis game provides an example of this style of learning game [22]. Taiga Park has the students investigate problems pertaining to water quality in a simulated state or national park and weigh the concerns and interests of loggers, park rangers, visitors, and scientists among others. Students in a 4th grade class playing through the Taiga Park module were engaged in very high quality discourse regarding issues like water quality and conservation because these issues were tied directly to the game's content [10]. Barab *et al.* [10] specifically credit the situated nature of learning in Quest Atlantis in allowing students to become legitimate participants in scientific practices.

C. Feedback

All games are driven by some form of feedback. Winning or losing, granting achievements, assigning a score, and allowing a player to progress are all ways in which games can provide feedback about a player's in-game performance and, therefore, their mastery of the game. This is a type of formative assessment, since it encourages players to reflect upon their in-game actions and choices. Game players are engaged in the type of metacognition that Winne and Azevedo [23] called self-regulated learning, where players evaluate their gameplay and, when failing, are able to then revise their play style and try a task again.

Failure can also be motivating to video game players. In a study of game players' biological reactions to events in the video game Super Monkey Ball 2, players actually had positive reactions to ostensibly negative in-game events such as player death while playing the game, but still had negative reactions when shown replays of that same event [24]. They further note that these neurological reactions could be responsible for maintaining player interest and encouraging persistence in the face of failure. This increased likelihood to persist at a task is recognizable as an indicator of motivation [25]. In this way, the feedback in games not only plays the role of providing assessment, but, situationally, can also serve as bait.

An important part of the freedom to fail that video games provide is the feedback that comes with the failure. In a sense, the failure itself is a form of feedback; it is the game's way of telling the player "no, that didn't work; try something else." This is what Plass, Homer, and Hayward [26] call "corrective feedback," contrasting it with "explanatory feedback" in which the player receives an explanation of why they are incorrect (p. 46). While both forms of feedback are useful, explanatory feedback supports transfer and retention better than just corrective feedback [26]. Ke [27] observed that students playing games that only featured corrective feedback were far less reflective regarding their failures than those receiving explanatory feedback.

D. Interactivity

Across the literature on video games, authors seem to universally agree that one of the most important features of video games that can be used in educational contexts is that they are fundamentally interactive media. A video game's interactive nature means that players can have a say in designing their own play experiences, a principle that Gee [28] calls co-design. In the Quest Atlantis learning game, students were not just idle consumers of the material and the story being presented [10]. Rather, they became engaged participants through their interactions with the game.

As demonstrated by the high levels of student engagement in the studies cited above, interactivity, when done well, is bait. It is a way to get students invested in the material, so much so that they may even find themselves having fun. To be most effective, the interactions must be meaningful to the students. One way to create these meaningful experiences is through narrative.

A learning game's narrative ideally serves as both content delivery and bait for players. We say ideally because this depends entirely on how the content is tied to the narrative and how much players actually find the narrative itself engaging. To describe the connection between the content and the narrative of a game, Rieber [2] uses the terms "exogenous" and "endogenous" (p. 50). An exogenous game is one in which the content is not relevant to the gameplay, while endogenous games feature content and gameplay that are intrinsically related to one another. In exogenous games, the narrative and gameplay only serve as bait for players to get the content that is generally viewed as being tacked on to the game. Perhaps unsurprisingly, exogenous games are less effective as learning tools than endogenous games [27].

III. METHODS

In this section, we provide a brief description of the online introductory astronomy course, ASTRO 001V, the study participants, and our methods and instruments for data collection and analysis. Students enrolled in ASTRO 001V, as in all introductory astronomy courses at Penn State, are given identical online surveys during the first and last weeks of class to provide an objective measure of their learning that can be reliably compared across all ASTRO 001 instructors and course formats. Because the surveys contain distractor-driven multiple choice questions, we adopt a conceptual change framework that incorporates considerations of the motivations of the students and recognizes the social contexts in which student learning takes place [25]. We analyzed the results of these pre- and post-tests from two sections of ASTRO 001V and focused our analysis on questions from the survey that covered material from the most gamified portions of ASTRO 001V.

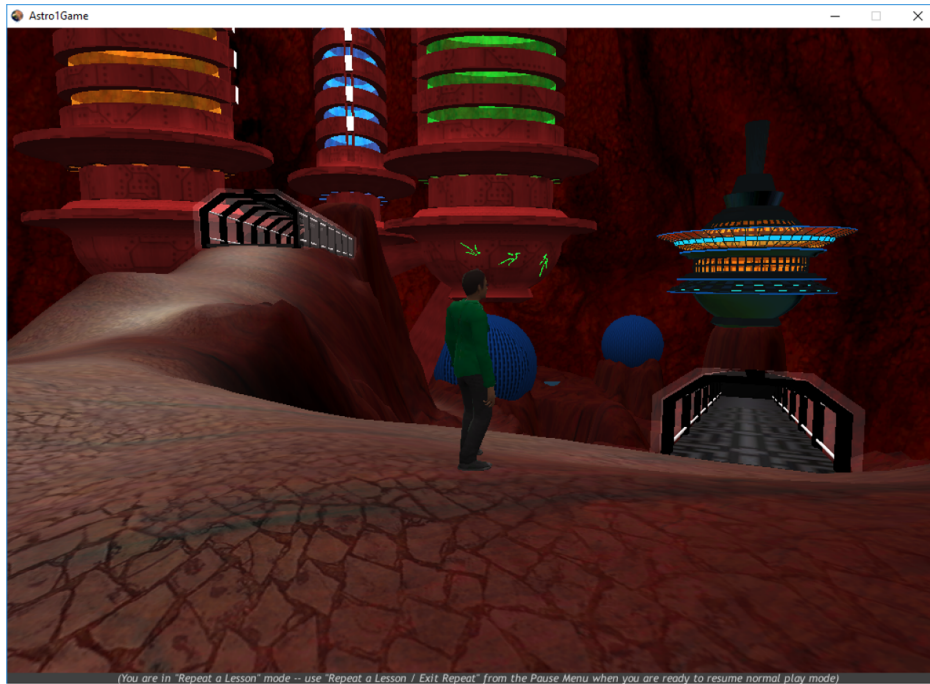


FIG. 1. Screenshot from ASTRO 001V showing part of the University of Mars underground complex.

A. Content and design of ASTRO 001V

In Spring 2014, the Penn State Department of Astronomy & Astrophysics debuted an online section of ASTRO 001: The Astronomical Universe that was run through a single-player digital world. ASTRO 001V was designed and programmed by Jane Charlton, Andrew Mshar, Nahks Tr'Enhl with further work contributed by Daryl Branford, Scott Given, Brendan Mullan, Brian Pomerantz, and Tim Schneider. ASTRO 001V presents the course material through an interactive science fiction story in which the customizable player character is a student at the University of Mars.² Figure 1 shows part of the underground portion of the digital University of Mars consisting of the dormitory building (left) and the Virtual Reality Lab, where most of the minigames can be accessed.

ASTRO 001V delivers what would otherwise be lecture or reading material through the player character's conversations with nonplayer characters (NPCs), primarily your fellow students Fabian and Thomasin and your alien instructor Lerpz. The lecture material is interspersed with questions and individual lessons are separated by in-game quizzes. These quizzes are ungraded and used for student self-assessment only, and may be retaken at any time. For both the quizzes and in-lecture questions, students must answer the question correctly to proceed. When a question is answered incorrectly, students are given feedback that does not actually give them the answer to the question, but

ideally serves to prod them into the right line of thinking. The quizzes, however, do not use explanatory feedback. The feedback is identical regardless of which incorrect choice was made. As with the quizzes, lessons can also be repeated at any time, which is specifically recommended to the students by the quiz-administering NPC if they do poorly on the end of lesson quiz.

By our definition of video games provided above, ASTRO 001V is not a game in its own right. Rather, it is a virtual class environment that contains gamified elements such as existing in an entirely digital world setting and offering the ability to replay or redo lessons or activities at will [15]. It also contains brief, small-scale games embedded in the larger digital world that we refer to as minigames, which are the most gamified feature of the course. The reason ASTRO 001V as a whole cannot be considered a game is because the program itself contains no means of evaluation or feedback regarding the overall course goals. These forms of evaluation, specifically homework and exams, are external to the digital world. The feedback students do receive also does not affect their progression through the course in this digital setting because students do not have to become better at astronomy to progress through the course. Rather, students can find the correct answer to most of the in-game questions through clicking on answer choices for a given question until they come across the correct answer.

1. The minigames

The minigames are included after the relevant material is presented through the NPCs and are intended to act as both

²A demo of the ASTRO 001V program can be requested at <http://www.psuastrogame.com/demo/>.

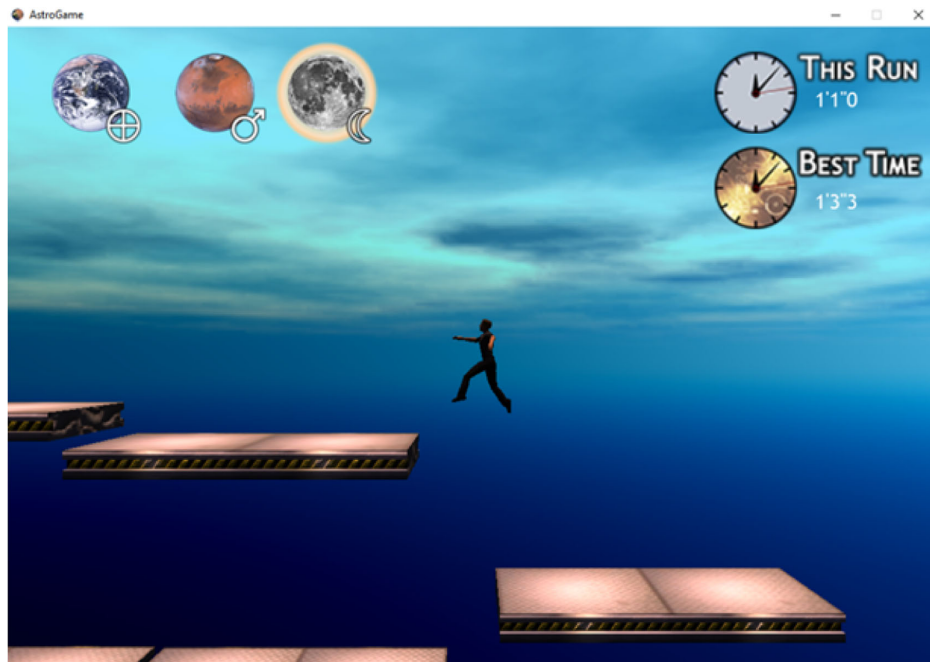


FIG. 2. Screenshot from the Gravity Platformer minigame showing the player avatar jumping between platforms in Moon gravity.

interactive teaching and self-evaluation tools. After students go through the minigame tutorial, they are scored on how well they do in game. Here we discuss the three minigames that are the focus of our analysis.

The first minigame that students encounter is a timed platform jumper made to demonstrate the effects of surface gravity. The Gravity Platformer directly follows the lesson on surface gravity, and aims to reinforce the students' understanding of gravity by simulating its effect on the player's avatar. The student interacts with the game by moving the avatar around within the level to reach the end point. The student can also change the strength of gravity in the simulation, which is necessary to move through the level because the platforms have different height separations, as seen in Fig. 2. The timer in the upper-right corner provides a potential source of motivation for the student; a faster run is better. While students can just run through the level once, seeing that they are timed could encourage students to replay the minigame to try and get a better time.

The Moon Phases minigame is based around a simulation of the Earth-Moon-Sun system, as shown in Fig. 3. The simulation shows a static Earth and Sun with the Moon orbiting Earth. Players stop the simulation in accordance with prompts that appear in the text window portion at the bottom of the screen. The prompts tell the player to stop the motion of the Moon when it is in a position that would make the statement true. The prompts refer to things like the phase of the Moon, the time at which it rises, sets, or is at peak altitude, or a type of eclipse that could happen with the Moon in that position. The player is scored on how close to the correct position they are able to stop the Moon, with a maximum score of 30 on each attempt.

Players are given 15 attempts to accumulate as many points as possible. The score in the Moon Phases minigame is a more endogenous feature than the timer in the Gravity Platformer, as it provides feedback that relates directly to the student's understanding of Moon phases. As with the timer in the Gravity Platformer, the score in the Moon Phases minigame could motivate students to try to maximize their score.

The Spectroscopy minigame has two parts to it. First, the student must match a given spectrum (either absorption or emission) to the conditions (emission spectrum from a hot diffuse gas, absorption spectrum from a hot object behind a cooler, more diffuse gas) and elements that create it. Then the simulation zooms in on the gas to show what is occurring on the level of individual atoms, which is depicted in Fig. 4. Here, the students match the incoming photon to the energy level jump of the electron in the atom or the outgoing photon to the emission line in the spectrum that it would create. The first time that students encounter the second part of the simulation, they are given a number of questions about which particles are which (e.g., making sure they know that electrons specifically are responsible for absorbing or emitting photons) and the relationship between the energy and wavelength or frequency of a photon). Of the three minigames we focus on, the Spectroscopy minigame features the least student control over the progression of the simulation. Here, the score reflects a student's ability to correctly match the numbers associated with the photon's energy to the absorption or emission lines in the spectrum created by the given atom, a skill that we feel only weakly reflects a student's understanding of spectroscopy.

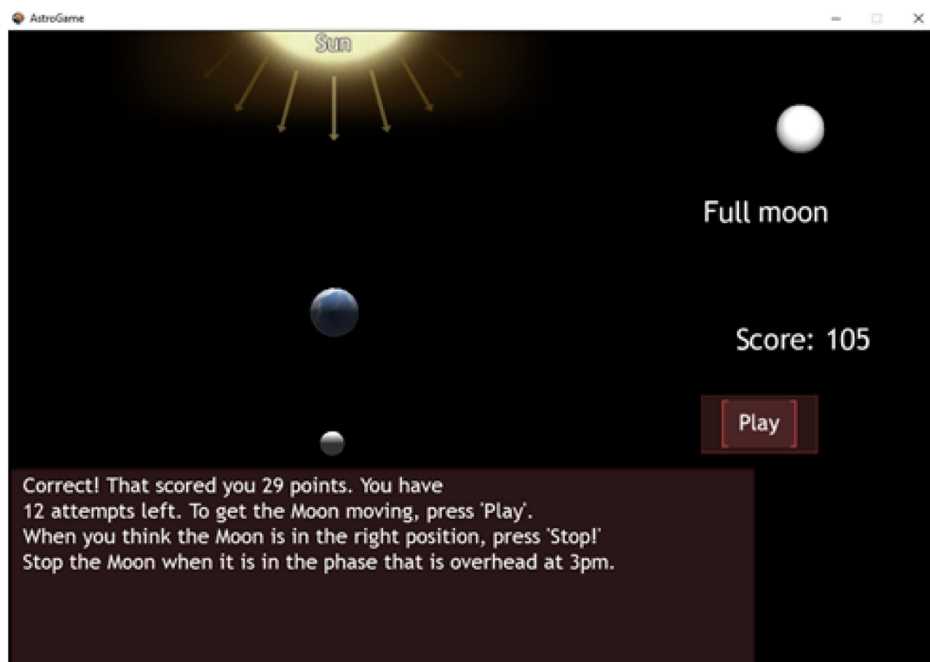


FIG. 3. Screenshot of the Moon Phases minigame taken after a question was answered correctly.

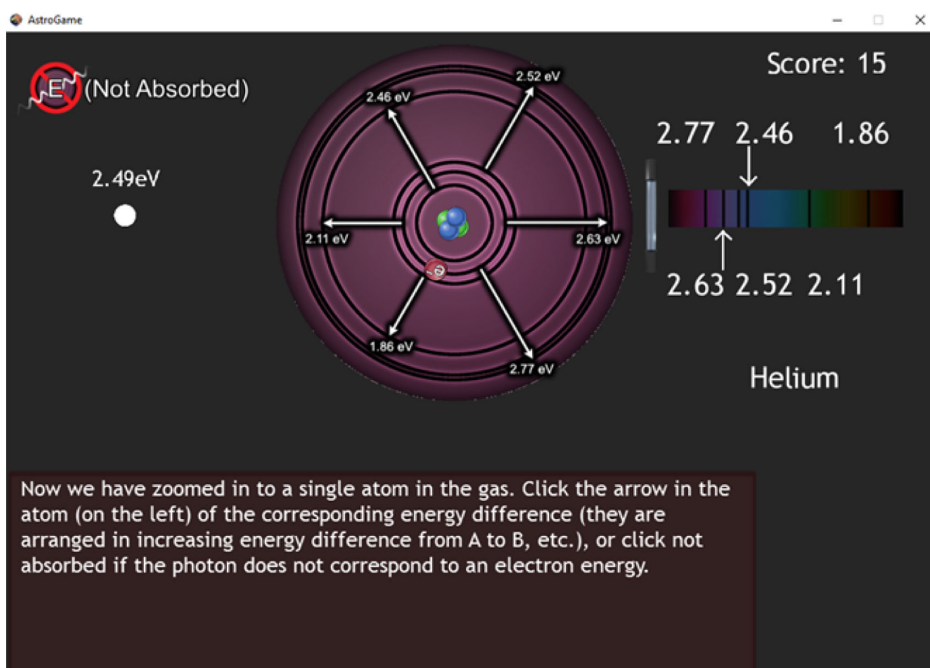


FIG. 4. Screenshot from the part of the Spectroscopy minigame where the students see how absorption (shown here) and emission occur on the scale of individual atoms.

B. Subjects and setting

The target population for this study was students enrolled in ASTRO 001V in Spring 2014 and Fall 2014. These were the first two times the course was offered. ASTRO 001 is the course number for general education, nonscience major, introductory astronomy at Penn State; therefore, students

enrolled in this course are likely to represent primarily students needing to take this course to fulfill a general education science requirement and are not likely to be astronomy majors. All students in the study voluntarily chose to enroll ASTRO 001V, rather than other sections of ASTRO 001 which did not have a video game component.

C. Data collection

We collected preinstructional (pretest) and postinstructional (post-test) data from both the Spring 2014 and Fall 2014 semesters using the Test Of Astronomy STandards (TOAST). The TOAST is a research-validated assessment instrument designed to assess students' general astronomy content knowledge [29]. The questions and answer choices on the TOAST are written using as little scientific language as possible, and distractors are written to "reflect students' most common alternative conceptions" [29]. Unlike a concept inventory, TOAST is written to cover the wide range of topics typically covered in an introductory astronomy class, including gravity, electromagnetic radiation, stellar evolution, and moon phases.

The TOAST surveys are conducted online through Penn State's course management system during the first and last weeks of class. The course management system also logs the amount of time a student spends on a particular item within the system. We matched student pre- and post-tests with their associated time stamps and purged student identifiers to maintain anonymity. Any tests without a match were removed from the data set.

Students were required to complete the TOAST assessment, but their performance on the assessment was not included in their final grade. Thus, some students may have provided random answers to the questions, to complete the survey quickly, rather than reading the questions before selecting an answer. Therefore, we solicited graduate student volunteers in the Astronomy & Astrophysics department to take the TOAST to establish a baseline for how quickly an expert can complete the assessment. This establishes a best-case scenario for test completion, and therefore a minimum feasible time for a student who puts in a valid attempt to read and answer the questions. We received 7 volunteers who had an average completion time of 10:35, with a minimum time of 5:37. To be on the conservative side, we rounded the minimum cutoff time down to 5:00.

We therefore removed any pre- and post-tests from our sample that were completed in less than five minutes and more than one hour (most of which were well over an hour). We also removed any answer keys showing repeated patterns (e.g., all the same letter, or ABCDABCD) and those that had more than three answers left blank. Applying

these constraints to our data from the first two sections of the Astro 001 Game class cut the total number of valid responses from 623 to 466, as shown above in the column labeled "Matched, postcut" of Table I.

D. Data analysis

Because we want to focus on the gamified features of Astro 001V, we tried to isolate the learning that is thought to occur primarily from the embedded minigames and interactive simulations. To look at this specifically using the TOAST, we asked the two primary game developers (Mshar and Tr'Ehnl) to rate the TOAST questions based on how relevant they were to in-game material and activities. Both of the developers, beyond their expertise in Unity, have undergraduate degrees in Astronomy from Penn State. The lead author also rated the questions in the same manner based on his experience playing through the game. The questions were rated on a descending letter scale from "A: Addressed directly in a minigame" to "G: Not covered in course at all." We chose to focus particularly on any question that received at least two "A" ratings. Six TOAST questions met this criterion. The topics these six questions covered were moon phases (question 5), structure of the solar system (question 8), gravity (question 20), and the nature of electromagnetic radiation (questions 22, 25, 26). Questions 25 and 26 specifically ask about wavelength and energy of an emitted or absorbed photon in a Bohr model atom, while question 22 asks about what causes energy to be emitted by atoms. The full table of question ratings and the normalized gains associated with those questions (separated by class section) can be found in Table II.

To compare the learning gain scores from our selected questions with the student gains on non-mini-game questions, we created a comparison set by removing our questions of interest along with the questions not addressed in the curriculum at all (those receiving at least two G ratings: questions 4, 6, and 21 as seen in Table II) from our TOAST results. We combined both classes and averaged the normalized gains from the remaining eighteen questions, giving us a comparison score of 0.19 (SD = 0.17). We used this gain score and standard deviation as our point of comparison when discussing the gain scores for the individual questions of the TOAST. This allowed us to directly compare the students' performances on questions

TABLE I. Summary of the collected TOAST surveys used in this study. Summary of data obtained from TOAST surveys of two sections of ASTRO 001V. The "Total" learning gain is calculated by combining the data from the two classes.

Class section	Pretests	Post-tests	Post-tests (matched)	Matched, postcut	Learning gain $\langle g \rangle$
Total	1052	626	623	466	0.17
Spring 2014	495	292	290	224	0.17
Fall 2014	557	334	333	224	0.18

TABLE II. Relevance of TOAST questions to ASTRO 001V course material and learning gains associated with each question. TOAST questions were rated according to their relevance to the material covered by ASTRO 001V. The rating scheme is as follows: (A) Addressed directly in a minigame, (B) Addressed indirectly in a minigame or not emphasized, (C) Addressed directly through applets, (D) Addressed indirectly through applets or not emphasized, (E) Addressed directly through NPC discussions and reading, (F) Addressed indirectly through NPC discussions and reading or not emphasized, (G) Not covered in course at all.

TOAST Question No.	Question relevance rating			Normalized gain	
	Tr'Enhl	Mshar	Barringer	Spring 2014	Fall 2014
1	B	B	B	-0.05	0.01
2	A	B	B	0.55	0.47
3	B	A	B	0.08	0.28
4	C	G	G	0.13	0.04
5	A	A	A	0.16	-0.02
6	G	F	G	-0.01	-0.02
7	C	C	E	0.36	0.29
8	E	C	E	0.19	0.22
9	E	E	E	0.21	0.20
10	D	A	E	0.20	0.26
11	C	B	E	0.12	0.15
12	E	E	E	0.29	0.30
13	A	E	E	0.21	0.13
14	E	E	E	0.19	0.22
15	E	E	F	-0.13	0.01
16	E	E	F	0.43	0.44
17	E	E	E	0.30	0.39
18	B	A	A	0.14	0.22
19	E	E	E	0.19	0.11
20	A	A	B	0.31	0.36
21	G	G	G	0.07	0.03
22	A	A	A	0.35	0.32
23	E	E	E	0.02	0.17
24	E	E	F	0.03	0.04
25	A	A	A	0.17	0.08
26	A	A	A	-0.04	0.06
27	E	E	F	0.01	0.05

that were covered by minigames and those that were not. If the learning gain score for one of the minigame-aligned questions is above 0.19, then students learned that gamified material better than the material from the non-gamified portion of the course. Conversely, if the normalized gain for a question is below 0.19, then students learned that material less effectively than we would expect. Effect sizes were calculated using Cohen's d [30].

IV. RESULTS

A. Gain scores by minigame topic

The class gains for the TOAST questions deemed most relevant to the minigames are shown in Table III. The learning gains for question 20 on surface gravity and question 22 on the Bohr model are both above 0.19 with a large effect size. Therefore, we conclude that the

TABLE III. Learning gains and effect sizes for selected questions. Learning gains by class for TOAST questions deemed to be closely tied to material covered in ASTRO 001V minigames. Questions selected for further analysis indicated in boldface. Effect sizes (ES) come from comparison to average gain score of 0.19, discussed above.

Question No.	Spring 2014	ES	Fall 2014	ES	Question topic
5	0.16	0.17	-0.02	1.2	Moon phases
18	0.14	0.29	0.22	0.17	Solar system structure
20	0.31	0.71	0.36	1.1	Surface gravity
22	0.35	0.94	0.32	0.81	Electromagnetic radiation
25	0.17	0.13	0.08	0.69	Electromagnetic radiation
26	-0.04	1.4	0.06	0.81	Electromagnetic radiation

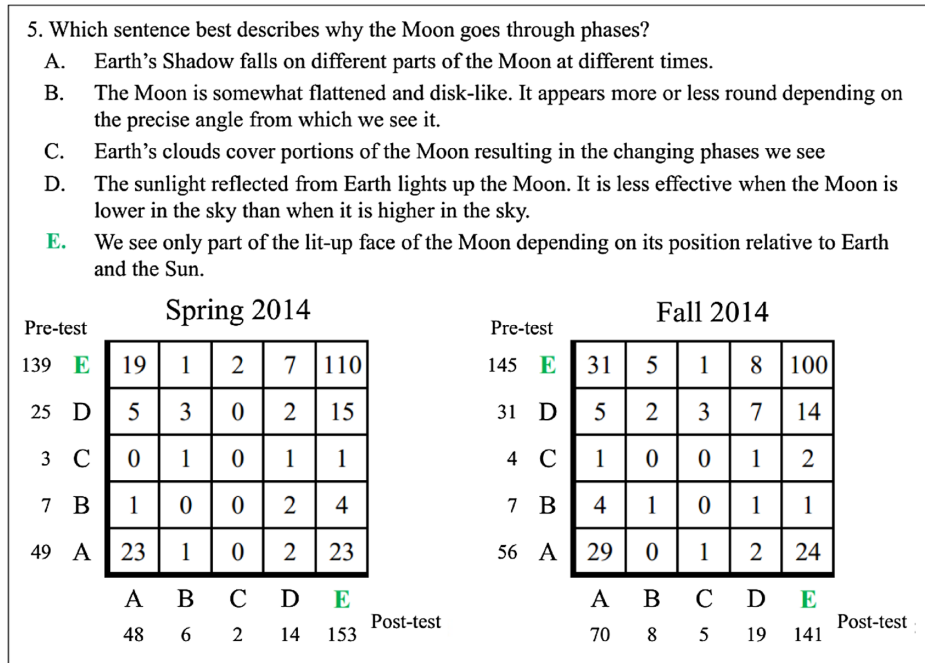


FIG. 5. Comparison of how students responded to TOAST question 5 on the pre- and post-test. Question adapted from Ref. [29].

minigames that address surface gravity and basics of the Bohr model do so in a manner that positively contributed to student learning.

The other questions we focus on, 5, 25, and 26, show surprisingly low, even negative in a few cases, learning gains. With the exception 25 in Spring 2014, all of the learning gains for these questions were significantly below the overall normalized gain score of 0.19 we are using as our threshold. While the learning gain score on question 18 also fell below our threshold, the small effect size led us to exclude it from more detailed analysis. The learning gains for these three questions do not support our initial hypothesis that the most gamified elements in ASTRO 001V, the minigames, would lead to effective learning on the topics they cover. Therefore, we analyzed how students’ answer choices changed between the pretest and the post-test. The answer choices on the TOAST were written using research on students’ common alternate conceptions, so by analyzing how students’ answer choices changed, we are able to suggest how students’ conceptions were changed, if at all, as a result of playing the minigames.

B. Analysis of TOAST question answer choices

Each of the following figures contains the relevant question and answer choices as they appear on the TOAST for the three items where students’ gain scores were below the overall normalized gain score. The panels below will show the comparison of pretest and post-test responses for each section separately. An example of how to read the graphs is as follows: the number in column A, row B tells you the number of students in the section who

chose answer B on the pretest and answer A on the post-test. The correct answer choice is indicated in green.

TOAST question 5, which addresses Moon phases, is shown in Fig. 5. As shown in Table III, this question showed low gains in both sections, but the student responses show an unsurprising pattern. While the overwhelming number of students got and remained with the correct answer (Moon phases caused by relative positions), the second most popular response (Moon phases caused by Earth’s shadow) is the most common misconception about the cause of Moon phases. One potential limitation in the gain we found on this item relating to the lunar phases minigame is due to a ceiling effect as most students began by selecting the correct response. However, a portion of the students continued to use the common alternative conception that the Earth’s shadow causes the lunar phases [31]. While about the same number of students who had this conception improved as remained using this idea, this suggests that the minigame did not fully support all learners in changing their explanation of this challenging phenomenon.

The Moon Phases minigame is based around a simulation of the Earth-Moon-Sun system. The simulation shows a static Earth and Sun with the Moon orbiting Earth. Students are given prompts that describe a particular Moon phase, and must stop the simulation when the Moon is in the position indicated by the prompt. Students are then scored on how close to the correct position they are. The gameplay directly evaluates the student’s understanding of how the relative positions of the Earth, Moon, and Sun contribute to producing Moon phases, and the score provides direct feedback on their understanding of this

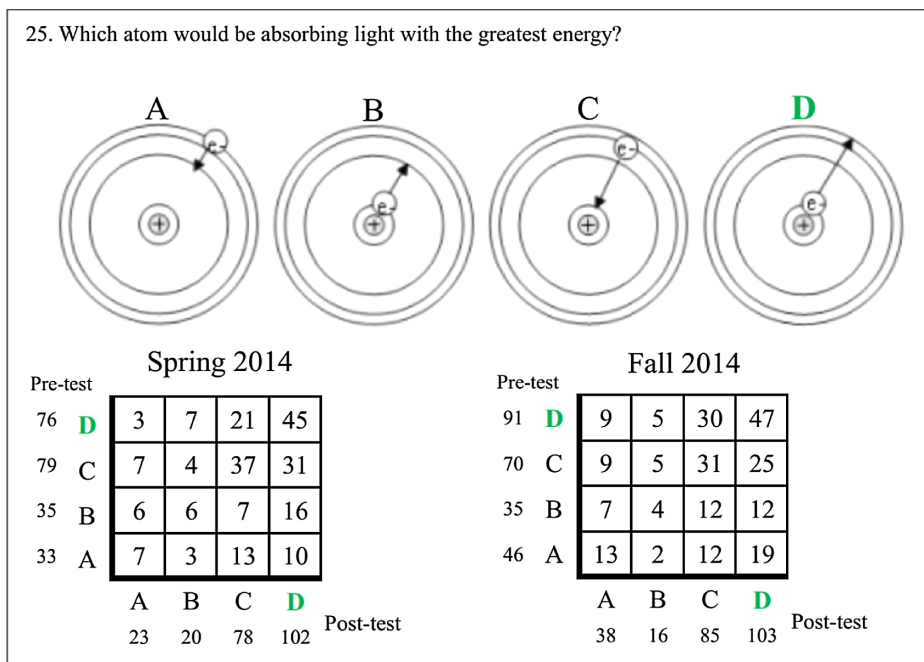


FIG. 6. Comparison of how students responded to TOAST question 25 on the pre- and post-test. Question adapted from Ref. [29].

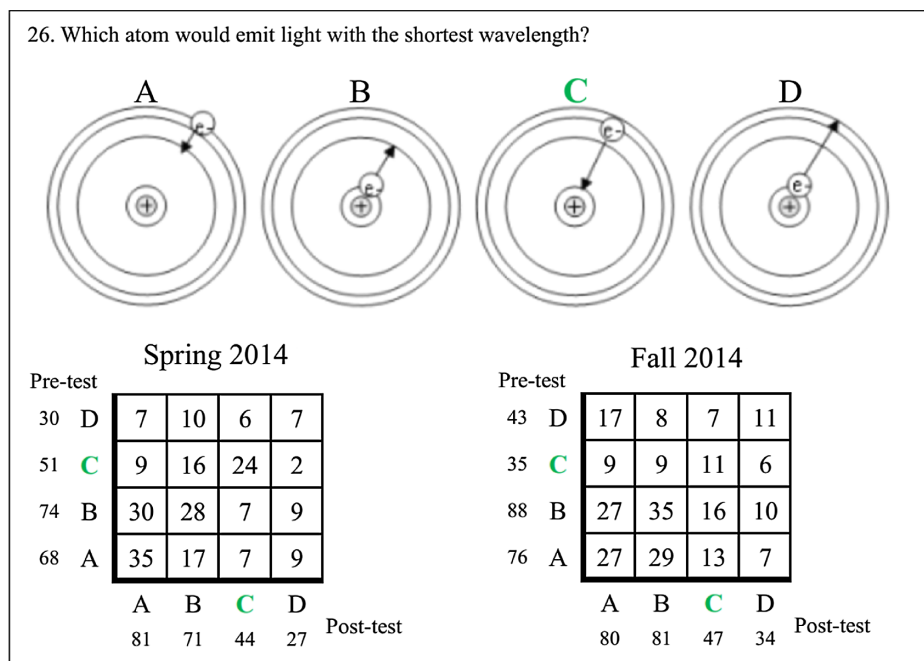


FIG. 7. Comparison of how students responded to TOAST question #26 on the pre- and post-test. Question adapted from Ref. [29].

concept. But as our results show, students still held on to incorrect conceptions of the cause of Moon phases.

Both questions 25 (Fig. 6) and 26 (Fig. 7) relate to students' understanding of the emission and absorption of light by atoms. The students' answer selections on question 25 show that, for the transition with the greatest energy, the students generally knew to pick the drawing with the

longest arrow, corresponding to the largest energy transition. In both classes, there was also a fair amount of switching from A or B answers on the pretest to C or D on the post-test, which indicates that the students in both classes were able to make the larger transition = higher energy after instruction. However, they seem to have been uncertain about which drawing with the longest arrow to

choose, likely forgetting that the electron goes up in energy levels when it absorbs light. Based on the low gain scores across both sections, question 26 is more problematic, though the mistake the students made is a common one. In both classes, A and B were the clear favorite choices, meaning that students likely made the incorrect connection between shorter wavelength (higher energy) and shorter arrows (lower energy), a common mistake in introductory physics and astronomy.

In the spectroscopy minigame, students must match a given spectrum (either absorption or emission) to the conditions (emission spectrum from a hot diffuse gas, absorption spectrum from a hot object behind a cooler, more diffuse gas) and elements that create it. Then the simulation zooms in on the gas to show what is occurring on the level of individual atoms. Here, the students match the incoming photon to the energy level jump of the electron in the atom or the outgoing photon to the emission line in the spectrum that it would create. Students are scored on their matching abilities, and the total score is kept for the game. The depictions of the individual atoms in the simulation are similar to the depictions on the TOAST, so these should be familiar to students. The feedback received is also directly relevant to the students' understanding of spectroscopy. Despite the relevance to the subject material and the presence of educationally beneficial features, the spectroscopy minigame did not have the expected effect on student conceptions.

V. DISCUSSION

Our results indicate that, as a whole, the minigames did not contribute to student learning in ASTRO 001V in the manner we expected; students made lower gains on most of the questions (three of the six TOAST questions) that corresponded conceptually with the minigames than the overall normalized gain score in the course, while the fourth (question 18) showed student performance that is comparable with the average of the remaining questions. We hypothesized that because the minigames were interactive, provided just-in-time feedback, and simulated phenomena to the necessary degree of accuracy, these features would help students learn the intended material. However, our analysis suggests learning science through gamified elements in online courses requires a consideration of more than just these elements from the literature.

In looking at TOAST questions that were deemed to be closely related to minigame material, we found that the questions on surface gravity and the basics of the emission of electromagnetic radiation showed positive learning gains on par with those from in-person lecture classes using interactive engagement teaching methods. However, students performed significantly less well on the questions on Moon phases and the Bohr model of absorption and emission in atoms. Analysis of student responses on these questions suggests that students either learned or continued

to hold common alternate conceptions in these areas after instruction. This suggests that the minigames did not support a large proportion of the students in reconceptualizing their understanding of these astronomy concepts.

While gamification of education and creating and using video games in educational settings clearly have many benefits, our work demonstrates that such benefits do not come automatically. Just because content is delivered through an endogenous game with strong alignment between the game's learning goals and mechanics does not mean it will be effective at teaching that material. Like any educational tool, games require study, iteration, and constant improvement [3]. They also require appropriate activities and materials within the class and reflection by the students [13]. Additionally, our results suggest that a further element that should be considered when designing video games in educational setting is students likely alternative conceptions and how the game will explicitly address and support conceptual change of that particular aspect of students' conceptualization of science.

One limitation of this study is the extent to which we are able to disentangle the effects of playing the minigames addressing these topics from other factors in the course design. We are also limited in the extent of how we analyzed student learning using only a single question matched to each minigame. Future research could take a more in-depth analysis of the low student gains on TOAST questions 5, 25, and 26 through a controlled experiment (to isolate the potential factors in the student's performance) along side interviews with students regarding their understanding of the material. Further, the use of detailed testing instruments such as the Lunar Phases Concept Inventory and the Light and Spectroscopy Concept Inventory could also give us a better idea of where the problem areas are in students' understanding of the material addressed in these minigames [32,33].

For future studies, we also suggest a mixed-method approach that combines the quantitative data from diagnostic tests and concept inventories with observations of how students actually play the games. Because of the preliminary nature of our study, we were unable to gather data on this ourselves. Direct data on how students interact with the ASTRO 001V program, such as from recorded sessions with students thinking aloud and interviews, would illustrate how students progress through the minigames in particular and the overall program itself. This information would be essential to evaluate how the minigames could be redesigned to promote more effective student learning. Without these data, however, there is little we can meaningfully suggest for the improvement of ASTRO 001V.

The study of learning games and gamification for learning is young, which means there are still gaps in the body of education research on the benefits of video games for learning. This is not helped by the inherently broad nature of "video games" and the wide variety of uses for which they can be adapted and created. What this means

for developers of educational games is that there are no universal best practices one can follow in either design or implementation. While we can point out things that clearly do not work, like making games exogenous from their intended content, ideal practices appear to be mostly context dependent and perhaps content dependent.

The authors would like to thank the development team of ASTRO 001V for the years of work and support they put into developing this class. The lead author would also like to thank Dr. Gabriella Richard for valuable in class discussions that contributed to his academic understanding of the video game medium.

-
- [1] L. S. Vygotsky, Play and its role in the mental development of the child, *Sov. Psych.* **5**, 6 (1967).
- [2] L. P. Rieber, Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games, *Educ. Technol. Res. Dev.* **44**, 43 (1996).
- [3] Board on Science Education, Division of Behavioral and Social Sciences and Education, National Research Council, Introduction, in *Learning Science Through Computer Games and Simulations. Committee on Science Learning: Computer Games, Simulations, and Education*, edited by Margaret A. Honey and Margaret L. Hilton (The National Academies Press, Washington, DC, 2011).
- [4] M. Duggan, *Gaming and Gamers* (Pew Research Center, Washington, DC, 2015), <http://www.pewinternet.org/2015/12/15/gaming-and-gamers/>.
- [5] J. Feng, I. Spence, and J. Pratt, Playing an action video game reduced gender differences in spatial cognition, *Psychol. Sci.* **18**, 850 (2007).
- [6] K. Squire, B. DeVane, and S. Durga, Designing centers of expertise for academic learning through video games, *Theory Into Practice*, **47**, 240 (2008).
- [7] F. W. M. Yip and A. C. M. Kwan, Online vocabulary games as a tool for teaching and learning English vocabulary, *Educ. Media Int.* **43**, 233 (2006).
- [8] M. Peterson, Computerized games and simulations in computer-assisted language learning: A meta-analysis of research, *Simulation and gaming*, **41**, 72 (2010).
- [9] L. A. Annetta, M. T. Cheng, and S. Holmes, Assessing twenty-first century skills through a teacher created video game for high school biology students, *Res. Sci. Technol. Educ.* **28**, 101 (2010).
- [10] S. A. Barab, T. D. Sadler, C. Heiselt, D. Hickey, and S. Zuiker, Relating narrative, inquiry, and inscriptions: Supporting consequential play, *J. Sci. Educ. Technol.* **16**, 1 (2007).
- [11] K. Squire, M. Barnett, T. Higgenbotham, and J. Grant, Electromagnetism Supercharged!, in *DIGITAL LEARNING GAMES 18, Proceedings of the 2004 International Conference of the Learning Sciences*, edited by Y. Kafai, W. Sandoval, N. Enyedy, A. Dixon, and F. Herrera (Lawrence Erlbaum, Mahwah, 2004).
- [12] A. W. Przybylski, C. S. Rigby, and R. M. Ryan, A motivational model of video game engagement, *Rev. Gen. Psychol.* **14**, 154 (2010).
- [13] C. Steinkuehler and K. Squire, Videogames and Learning, *The Cambridge Handbook of the Learning Sciences*, 2nd ed., edited by R. K. Sawyer (Cambridge University Press, New York, 2014), pp. 377–389.
- [14] S. Deterding, M. Sicart, L. Nacke, K. O’Hara, and D. Dixon, Gamification: Using game-design elements in non-gaming contexts, *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems* (Association for Computing Machinery (ACM), New York, 2011).
- [15] K. M. Kapp, It’s in the Game: Understanding Game Elements, in *The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education* (Pfeiffer, San Francisco, 2012).
- [16] R. McDaniel, What we can learn about digital badges from video games, in *Foundation of Digital Badges and Micro-Credentials: Demonstrating and Recognizing Knowledge and Competencies*, eBook, edited by D. Ifenthaler, N. Bellin-Mularski, and D.-K. Mah (Springer International Publishing, Switzerland, 2016), pp. 325–342.
- [17] B. J. Arnold, Gamification in Education, in *Proceedings of the American society of business and behavioral sciences (ASBBS)* (American Society of Business and Behavioral Sciences, San Diego, 2014), Vol. 21, pp. 32–39.
- [18] C. B. Eiben, J. B. Siegel, J. B. Bale, S. Cooper, F. Khatib, and B. W. Shen, FoldIt players, B. L. Stoddard, Z. Popovic, and D. Baker, Increased Diels-Alderace activity through backbone remodeling guided by FoldIt players, *Nat. Biotechnology* **30**, 190 (2012).
- [19] J. Cox, E. Y. Oh, B. Simmons, C. Lintott, K. Masters, A. Greenhill, G. Graham, and K. Holmes, Defining and measuring success in online citizen science: A case study of zooniverse projects, *Comput. Sci. Eng.* **17**, 28 (2015).
- [20] D. Dicheva, C. Dichev, G. Agre, and G. Angelova, Gamification in education: A systematic mapping study, *Educ. Technol. Soc.* **18**, 75 (2015).
- [21] J. Martí-Parreño, E. Méndez-Ibáñez, and A. Alonso-Arroyo, The use of gamification in education: A bibliometric and text mining analysis, *J. Comput. Assist. Learn.* **32**, 663 (2016).
- [22] S. Barab, M. Thomas, T. Dodge, R. Carteaux, and H. Tuzun, Making learning fun: Quest Atlantis, a game without guns, *Technol. Res. Develop.* **53**, 86 (2005).
- [23] P. H. Winne and A. Azevedo, Metacognition, in *The Cambridge Handbook of the Learning Sciences*, 2nd ed.,

- edited by R.K. Sawyer (Cambridge University Press, New York, 2014), pp. 21–43.
- [24] N. Ravaja, T. Saari, M. Salminen, J. Laarni, and K. Kallinen, Phasic emotional reactions to video game events: A psychophysiological investigation., *Media Psychol.* **8**, 343 (2006).
- [25] P. R. Pintrich, R. W. Marx, and R. A. Boyle, Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change, *Rev. Educ. Res.* **63**, 167 (1993).
- [26] J. L. Plass, B. D. Homer, and E. O. Hayward, Design factors for educationally effective animations and simulations, *J. Comp. High. Educ.* **21**, 31 (2009).
- [27] F. Ke, A Case Study of Computer Gaming for Math, *Comput. Educ.* **51**, 1609 (2008).
- [28] J. P. Gee, Learning by design: Good video games as learning machines, *E-Learning* **2**, 5 (2005).
- [29] S. J. Slater, The development and validation of the Test of Astronomy STandards, *J. Astrono. Earth Sci. Educ.* **1**, 1 (2014).
- [30] J. Cohen, The t Test for Means, in *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. (Routledge, Abingdon, 1988).
- [31] J. Baxter, Children’s understanding of familiar astronomical events, *Int. J. Sci. Educ.* **11**, 502 (1989).
- [32] R. S. Lindell and S. R. Sommer, Using the Lunar Phases Concept Inventory to investigate college students’ pre-instructional mental models of lunar phases, *AIP Conf. Proc.* **720**, 73 (2004).
- [33] E. M. Bardar, Ph. D. thesis, Boston University, Boston, MA, 2006. Retrieved from <http://search.proquest.com/docview/305364831/E60D999EF5D3464BPQ/2?accountid=13158>.