Ice Paddles, CO2 Invaders, and Exploding Planets

How Young Students Transform Climate Science Into Serious Games

Giovanni Maria Troiano
g.troiano@northeastern.edu
Northeastern University
Boston, USA

Dylan Schouten
d.schouten@northeastern.edu
Northeastern University
Boston, USA

Michael Cassidy
michael_cassidy@terc.edu
TERC
Boston, USA

Eli Tucker-Raymond
etuckerr@bu.edu
Boston University
Boston, USA

Gillian Puttick
gillian_puttick@terc.edu
TERC
Boston, USA

Casper Harteveld
c.harteveld@northeastern.edu
Northeastern University
Boston, USA

ABSTRACT

In game-based curricula that leverage game design, students can learn creatively by transforming serious topics into video games (i.e., serious games). However, as these games remain mostly underexplored, we know little about how students design them and the extent to which they reflect content uptake. Here, we leverage a framework for serious games called *Triadic Game Design* (TGD) to analyze 391 games on climate science, which were designed by 8th-grade students with *Scratch*. Based on a large-scale TGD-based analysis, we provide an overview of design outcomes emerging from student games, and analyze how *reality*, *meaning*, and *play* are articulated in these games to reflect content uptake. Then, we ask two experts in game design and education to assess a subset of the 20 most representative games, to reflect on further design and pedagogical insights that may have not been captured by the large-scale analysis. Our results reveal a wide range of design outcomes, where Pong-like games teach players about the ice-albedo feedback loop, and CO_2 molecules become targets to be shot in games like *Space Invaders*. Our work can serve as guidance and inspiration to help both researchers and educators evaluate student-designed games, as well as reason about how to use them as assessment tools in game-based constructionist curricula.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; • **Applied computing →** *Interactive learning environments.*

KEYWORDS

Game-based learning, Scratch, TGD, constructionism, artifacts

ACM Reference Format:

Giovanni Maria Troiano, Dylan Schouten, Michael Cassidy, Eli Tucker-Raymond, Gillian Puttick, and Casper Harteveld. 2020. Ice Paddles, CO2 Invaders, and Exploding Planets: How Young Students Transform Climate Science Into Serious Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '20), November 2–4, 2020, Virtual Event, Canada.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3410404.3414256

CHI PLAY '20, November 2–4, 2020, Virtual Event, Canada

ACM ISBN 978-1-4503-8074-4/20/11…$15.00
https://doi.org/10.1145/3410404.3414256

1 INTRODUCTION

Modern education is gradually moving towards creative [19] and sustainable learning [107], with emphasis on problem-solving [40] and digital literacy [36]. In that respect, game-based learning is showing promise and the benefits of using games for such learning have been demonstrated extensively (e.g., [7, 34, 66, 105, 108, 118]). For instance, video games have been used to effectively teach various concepts to students in science and engineering [64], or help young adults learn and explain complex forces in physics [34]. Furthermore, prior work in game-based learning that leverage game design tools (e.g., Scratch [95]) showed how such an approach enhances peer collaboration [100], and motivates students to improve their learning performance throughout the curriculum [79].

In this paper, we focus on games in education in the context of constructionist curricula (e.g., [14, 91, 91, 116]) that leverage game design for creative and student-centered learning. Despite the massive number of artifacts designed by young students within and outside game-based constructionist curricula (see [4]), prior work focused mainly on how problem-solving [102] and computational thinking (CT) [70, 117] can effectively develop through game design [5, 6, 28, 48, 61, 72, 115, 116, 122]. Therefore, few have looked at or attempted to assess existing student-designed artifacts in these contexts [16, 28, 49], and it remains unclear how educators can use such games as tools for assessing student learning. Furthermore, previous work mostly looked at small datasets (e.g., three games in [16]). Hence, we lack comprehensive overviews and extensive analyses of design outcomes, which emerge from games designed by students in constructionist curricula. For instance, we know little about how students organize the various design elements composing their games (e.g., the title or game genre that they choose, the gameplay mechanics that they program, and how they make use of sprites and art assets to create the game aesthetics), and the extent to which their game design impacts learning outcomes. Such knowledge may help us understand how students approach game design in constructionist curricula, as well as reflect on how these games could be used as "assessment tools" by constructionist educators.

Here, we look at games designed by students to meet the educational goals of a game-based constructionist STEM curriculum. In this curriculum, students were tasked to design "serious games", i.e. games that are intentionally designed with a purpose beyond entertainment (e.g., education, marketing) [45]. We analyze 391 student-designed games in Scratch [95], and explore how students transform learning content (here climate science [14, 90, 91, 116]) into serious games, as well as how they convey educational content to players through the gameplay. These games designed by dyads of 8th-grade students as part of a three-years STEM curriculum[1] on climate science and systems thinking.

First, we analyze 391 games to identify common design trends among student games. Specifically, we use the serious game design framework *Triadic Game Design* (TGD) [45] to explore how the games articulate *reality*, *meaning*, and *play*, and to what extent these three are "balanced" to help reflect content uptake. The TGD framework helps systematize the analysis of a large dataset (i.e., 391 games), and its results provide an overview of (1) design trends in student games and (2) how and why a student game can be deemed as "balanced" or "imbalanced". Second, we run an in-depth analysis with both a game design and an education expert. We have the two experts look at a subset of the 20 most representative games (i.e., 10 "balanced" and 10 "imbalanced") from the whole dataset (i.e., 391), and ask them to provide further design and pedagogical insights.

With our analyses, we provide empirical considerations that align with theoretical underpinnings supporting constructionist perspectives in game-based education [88], and help demonstrate the effectiveness of constructionist teaching and learning [63]. Also, in formalizing "balance" in games via TGD, we speak to research concerned with uncritical relativism [39] and the lack of "rigorous assessment strategies" in constructionist teaching (see [120], p. 139).

2 GAME-BASED CONSTRUCTIONIST LEARNING

The use of games as educational tools has a long tradition. Notably, in the early 80s Seymour Papert and his team at MIT introduced LOGO [82, 93], a programming language used by young students to learn creatively through the design and programming of video games and digital artifacts [81]; LOGO has recently evolved into Scratch [95]. Nowadays, game-based learning approaches [3, 92, 113] are spreading across modern school curricula, where young students learn by either *playing* [35] or *designing* games [89]. As such, game-based learning approaches can be divided in:

Instructionist (or *play-centric*) approaches [35], where students play *serious* games [9, 53, 66, 97, 108] designed by expert game designers [50] to learn a variety of topics, including health education [80], mathematics [57], and problem-solving [102].

Constructionist approaches [51, 54, 111], where students "construct" knowledge [86] while creating (often digital) artifacts [52, 81, 94], and become designers themselves.

While the benefits of constructionist approaches to learning are well-known [81, 93, 98], previous work pointed out how instructionist approaches are still preferred in game-based curricula [53]. However, modern classrooms are increasingly integrating constructionist approaches [51, 52, 54, 55, 94] within scholar activities, where young students design video games to learn about, amongs others, computer science [78], mathematics [57], and climate science [14, 90, 91, 116]. In this paper, we focus on the constructionist approach to game-based learning, one that leverages game design to learn (e.g., [6, 14, 115, 116]) and that is referred to as game-based constructionist learning [115] or as constructionist gaming [54]. With this focus in mind, we provide background on the use of game design as learning tool for young students next.

[1]https://external-wiki.terc.edu/display/SCRATCH/

2.1 Game Design as Learning Tool

Game design is promising for effective learning in modern education [29, 55, 60]. Among its benefits, it offers a unique opportunity for effectively engage students with the learning of programming [103] and computational thinking (CT) [38, 70–72, 116, 117], which are key 21st century skills [31, 36]. For instance, previous work [22] effectively taught Object-Oriented Programming (OOP) to students through game design, showing how it keeps learners engaged and helps broadening OOP skills. Yulia and Adipranata [124] showed how through game design students with different programming backgrounds can be equally engaged in. They showed how students who are proficient at programming can develop complex games to challenge their skills, while less proficient students can develop simpler games and consolidate the basics of programming. The use of game design can also encourage students to re-elaborate learning concepts based on their personal understanding (cf. [57]), and make them feel engaged, empowered, and proficient. Recently, research on CT has shown that game design leads to higher CT proficiency compared to other design practices [69], and that game genre has impact on how students develop in CT [115]. In sum, the use of game design as tool for learning complementary skills has been widely investigated. However, previous research has under-considered the design outcomes produced by students when designing games in constructionist curricula, as well as what design outcomes may best reflect content uptake.

2.2 Evaluating Student-Designed Games

Evaluating artifacts is challenging [114], as they often involve measuring creativity, which is hard to quantify [87]. Previous work, however, has proposed systematic approaches for evaluating creative artifacts, including domain-independent metrics [33] and machine learning techniques [106]. In the context of constructionist learning, *educational data mining* [17] and *learning analytics* [41] are emerging as promising for evaluating creative artifacts vis-à-vis learning outcomes. However, few studies have attempted to evaluate creative artifacts in game-based learning. In this paper, we align with two studies involving experts in student-designed games evaluation. Iacovides and Cox [49] proposed a mixed-method approach, where four educational games were evaluated through experts judgment, playtesting, gameplay observations, questionnaires, interviews, and online surveys. A similar approach was used by Baytak and Land [16], where three student-designed games were evaluated also through interviews, observations, and play-testing. However, compared to the studies mentioned above, we (1) conduct our analysis based on an established game design framework (i.e., TGD [45], see next section) and (2) evaluate a much larger dataset (i.e., 391 games).

3 METHOD

We qualitatively analyze 391 designed by dyads of 8th-graders to identify trends among students' design outcomes (e.g., gameplay mechanics, aesthetics), and their strategies to frame the learning content and convey it to players. We use a grounded theory approach [37] via consensus coding [96], as well as the Triadic Game Design (TGD) framework [45] to analyze and evaluate (1) how students transform climate change topics (e.g., global warming) into games, and (2) how the three TGD constructs of *reality*, *meaning*, and *play* are articulated by students to "balance" game design and learning content. Next, we describe (1) the TGD framework, (2) our dataset, and (3) our qualitative analysis.

3.1 Triadic Game Design

The TGD framework [45] postulates that (1) any serious game is based on three main constructs, namely *reality*, *meaning*, and *play*, and that (2) by emphasizing these three constructs when analyzing serious games one can asses how "balanced" a game is with respect to its intended goals (here, teaching players about climate science). In that respect, the TGD framework is similar to other design and evaluation frameworks for serious games, which emphasize the need for *balancing*, for instance, pedagogical (i.e., meaning) and gameplay constructs in serious games (see [45], p. 87). The TGD framework describes reality, meaning, and play as follows:

Reality - The connection between the game world and the real world; the TGD framework suggests that any game contains (to a certain extent) an underlying model of reality, often deployed through the representation in-game of real objects (e.g., cars), or through the implementation of real-life physics and mechanics (e.g., it is possible to jump "over" things but not "underneath" them). When looking at games from this perspective typical questions are "How does it work in the real world?" or "Is this realistic?".

Meaning - Represents the underlying message or purpose of a game and how this is accomplished in a game format. What distinguishes serious games from entertainment games is their intention to achieve a "serious" purpose such as changing a behavior (e.g., exercising, see exergames [83]), collecting behavioral data to learn about human behavior (e.g., [68]), or to educate players. When looking at games from this perspective typical questions are "Is this effective?" and "How does this achieve our goal?".

Play - Pertains to the game design elements (or outcomes), such as the gameplay mechanics, the game genre, or the challenges that players need to overcome as they play a game (e.g., solving puzzles). When looking at games from this perspective typical questions are "Is this fun or engaging?" and "How user-friendly is this?".

Then, a "balanced" serious game is one that appropriately considers each construct, i.e. a game that is realistic, achieves its intended purpose, and is engaging. Balancing requires designers to carefully consider each construct, make trade-offs, and find ways in which the different design components reinforce each other.

In our analysis, we chose TGD as the analytical framework to analyze student games for three main reasons. First, we analyze student games that are serious games about climate change, thus favoring a framework like TGD that focuses on serious games. Second, we look at the design elements of student-designed games to reason about them as "evidence" of learning; the TGD framework focuses specifically on capturing and analyzing such design elements (i.e., constructs) in serious games, unlike other frameworks, which, for example, focus on measuring engagement [8], learning outcomes [32], or support and guide the game design process [25].

3.2 Curriculum and Student-Designed Games

The games were designed by dyads of 8th-grade students as part of an innovative STEM curriculum, which focused on the learning of climate science, CT, and systems thinking. The curriculum was implemented and carried out by TERC and Northeastern University, with the help and collaboration of four middle schools in the greater Boston area. In short, the students were instructed about various phenomena that surround climate science (e.g., global warming, CO_2 emission), and were then asked to design serious video games in Scratch to creatively represent (and re-elaborate) the content that they had learned. The curriculum was structured as follows:

(1) First, the teachers introduced and instructed their students on different climate change topics using *Power Point* slides; the first classes allowed students to learn and consolidate knowledge of climate science before moving to mind-mapping and brainstorming for designing their game.
(2) Then, the teachers tasked students to make mind-maps that represent climate change phenomena and their explicit relationships as cause-and-effect (e.g., fossil-based fuels increase CO_2 emission and lead to global warming)
(3) Finally, the teachers let students play and critique existing serious video games on climate science (e.g., NASA's *Offset* [58], and *Power Up* [59]) and then introduced them to Scratch; before designing their own video game, students were prompted to program their first Scratch project through a "10-blocks challenge", where students were allowed to use only ten Scratch blocks to build a simple application.

Once the teachers ensured that students had a solid understanding of (1) climate science phenomena and (2) how to program in Scratch, they allowed them to design their games in dyads. During the design process, the teachers assisted the students in developing their games through focus groups and by supervising their design iterations. At the end of the curriculum, the dyads of students had to present the finalized games in front of their peers and teachers. The curriculum was implemented in more than 50 science classes over three years at three different schools by nine teachers, and involved several 8th-graders (13 to 14 years old, approximately 19 students per class). The total duration of the curriculum was three weeks, with students working on their game design projects an amount of hours (in class) relative to four days of work.

3.3 Data Collection (Scratch Video Games)

The 391 video games were designed and programmed by dyads of 8th-grade students in Scratch and uploaded on scratch.mit.edu with a unique ID. All student-designed video games were also stored in an external repository from which we could retrieve the compressed .sb2 files for our analysis.

3.4 TGD Analysis

We qualitatively analyzed the student-designed games using a combination of TGD and a grounded theory approach [37], particularly emergent and consensus coding [67]. First, we labelled the *game genre* of student-designed games (e.g., *storytelling*) by referring to Heintz and Law's [46] classification of game genre in HCI as the basis for our coding. We discussed possible genre labels focusing on (1) game dynamics and (2) the goal of the game; for instance, a fast-paced game where the scope is to avoid obstacles, and/or shoot at targets, would be labeled *action*. Two authors with expertise in games and HCI started the first round of coding, and independently coded half of the dataset (i.e., ~195 games) to identify game genres. After the first round of coding, the two authors identified six game genres: (1) *action*, (2) *simulation*, (3) *puzzle*, (3) *strategy*, (4) *quiz*, (5) *adventure*, and (6) *storytelling*. However, as suggested by Rolling and Adams [99], game genres should not be labeled just on the basis of game dynamics and content, but also on the type of challenges that they provide. Hence, we started considering how student-designed games intended to challenge players through the gameplay mechanics, and refined our labels in a second round of coding; for instance, a game where the main challenge is to jump on platforms and avoid enemies, much like *Super Mario Bros* [76], would be labeled *platform*. Two authors repeated the coding procedure over the entire dataset and considered gameplay challenges to finally identify 10 game genres: (1) *clicker*, (2) *maze*, (3) *platform*, (4) *pong*, (5) *puzzle*, (6) *quiz*, (7) *shooter*, (8) *simulation*, (9) *storytelling*, and (10) *swipe elimination*; the emerging game genres are described into details in section 4.1. After labeling the game genres, we analyzed the games based on TGD [45]. We focused the TGD analysis on how students articulated the three constructs of reality, meaning, and play in their games about climate science. For each TGD constructs, we looked at the following qualitative characteristics in the games:

- For *reality*, we looked at how the games represent the reality of climate science [90, 91] and deploy systems thinking [21]. In particular, we considered (1) how the art assets in Scratch were used to re-construct the reality of climate change and situate players in it, and (2) the system complexity demonstrated in the games (e.g., a game with several feedback loops and complex human-to-environment interactions would be deemed as *highly complex*).
- For *meaning*, we looked at what strategies (e.g., persuasion, lecturing) the game used to convey the educational message to players and to inform them about climate change.
- For *play*, we looked at the gameplay characteristics (e.g., is the game engaging, complete, and does it have a clear end goal?); furthermore, we considered whether the games would suffice on the gameplay mechanics alone to deploy climate science, or if they relied much on debriefing players about climate science via textual explanations.

Also, we coded whether the three TGD constructs were well-balanced in the games, and deemed the games as *balanced* if all three constructs were appropriately considered, or as *imbalanced* if one of the three constructs was missing or tackled insufficiently. Based on the criteria explained above, two authors started analyzing the whole dataset and developed a coding scheme to rank the student-designed games on all three TGD constructs. The coding scheme was based on a trichotomous *negative* (-), *neutral* (=), and *positive* (+) judgment metrics, where a game would be scored + if all three TGD constructs were deemed as well-addressed, = if at least one of the TGD construct was ranked positive, and - if none of the TGD constructs was considered sufficiently tackled. In the first round of coding, two authors coded a subset of games (i.e., 100 games) in pair to reach consensus on which games should be considered +, =, or -. Finally, the two authors independently coded the rest of the dataset (e.g., ~145) in the second and final round of coding.

3.5 Expert Game Analysis and Evaluation

We further analyzed the games with the help of two experts (1) a game design expert with over 10 years experience designing serious games, and (2) a climate science expert with over 20 years experience in education. The rationale for involving experts in our game analysis was threefold: (1) validate our own analysis based on the opinion of experts (similar to [71]), (2) gather deeper design and pedagogical insights into student-designed games, and (3) use such expert's insights to inform research and assessment practices in game-based constructionist education.

For the expert analysis, we selected a subset of student-designed video games because having the experts evaluate the entire dataset would have been time-consuming, may have yielded unfocused insights, and objectively impractical. Hence, prior to the expert analysis, we identified a subset of 20 games which were best representing the breadth of design outcomes produced by the students in their game design: two examples for each game genre, one *imbalanced* and one *balanced* game. For the *balanced* games, we discussed the games ranked +, and selected 10 examples that we deemed most representative; for the *imbalanced* games, we made sure to select at least one game for each degree of imbalance observable, namely five games ranked - and five ranked =.

First, we explained our coding procedure in details to the experts, including the scoring metrics to rank the games. Then we provided each expert with a Google Sheet that included (1) 20 links to the selected student-designed games, (2) the games/projects IDs, (3) the games' names, (4) the game genres, (5) three columns for ranking the TGD constructs separately, (6) a column for the overall ranking of the games, and (7) a column to put detailed notes/descriptions of the games' evaluations. We had each expert perform their coding and evaluation separately. After the experts had completed their analysis, we had a group discussion with them to reach consensus on their coding and evaluation of the games; we were able to reach consensus within the first session of group discussion.

4 RESULTS

First, we provide context by describing the 10 game genres emerging from our analysis. Second, we provide an overview of the main design trends among student-designed games.We describe the main trends for (1) the design of game mechanics, (2) use of graphical elements (i.e., sprite and backdrops in Scratch), and (3) strategies employed to convey educational content about climate science to players. Finally, we describe the results of the experts analysis. Notice that, a part of these results extend findings of our previous research [115], and thus some overlap may be observable.

4.1 Context: Game Genres

Below, we explain the definition of each game genre into details, based on their content (i.e., climate science) and gameplay mechanics (i.e., in-game challenges for the players).

Clicker games require players to simply click (through a touchscreen, keyboard, or mouse) to perform actions in the gameplay, including advancing scenes, destroying objects or selecting virtual objects. We found at least 94 (i.e., 23.3% of the whole dataset) clicker games in our dataset.

In **Maze** games the scope is to navigate within spaces restricted by walls and other obstacles to eventually find a way out. Generally, the maze games from our dataset would task players to find the way out of a maze to help reduce CO_2 emission, save the forest, or reduce global warming. We found 32 (i.e., 7.9%) maze games.

Platform games have game mechanics similar to the popular Nintendo® game *Super Mario Bros* [76], where the scope is to jump between platforms, avoid or kill enemies, and gather special objects to progress through the gameplay. Examples of platform games from our dataset are *Jump Into Climate Change!*, where the player controls an in-game character that rides a bike and the scope is to avoid colliding with objects that pollute the environment (e.g., TVs, cars). We found 23 (i.e., 5.7%) platform games.

Pong games resemble the popular Atari™ video game *Pong* [10]. The gameplay is simple: the player(s) control a paddle to deflect a ball to gather points when competing against other players or the computer. *Albedo Pong* (Figure 1d) is one such game included in our dataset, where the paddle becomes a block of ice, which will slowly melt if the sun rays enter the surface of the ocean. Pong games are the majority in our dataset (i.e., 96 games or 24.5%).

Puzzle games challenge players with riddles that need solving to progress through the gameplay. *Comfortable Climate* is an example of student-designed puzzle game, where the player moves between various rooms of an apartment to find and turn off as many appliances that help increase CO_2 emission. We found 17 (i.e., 4.2%) puzzle games in our dataset.

Quiz games are based on a simple question-and-answer gameplay structure. Overall, quiz games from our dataset task players to respond correctly to questions about climate science, such as choosing between fruits or meats, or between riding a bike or driving a car; the result of the quiz will teach players which choices are eco-friendly and which ones are not. We found 50 (i.e., 12.4%) quiz games in our dataset.

Shooter games challenge players' eye-hand coordination with moving targets that must be shoot to advance in the gameplay. Most student-designed shooter games in our dataset were inspired by games like *Space Invaders* [109] or *Aero Fighters* [110]. Shooter games were among the fewest in our dataset (i.e., 11 games or 2.8%).

Simulation games emulate real-life in virtual settings. For instance, in simulation games players are tasked to perform actions like driving cars or flying airplanes (see [45], p.75-76). However, in our dataset simulation games focus on simulating real-life situations that involve climate change, like in *Government Simulator* (Figure 2h), where players act like politicians to decide between increasing economical wealth but harm the environment or improve sustainability to save the environment. We found 22 (i.e., 5.5%) simulation games in our dataset.

Storytelling games emphasize stories and their narrative, which players navigate through to learn about particular topics or stories. In our dataset, student-designed storytelling games like *Extreme Climate* teach players about how CO_2 emission contributes to atmospheric pollution, hence help causing global warming. Storytelling games represent 6.2% of the entire dataset (i.e., 25 games) .

Swipe Elimination games are popular in mobile gaming (see *Fruit Ninja* [43]); they are based on the gameplay mechanics of swiping on the screen to delete, chop, or destroy in-game objects. *Extreme Weather* is an example of swipe elimination game from our dataset, where players swipe-away CO_2 clouds from the sky to "clean" the atmosphere. We found 21 (i.e., 5.2%) swipe elimination games in our dataset.

4.2 TGD Rankings

Based on our TGD analysis, we ranked 270 games (i.e., 69%) as imbalanced in all three TGD constructs (i.e., ranked as -); this means that the majority of student-designed games in our dataset were either incomplete or poor in their game design, not representing the reality of climate change sufficiently, and not conveying the educational content in a clear way. We found 93 (i.e., 24%) to be somewhat balanced (i.e., ranked =); this means that such games tackled at least one TGD construct sufficiently (e.g., the game had engaging gameplay mechanics but the educational content was not delivered to the player). Finally, we ranked 28 (i.e., 7%) as balanced in all three TGD constructs (i.e., ranked +).

However, besides our TGD ranking, we found that all student-designed games had merit, and since these games were also meant to be educational (i.e., serious games), we acknowledge that designing such games successfully, while also learning climate science in parallel, might have been an arduous task for K-12 students. As such, our rankings should not be interpreted as reflective of students' design capabilities, but rather as heuristics that can reflect why and where such games are balanced (or imbalanced), and possibly use these heuristics to help both educators and students in game-based constructionist curricula.

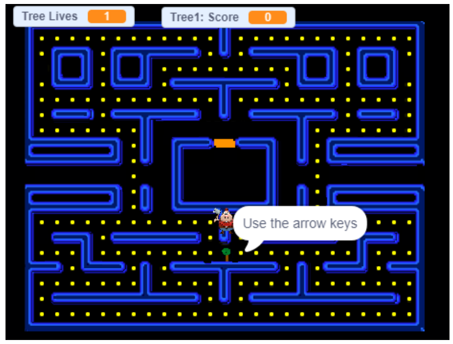
4.3 Design Outcomes from Student Games

Overall, the games showed similar design trends and outcomes. Below, we summarize the main trends we identified using the 20 games that we deemed as the most representative of the entire dataset, which we also used for the expert analysis (see Table 1 and Figure 1 and 2). As we used the TGD as our analytical framework, we will describe the design outcomes of student games following the reality, meaning, and play structure of TGD.

4.3.1 Reality. The student-designed games showed various levels of complexity in their climate systems representations (i.e., potentially hinting to students' uptake of systems thinking [20, 47, 91, 123]), going from *simple* to *highly* complex climate systems. Furthermore, they showed three main uses of art assets (or sprites) in Scratch, which aimed to transform the reality of climate science (e.g., CO_2 molecules) into game aesthetics (e.g., CO_2 "invaders").

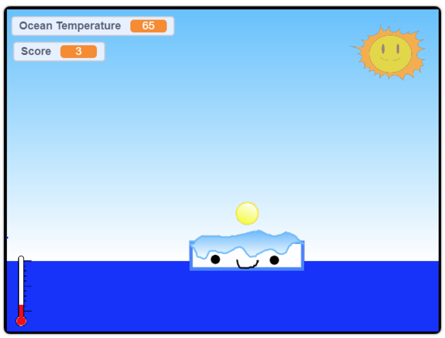
Climate Phenomena as Playable Systems – Puttick et al. [90] describe three complexity levels through which students deploy systems thinking in game design, i.e. *simple*, *moderate*, and *high*. In our dataset, *Albedo Pong* (Figure 1d) and *Save the Animals!*

(a) Plants vs. Zomseas



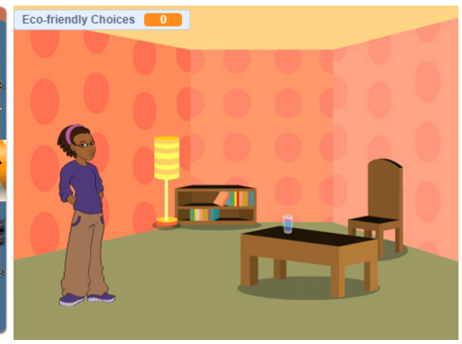
(b) CO2 Pacman

(c) Adventures of TreeGuy

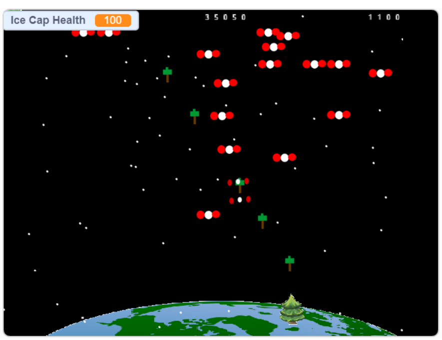


(d) Albedo Pong

(e) Carbon Crusher

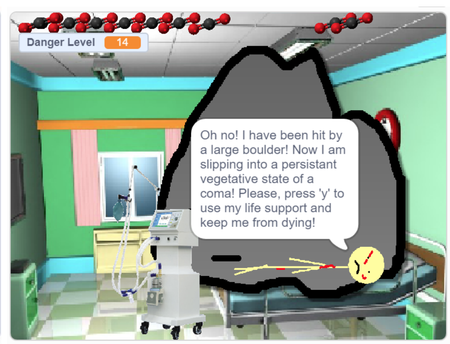


(f) Eco-Friendly Choices

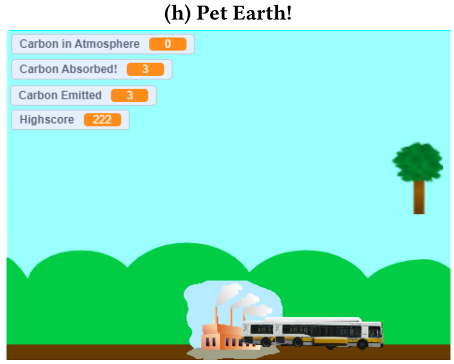


(g) Ice Cap Hero!

(h) Pet Earth!



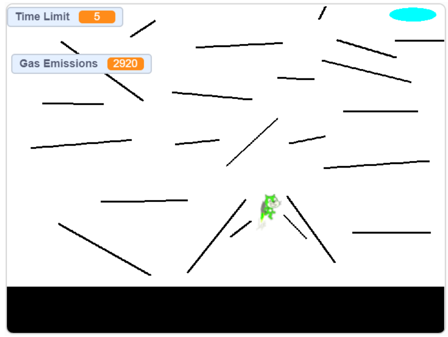
(i) Carbon Clicker



(j) Save the Atmosphere!

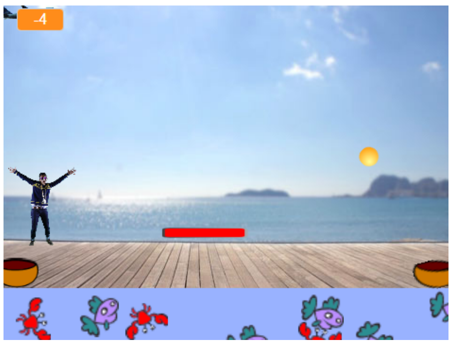
Figure 1: The ten most representative student games that were deemed "balanced".

(a) Defrostation



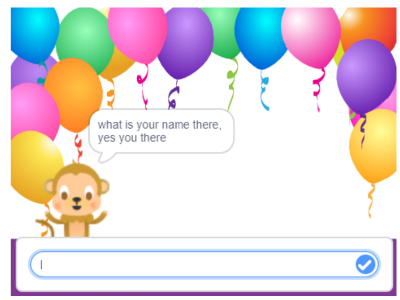
(b) Jetpack Cat

(c) GAMEGAMEMEGAMEGA



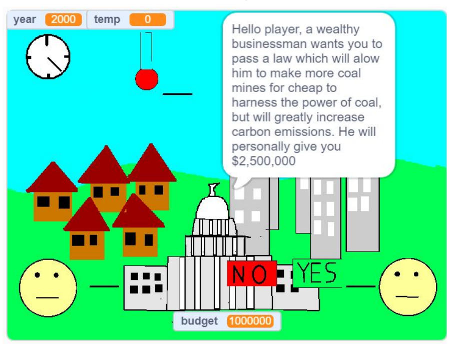
(d) Save the Animals!

(e) Climate Change Sort Game



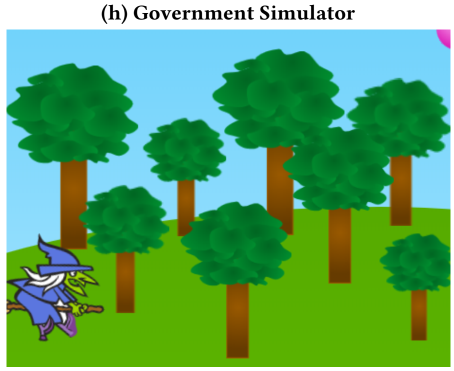
(f) Monkey & Balloons

(g) Climate Change Game



(h) Government Simulator

(i) Gerald the Squirrel



(j) Carbon Witch

Figure 2: The ten most representative student games that were deemed "imbalanced".

(Figure 2d) are examples of *simple* systems complexity. In these games, the climate phenomenon of ice-albedo feedback [62] is represented by the cause-effect of four interconnected climate elements (i.e., sun ray-hits—>ocean-warms-up—>ice-melts—>raises-ocean level), which are directly mapped to game mechanics; the player controls an ice paddle, which shrinks when a sun ray (represented by a light-yellow ball) strikes the surface of the ocean, and this will raise the temperature and provoke the ice to melt.

Carbon Clicker (Figure 1i) instead deploys a *moderate* system complexity; this is a storytelling game, which teaches the player how using fossil-based energy to sustain power plants, fuel cars and aeroplanes, and power medical technologies (e.g., respirators) all interconnect and contribute to CO_2 emission, global warming, and will eventually destroy planet Earth (e.d., "in a big explosion"). Compared to *Albedo Pong*, *Carbon Clicker* displays a *moderate* systems complexity because, although not representing feedback loops, it represents a more complex relationships of cause-and-effect in climate systems (e.g., choice of energy source, use of technology, pollution, effect on Earth atmosphere), and accounts for the human influence on climate change.

Finally, *Government Simulator* (Figure 2h) has a *high* system complexity. The game simulates a real-life scenario, in which politicians have to make decisions regarding budget management, investing in various energy resources, building new factories, all of which will have an (either positive or negative) impact on both people's lives and the environment. This game represents multiple climate systems and the impact of human decisions on those, similar to *Carbon Clicker*. However, as players can influence human decisions in *Government Simulator*, this game shows an even more complex overview of cause-and-effect relationships in climate change compared to *Carbon Clicker*.

The Default, The Mix-And-Match, and The Custom – The aesthetics of student games seem to leverage three main approaches: (1) using the *default* art assets (i.e., sprites, backdrops) from the Scratch library, (2) using a *mix-and-match* between Scratch art assets and drawn or realistic pictures, or (3) using completely *customized* art assets for creating unique aesthetics.

The game *GAMEGAMEMEGAMEGA* (Figure 2c) uses *customized* art assets to create cartoon aesthetics similar in style to the arcade games from the 80s, for instance like *Super Mario Bros* [75] and *Donkey Kong* [74]. This game achieved good results at creating custom aesthetics that are pleasing and seamless; however, not all games in our dataset that used custom art assets achieved good results. Another approach to design aesthetics we found in student-designed games was a *mix-and-match* between the default Scracth art assets and customized ones. For instance, *Defrostation!* (Figure 2a) and *Save the Animals!* (Figure 2d), super-impose Scratch sprites (e.g., characters, objects) to realistic pictures (e.g., a forest, an ocean, etc.) to create seamless game aesthetics; this strategy shares similarities with *Augmented Reality* applications (see [11] for a review), where 3D or cartoon pictures are super-imposed to real environments to create immersive user experiences.

Finally, games like *Monkey and Balloons* (Figure 2f) and *Jetpack Cat* (Figure 2b) use *default* sprites only to create the games' aesthetics. While games' aesthetics may be a hint to balance in a student game, it is important to notice that we found some games with *default* aesthetics to be balanced, and some with *customized* aesthetics to be imbalanced. As such, the extent to which aesthetics can be used as a heuristic or an early identifier of balance in student-designed games is unclear and requires further scrutiny.

4.3.2 Meaning. We followed the work of Barnes et al. [14] in our analysis to identify two ways in which students convey climate science to players, i.e. by (1) *lecturing*, or by (2) *raising awareness*.

Lecturing on Climate Science – An approach to convey climate science that we observed in student games was the "interactive lecture" approach, namely games designed to educate players mainly through debriefing, textual explanations, or question-and-answer (Q&A). Not surprisingly, we found such an approach to be common in student games that are either storytelling or quiz (e.g., *Eco-Friendly Choices*, *Gerald the Squirrel*; Figure 1f and 2i), which game genres are notably text-heavy. *Monkey & Balloons* (Figure 2f), for instance, asks players to answer questions about climate science (e.g., what's a cause for climate change?) through multiple choices (e.g., A: too much CO_2 in the atmosphere; B: digging up the ground; C: flying aircraft into space), of which only one is correct (i.e., A); the game keeps asking the same question until the player provides the correct answers, and then moves on to the next question.

Most quiz games in our dataset that lectured players were similar to the above; however, other quiz games only revealed the correct answers at the very end of the game, and required the player to re-play the game if they wished to provide the right answer. *Carbon Clicker*, instead, does not give agency to players to choose between a right or a wrong answer; rather, the game progresses linearly as the player simply clicks to move from the previous scene of the game onto the next one. This somewhat "interactive" gameplay resembles how one would interact with a Power Point lecture, simply clicking to move from one slide to another.

Raising Awareness to Alter Players' Behavior – Differently from the games described in the subsection above, the games designed by students to raise awareness about climate change urged players to make decisions, take actions, and modify climate scenarios while playing, with the scope of letting players reflect on their behaviors in real-life and possibly change them for the better (e.g., developing awareness about fossil-fuel cars polluting the atmosphere and increasing CO_2 emission—>use bikes rather than cars to help reduce CO_2 emission). We identified three main strategies used by students in their games to raise awareness in players: (1) *procedural rhetoric* [121], (2) *reverse psychology* [104], and (3) *activism* [101, 112]. In the work of Bogost [18], procedural rhetoric in games is described as a "persuasive practice" through which game designers have players make specific decisions, which are not forced, but rather "provoked" indirectly. For instance, the game *Pet Earth!* (Figure 1h), asks the player to dress up, feed, and wash planet Earth in a *Tamagotchi*-like game. Here, players assume that they are giving good service to Earth, whose innocent look reminds of a cute pet. However, as clothing, food, and water represent massively-consumed resources in the context of the game, instead of giving good service to Earth, the player ends up destroying it for indiscriminate resource consumption.

Pet Earth! is an example of a game that uses both procedural rhetoric (i.e., persuades the player to feed and wash the Earth), as well as reverse psychology (i.e., the actions of feeding and washing result in destroying Earth, which is not the outcome expected by players) to convey climate science to players. In *Pet Earth!*, procedural rhetoric is used in a subtle way; however, other games in our dataset were more explicit in their persuasive strategies, where players making decisions that harm the environment (e.g., building power plants) would be "punished" by either cutting their in-game budget or by lowering the morale of the population (in *Government Simulator*). Finally, games like *Save the Atmosphere!* (Figure 1j) leverage the dialectic of climate activism (e.g., [112]), to raise awareness about environmental challenges by letting players make proactive decisions (e.g., swap nuclear plants with solar energy-powered plants) to help save the environment.

4.3.3 Play. Most games designed by students re-purposed the game mechanics of existing video games, and often changed the original in-game elements (e.g., aliens in *Space Invaders* [109]), with climate science ones (i.e., aliens become CO_2 molecules in *Ice Cap Hero!*; Figure 1g). Many games in our dataset made consistent use of textual explanations to teach players climate science. As explained earlier, these games felt more like a "debriefing" or a "lecture" on climate change, rather than an interactive re-adaptation in form of a video game. Games like *Albedo Pong*, *Pet Earth!*, and *Government Simulator*, however, map the learned content and its educational message directly to gameplay mechanics; for instance, in *Albedo Pong* the ice-albedo feedback loop is cleverly mapped to the gameplay mechanics of *Pong* [10], which game mechanics are already framed within a loop system (i.e., deflect-the-ball—>score-or-opponent-scores—>deflect-the-ball).

While both gameplay design strategies are viable, delivering climate science to players via gameplay mechanics seems a more interesting approach, and also piggy backs on previous work that argue how successful serious games for learning manage to "educate" players mostly through their gameplay mechanics [42].

4.4 Expert Evaluation

The two experts' TGD scores (Table 1) reflected the authors' scores on 10 games (i.e., 50% among 20), while seven games (i.e., 35% among 20) reached some agreement, and three games did not reach agreement (i.e., 15% among 20). As such, we discussed the latter three games, and eventually reached consensus.

4.4.1 Reaching Consensus. The two swipe elimination games, which we initially selected as part of the subset of 20 games, were deemed imbalanced by both experts. As we wanted to have at least one balanced and one imbalanced example for each game genre, we selected three other swipe elimination games, and gave them to the experts for a second round of evaluation. The two experts scored and evaluated the three new selected games and then discussed with us the results. Eventually, they deemed *Save the Atmosphere!* balanced. As such, we replaced one swipe elimination game that was deemed imbalanced with the above one.

Initially, the two experts deemed *Pet Earth!* imbalanced and *Government Simulator* balanced. However, upon a clarifying discussion, both the experts and authors agreed that the *Pet Earth!* better integrates reality, meaning, and play compared to *Government Simulator*, which looks complex and well-made, but in fact presents issues (e.g., unresponsive input, misconceptions when presenting climate science phenomena), which eventually made us and the experts reverse the judgement for those two games. Below, we present the detailed results of the experts' evaluation organized by emergent themes, and report in quotation the insights they shared with us when discussing the games as a group.

4.4.2 Remixing Helps Student Focus on Educational Content. An initial, critical reflection made by experts on the student-designed games raised the following question: were all the games deemed as balanced in fact remixes of existing Scratch games? While to a greater (or lesser) extent, all student-designed games were inspired by the game mechanics of existing video games (e.g., pong, shooter, clicker) some games showed a strong resemblance with existing Scratch projects, which are a replica of well-known games (e.g., *Ice Cap Hero!* remixes a Scratch replica[2] of NAMCO's *Galaga* [73]). Besides *Ice Cap Hero!*, four other games remixed existing Scratch projects: *CO_2 Pacman* (i.e., Pacman[3]; Figure 1b), *Plants vs Zomseas* (i.e., Plants vs. Zombies[4]; Figure 1a), and *Carbon Crusher* (i.e., Candy Crush[5]; Figure 1e). These latter four games represent almost half of the games who were deemed balanced by the two experts. Besides scoring their gameplay positively, the experts found that these four remixed games integrated the climate science content in a smart and thoughtful way, and thus scored them positively on reality and meaning too. For instance, the experts pointed out how in *Carbon Crusher* the reward system from the original game (i.e., *Candy Crush* is cleverly mapped to climate science, where aligning four factories generates a wind turbine, while aligning four barrels of oil generates a power plant (i.e., the scale of the action is proportional to the benefit or reward in climate change). In *Plants vs. Zomseas* the zombies become "zomseas" representing the rising ocean level, which players can "defeat" by building walls and prevent cities from flooding. In sum, the experts acknowledged that remixing is common in creative practices [26], and that by remixing, students can focus more on refining the in-game educational content, as gameplay mechanics are already provided by existing games.

4.4.3 Contextualize and Reinforce are Key to Content Framing. Generally, the experts described the balanced games as *"doing a better job at framing and representing climate science content than the imbalanced ones"*. For instance, they explained how, reminiscent of the edutainment era [77] where content and gameplay where not well-integrated [45], student-designed games like *Monkey & Balloons* show a lack of integration between the climate topic(s) tackled (e.g., CO_2 emission, ice caps melting) and how these are represented in-game. The game has players answering questions that are asked by a monkey, which stands against a background with colored balloons. The experts argued that the use of such "festive" elements, if not properly contextualized, may deceive players and confuse them about the educational message on climate change, because *"climate change is not something we should celebrate"*; their reflection corroborate previous work arguing that game design that is not well-balanced *"will hinder learning"* [27].

By contrast, the experts noticed how in other quiz games (e.g., *Eco-Friendly Choices*; Figure 1f), players are situated in scenes that

[2]https://scratch.mit.edu/projects/11802482/
[3]https://scratch.mit.edu/projects/2345919/
[4]https://scratch.mit.edu/projects/346031626/
[5]https://scratch.mit.edu/projects/15144826/

Table 1: Coding of the 20 selected games by the initial coders, climate/education expert, and serious games expert.

Title[a]	Genre	Initial Coders	Climate/Education Expert	Serious Games Expert	Initial[b] Agreement	Consensus[c]	Remix[d]
Eco-Friendly Choices*	Quiz	+	=	+	Somewhat	Balanced	No
Monkey & Balloons*	Quiz	-	-	-	Yes	Imbalanced	No
Albedo Pong	Pong	+	+	+	Yes	Balanced	No
Save the Animals!	Pong	-	-	-	Yes	Imbalanced	No
Adventures of TreeGuy	Platform	+	+	+	Yes	Balanced	No
GAMEGAMEMEGAMEGA	Platform	-	=	-	Somewhat	Imbalanced	No
C02 Pacman*	Maze	+	+	+	Yes	Balanced	Yes
Jetpack Cat*	Maze	-	=	-	Somewhat	Imbalanced	No
Plants vs. Zomseas	Clicker	+	+	+	Yes	Balanced	Yes
Defrostation*	Clicker	-	-	-	Yes	Imbalanced	No
Ice Cap Hero!	Shooter	+	+	=	Somewhat	Balanced	Yes
Climate Change Game	Shooter	-	=	=	Somewhat	Imbalanced	No
Save the Atmosphere![e]	Swipe	=	NA	+	Somewhat	Balanced	No
Carbon Witch*	Swipe	-	-	-	Yes	Imbalanced	No
Carbon Crusher	Puzzle	+	+	=	Somewhat	Balanced	Yes
Climate Change Sort Game	Puzzle	-	-	-	Yes	Imbalanced	No
Pet Earth!	Simulation	-	=	+	No	Balanced	No
Government Simulator	Simulation	+	+	-	No	Imbalanced	No
Carbon Clicker	Storytelling	+	-	+	No	Balanced	No
Gerald the Squirrel	Storytelling	-	-	-	Yes	Imbalanced	No

[a] Many student games used their first names as title (e.g., Amanda & George). We provided the title for their games in such instances. All the starred (*) titles are games that we named.
[b] The "Initial Agreement" refers to how raters agreed without discussion.
[c] "Consensus" refers to the verdict after discussion.
[d] "Remix" refers to if the game was copied from another Scratch project.
[e] This game was added later because the positively rated selected game by the initial coders was judged negatively by both experts.

clearly contextualize climate change: a house, an office, a public street. Here players take the role of a character (an African-American teenager) who needs to choose between actions that will impact the environment either negatively (e.g., drive a car to work and emit CO_2) or positively (e.g., ride a bike to work and help reduce CO_2 emission). The context where these choices are made both leverage the medium (i.e., the game provides feedback to players about the outcome of their choices), as well as ensure that the player identifies with the in-game character. The experts also discussed instances of imbalanced games, where climate science *"could have made sense, but eventually did not"*. For instance, in *Defrostration* (Figure 2a) players use water to put out fires on trees to save the forest. As recent events in Australia showed [30], wild fires contribute to deforestation; however, the experts did not see this link sufficiently represented by the game, or emerging from either the gameplay or the narrative, and thought the player may interpret the game content as being about disasters in general.

In other games, like *Jetpack Cat* (Figure 2b), only one climate science element is represented, but the context is missing. In this game, players control the Scratch cat equipped with a jetpack through a maze. While using the jetpack to fly, the player will emit gas and thus CO_2 in the atmosphere. While the game does show a climate system (i.e., use jetpack—>CO_2 emissions—>raising temperatures), the context is unclear. The maze, the exit, or the Scratch Cat itself, none of those elements are clearly related to climate change. Notably, the players never see the consequences of their actions. Instead, the experts highlighted how *CO_2 Pacman* (Figure 1b) better integrated climate science in a maze game. The game shows clear relationships between climate science elements, where players control a tree, eat C02 pellets, and have to avoid a woodcutter to prevent deforestation and global warming; and the consequences are immediately visible, as getting caught by the woodcutter will release CO_2 in the atmosphere.

In short, the experts observed how the balanced games found a way to coherently integrate climate science content—both evidenced through the players' choices and actions, the environment they are situated in, and the consequences they observe—whereas the imbalanced once only did so partially, or not at all.

4.4.4 Clear Cause-and-Effect Let Meaning Emerge. As discussed earlier, the experts saw representing the consequences of human actions in games as one way to foster meaning about climate science; if feedback is provided, players are able to link their actions in-game as a cause-effect relationship related to climate change (e.g., choosing to eat meat will increase CO_2 emission, and thus lead to global warming). However, the experts also pointed out how simply including or visualizing such consequences is not sufficient for conveying meaning. Consequences need to be meaningful, namely they need to be *"interpreted by the player in a way that helps to accomplish the purpose the game was designed for"*. For instance, while *Government Simulator* (Figure 2h) is introduced as a simulation game by the students who designed it, its game mechanics are closer to a quiz game, as the main gameplay revolves around players answering "yes" or "no" to policy proposals, where players' decisions will impact budget, temperature, and the approval of citizens and companies. However, the game suffers from providing obvious choices, which consequences—except for the budget—are not clearly communicated to the player. Moreover, the educational content is provided through text, which makes the policy proposals lengthy to read but, more importantly, defeats the purpose of the game. A serious game is powerful for learning when the meaning emerges from the gameplay [42, 45], not by explicitly stating what needs to be learned through text, either at the start of the game, during, or at the end. As such, in their evaluation of *Government Simulator*, the experts agreed that textual debriefing helps clarify climate science concept to players, but recommend that students deliver meaning primarily through gameplay mechanics.

Pet Earth! (Figure 1h), on the other hand, illustrates how meaning can be delivered effectively through game mechanics, and with minimal use of textual debriefing. In this game, players have to take care of their "pet Earth", such as dressing it up, feeding it, and giving it a shower. After going through these "routine" tasks, the screen turns dark and pet Earth asks reflective questions ("Why would you do this to me?"), which players are asked to answer. Then, the game debriefs players on how each action relates to climate change and ends with the statement "I am Climate Change. Thank you for helping me destroy your Earth!" *Pet Earth!* uses multiple strategies identified by experts that together help deliver meaning. First, it uses anthropomorphism [23] by depicting the pet Earth as a human-like character, which strategy has been discussed as successful at creating empathy between players and in-game characters [12]. Second, it asks players to give their pet Earth a name and their pet Earth constantly refers to players as their friend. Third, it is a pet simulation game, so the default game mechanics are about "taking care" of your pet [119]. Fourth, it forces players to reflect on their actions by typing in responses. Fifth, by confronting players afterwards on the consequences of their actions, it makes use of "reverse psychology" (see Section 4.3.2), which is reinforced by the game's closing statement. Thus, the entire experience is carefully designed to deliver its meaning and various strategies and techniques are combined to do so.

4.4.5 Balance may be Already in the Name. The experts highlighted how student-designed games that were balanced clearly communicated their content already in the name/title; they explained that the name of a game is how *"players first identify and relate to the game and will influence how the context of the game is perceived"*. These insights are consistent with previous work with online gaming [24], which discusses how carefully-created character names are key for players to *"develop a persistent, pragmatic identity to maintain social relationships across games and related sites"* (p.1). Notably, six games (i.e., 30% among 20) were not given a name by student/designers, two games used generic names to provide climate context, such as *Climate Change Card Game* or *Climate Change Game*, while one game had a title that did not contextualize climate science (i.e., "GAMEGAMEMEGAMEGA"). It may not be coincidental that these games were all rated negatively, both by us and the experts. Instead, student-designed games that clearly communicated climate science were deemed balanced. For instance, *Albedo Pong* clearly relates game mechanics to the climate topic tackled (i.e., ice-albedo); *Plants vs. Zomseas* cleverly represents the threat of raising ocean levels through a popular horror character (i.e., a zombie); and *Pet Earth!* implicitly suggests that planet Earth is something we should take care of.

5 DISCUSSION

We analyzed 391 student-designed serious games on climate change, and evaluated a subset of 20 games with both serious games and climate science experts. Results from the initial TGD analysis unfolded common trends in design outcomes across student-designed games (e.g., alter players behavior using reverse psychology). The experts' evaluation revealed further insights that help identify effective design outcomes observed from the student-designed games. Next, we discuss implications of our results to student games evaluation and game-based constructionist learning.

5.1 Takes on Students' Games Evaluation

The qualitative evaluations of serious games for education have only recently emerged in game design research [1, 15, 84]. This work, however, focuses on evaluating serious games that are made by game designers for students, thus pertaining to *instructionist* learning [35]. As we focus on evaluating serious games designed by students, our work pertains to *constructionist* learning (e.g., [81]), and thus better aligns with previous evaluations of student-designed games carried out by experts, and in similar contexts [16, 49]. However, our TGD analysis and expert evaluations are, to a certain extent, also in line with emerging evaluations of serious games in instructionist learning [1]. Abdellatif et al. [1], for instance highlight *understandability*, *pedagogical aspects*, *learning outcomes*, and *engagement* (see [1], p. 114) as key aspects of successful and quality serious games. These aspects were also highlighted as characterizing a balanced student-designed serious game by the experts involved in our study. Understandability, for instance, is achieved in *Echo-Friendly Choices* (Figure 1f) by properly situating the player in a believable context, which reinforces the climate change message to the player; in contrast, *Monkey & Balloons* (Figure 2f), with its decontextualized representation of reality, may end up jeopardizing players' understanding of climate science content.

As mentioned in Section 2.2, evaluating student-designed games is an arduous task, for both researchers [16, 49] and educators [114]. To help make student-designed games evaluations more systematic and streamlined, we used multiple levels of depth in our analysis (i.e., initial TGD, then in-depth experts evaluation). By leveraging TGD [45] we could structure both the analysis and evaluation of a large dataset (i.e., 391 games), as well as of a small one (i.e., 20 games), and maintain the same evaluative framework throughout; such "scalability" is not characterizing previous evaluations [16, 49]. While being aware that we are not the first to leverage TGD for student-designed games assessment and evaluation (e.g., [13, 14, 48, 90]), to our knowledge we are the first to use this framework as a basis for developing evaluative metrics (i.e., the -, =, and + ranking system) for student-designed serious games. The ranking system we proposed in this paper is indeed simplistic and less elaborate compared to, for instance, metrics-based approaches that assess CT (e.g., *Dr.Scratch* [70]). However, with our ranking system we took a first step into making qualitative evaluations of student-designed serious games also partially quantifiable.

5.2 Implications for Game-Based Education

In this curriculum, K-12 students were asked to design and program a "serious game" about climate change in Scratch, which would be then played by students a grade below them (aged 8–10) to learn about climate science playfully. Creating serious games is very challenging [45], in large part because it requires designers to integrate the topic's content (Reality) with the game's mechanics (Play) in a way that delivers the message or accomplishes the game's goal (Meaning). This challenge was confirmed in our analysis as we were only able to identify 7% of a total of 391 games as a "balanced" game, i.e. a game that integrates reality, meaning, and play seamlessly to both evidence content uptake and support educational efficacy. This low percentage illustrates the importance to better guide and assist student designers in creating their serious games (see [44]).

While future research needs to ascertain the context (e.g., students' background, classroom setting) or the impact (e.g., learning, motivation) of the design process and its resulting artifact (i.e., game), evaluating the games tells us that a fair amount of students misrepresented a climate science topic or failed to get the message across, which both can be said to be indicative of not (fully) understanding the topic. This is essentially at the core of having students design serious games in the first place: to design a serious game well, one must have a complete understanding of its content [45]. The balanced games, on the other hand, illustrate how students were able to integrate—with care and thoughtfulness—all the elements and mechanisms that make up their games. This can only be done if these students mastered the content.

By looking at the students' design outcomes, and specifically contrasting both the imbalanced versus balanced games, we are able to retrieve a number of key insights that will help in improving game-based constructionist curricula, not only to better guide the student-designers, but also to make such modern learning environments more effective. In particular, there is the ongoing debate—not just with games, but with constructionist education at large—whether the time spent on "making" translates sufficiently (and efficiently) into appropriate learning outcomes [54]. In that respect, the pedagogical considerations emerging from our analysis point to four potential areas of improvement for constructionist educators:

- *Encourage remixing* and allow students to focus on integrating content, rather than spending much time designing the game. However, this may limit creativity;
- Have students thoroughly reflect on how to *integrate reality and game mechanics*, so as to deliver the educational content to players effectively;
- *Provide student with examples*; the students played existing serious games on climate science, but these were complex and designed by professionals. Showing students examples of other student-designed games may help setting the bar;
- *Use heuristics*, such as the name (or title) of a student game, how much text is used in-game, and the game aesthetics, to systematically structure the assessment of student games.

Last but not least, we were interested in understanding the extent to which student-designed games can be used as assessment tools (see [85]). Our results show promise in that regard. Our TGD-based approach provided structure to the experts, and facilitated a systematic analysis and evaluation of student games. We see constructionist educators using a similar approach. They can look at the articulation of reality, meaning, and play in student games as potential evidence of learning, and thus assess their students through the very games that they designed. Also, as already mentioned in the introduction, using the TGD-based metrics in this manner may help mitigate issues pertaining "uncritical relativism" in constructionist education [120].

5.3 Limitations and Future Work

The present work has a number of limitations, which should be tackled in future work. First, we evaluate student-designed games on climate science, a topic that involves several reality elements (e.g., CO_2, deforestation, energy consumption). Therefore, our evaluation puts emphasis on reality representation and contextualization. However, this may be less the case for student games that are designed to tackle different topics (e.g., math [65]); future work should explore our approach in other learning curricula to further validate its results. In game-based learning curricula that leverage game design, students do not always design serious games (e.g., [56]). As such, future work should explore how TGD performs as an analytical framework, also for student-designed games that are not necessarily meant to be "serious" or "educational" (i.e., games for entertainment). In such contexts, the *meaning* construct can focus on the "player experience" [2] the designers intent to achieve. Finally, we outlined a number of pedagogical considerations, and conjectured how both educators and researchers in game-based constructionist learning could be inspired by our TGD approach, to make sense and evaluate student games, as well as reason about the very student games as potential assessment tools. Future research efforts should test these pedagogical considerations to further validate or challenge them.

6 CONCLUSION

The large-scale TGD analysis yielded an overview of design outcomes from games designed by students in a STEM curriculum. These games leveraged existing game mechanics, by means of remixing the ones of well-known games (e.g., Pong); deployed meaning through persuasive strategies (e.g., reverse psychology, clear cause-and-effects); and represented reality at various systems complexity. The expert analysis points to the need of providing more guidance to students that design serious games for learning, given that that only 7% of the games were identified as well-designed (or "balanced"). Based on these insights, we provided pedagogical considerations for game-based constructionist curricula. Our work takes a step forward in evaluating student-designed games and provides inspiration for future research efforts.

ACKNOWLEDGMENTS

This project is supported by the NSF under Grant No. 1542954. We also acknowledge our collaborating teachers, school districts, and students, and thank them for their support throughout our research.

REFERENCES

[1] Abdelbaset Jamal Abdellatif, Barry McCollum, and Paul McMullan. 2018. Serious Games Quality Characteristics Evaluation: The Case Study of Optimizing Robocode. In *2018 International Symposium on Computers in Education (SIIE)*. IEEE, 1–4.

[2] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135 (2020), 102370.

[3] Emmanuel O Acquah and Heidi T Katz. 2020. Digital game-based L2 learning outcomes for primary through high-school students: A systematic literature review. *Computers & Education* 143 (2020), 103667.

[4] Efthimia Aivaloglou, Felienne Hermans, Jesús Moreno-León, and Gregorio Robles. 2017. A dataset of scratch programs: scraped, shaped and scored. In *2017 IEEE/ACM 14th International Conference on Mining Software Repositories (MSR)*. IEEE, 511–514.

[5] Jennifer Albert, Barry Peddycord III, and Tiffany Barnes. 2015. Evaluating Scratch Programs to Assess Computational Thinking in a Science Lesson (Abstract Only). In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education (SIGCSE '15)*. ACM, New York, NY, USA, 679–679. https://doi.org/10.1145/2676723.2691928

[6] A. Amato, Camillia Matuk, Dylan G M Schouten, Steven C Sutherland, Gillian Smith, and C Harteveld. 2020. What do Students Learn about Experimental Research by Designing Interactive Fiction Games?. In *The Interdisciplinarity of the Learning Sciences*, Vol. 1. International Society of the Learning Sciences, Nashville, TN, 553–556.

[7] Alan Amory, Kevin Naicker, Jacky Vincent, and Claudia Adams. 1999. The use of computer games as an educational tool: identification of appropriate game types and game elements. *British Journal of Educational Technology* 30, 4 (Oct. 1999), 311–321. https://doi.org/10.1111/1467-8535.00121

[8] Leonard A Annetta and Stephen Bronack. 2011. *Serious educational game assessment: Practical methods and models for educational games, simulations and virtual worlds*. Springer Science & Business Media.

[9] Leonard A. Annetta, Marshall R. Murray, Shelby Gull Laird, Stephanie C. Bohr, and John C. Park. 2006. Serious Games: Incorporating Video Games in the Classroom. https://er.educause.edu/articles/2006/1/serious-games-incorporating-video-games-in-the-classroom

[10] Atari. 1972. *Pong*. Game [Arcade]. Atari, New York, USA.

[11] Ronald T Azuma. 1997. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments* 6, 4 (1997), 355–385.

[12] Jaime Banks and Nicholas David Bowman. 2016. Emotion, anthropomorphism, realism, control: Validation of a merged metric for player–avatar interaction (PAX). *Computers in Human Behavior* 54 (2016), 215–223.

[13] Jackie Barnes and Casper Harteveld. 2017. When is a game not a Game?. In *Proceedings of the Games Learning and Society Center Conference*. 101.

[14] Jackie Barnes, Amy K. Hoover, Borna Fatehi, Jesus Moreno-Leon, Gillian Smith, and Casper Harteveld. 2017. Exploring Emerging Design Patterns in Student-made Climate Change Games. In *Proceedings of the 12th International Conference on the Foundations of Digital Games* (Hyannis, Massachusetts) *(FDG '17)*. ACM, New York, NY, USA, Article 64, 6 pages. https://doi.org/10.1145/3102071.3116224

[15] Nilüfer Bas, Alexander Löffler, Robert Heininger, Matthias Utesch, and Helmut Krcmar. 2018. Evaluation Methods for the Effective Assessment of Simulation Games. In *International Conference on Interactive Collaborative Learning*. Springer, 626–637.

[16] Ahmet Baytak and Susan M. Land. 2010. A case study of educational game design by kids and for kids. *Procedia - Social and Behavioral Sciences* 2, 2 (Jan. 2010), 5242–5246. https://doi.org/10.1016/j.sbspro.2010.03.853

[17] Matthew Berland, Ryan S Baker, and Paulo Blikstein. 2014. Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning* 19, 1-2 (2014), 205–220.

[18] Ian Bogost. 2008. *The rhetoric of video games*. MacArthur Foundation Digital Media and Learning Initiative.

[19] Chet A Bowers. 1995. *Educating for an ecologically sustainable culture: Rethinking moral education, creativity, intelligence, and other modern orthodoxies*. Suny Press.

[20] Derek Cabrera, Laura Colosi, and Claire Lobdell. 2008. Systems thinking. *Evaluation and program planning* 31, 3 (2008), 299–310.

[21] Peter Checkland. 1999. Systems thinking. *Rethinking management information systems* (1999), 45–56.

[22] W. K. Chen and Y. C. Cheng. 2007. Teaching Object-Oriented Programming Laboratory With Computer Game Programming. *IEEE Transactions on Education* 50, 3 (Aug. 2007), 197–203. https://doi.org/10.1109/TE.2007.900026

[23] Karen Collins and Ruth Dockwray. 2018. Tamaglitchi: A Pilot Study of Anthropomorphism and Non-Verbal Sound. In *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion*. 1–6.

[24] Nicole Crenshaw and Bonnie Nardi. 2014. What's in a name? Naming practices in online video games. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play*. 67–76.

[25] Sabrina Culyba. 2018. *The Transformational Framework: A process tool for the development of Transformational games*. Carnegie Mellon University.

[26] Richard Davis, Yasmin Kafai, Veena Vasudevan, and Eunkyoung Lee. 2013. The education arcade: crafting, remixing, and playing with controllers for Scratch games. In *Proceedings of the 12th international conference on interaction design and children*. 439–442.

[27] Jonathan DeHaan, W Michael Reed, and Katsuko Kuwanda. 2010. The effect of interactivity with a music video game on second language vocabulary recall. *Language Learning & Technology* 14, 2 (2010), 74–94.

[28] Jill Denner, Linda Werner, and Eloy Ortiz. 2012. Computer games created by middle school girls: Can they be used to measure understanding of computer science concepts? *Computers & Education* 58, 1 (Jan. 2012), 240–249. https://doi.org/10.1016/j.compedu.2011.08.006

[29] Darina Dicheva, Christo Dichev, Gennady Agre, Galia Angelova, et al. 2015. Gamification in education: A systematic mapping study. *Educational Technology & Society* 18, 3 (2015), 75–88.

[30] Hannah Etchells, Alison J O'Donnell, W Lachlan McCaw, and Pauline F Grierson. 2020. Fire severity impacts on tree mortality and post-fire recruitment in tall eucalypt forests of southwest Australia. *Forest Ecology and Management* 459 (2020), 117850.

[31] Charles Fadel and Bernie Trilling. 2010. 21st Century Skills: Learning for Life in Our Times. *Education Review // Reseñas Educativas* 0, 0 (2010). https://doi.org/10.14507/er.v0.1296

[32] Aroutis Foster. 2012. Assessing learning games for school content: The TPACK-PCaRD framework and methodology. In *Assessment in Game-Based Learning*. Springer, 201–215.

[33] Celso França, Luıs Fabrıcio W Góes, Alvaro Amorim, Rodrigo Rocha, and Alysson Ribeiro Da Silva. 2016. Regent-dependent creativity: A domain independent metric for the assessment of creative artifacts. In *Proceedings of the Seventh International Conference on Computational Creativity*. 68–75.

[34] Sara I. de Freitas. 2006. Using games and simulations for supporting learning. *Learning, Media and Technology* 31, 4 (Dec. 2006), 343–358. https://doi.org/10.1080/17439880601021967

[35] T. Fullerton. 2006. Play-centric games education. *Computer* 39, 6 (June 2006), 36–42. https://doi.org/10.1109/MC.2006.205

[36] Paul Gilster. 1997. *Digital Literacy*. John Wiley & Sons, Inc., New York, NY, USA.

[37] Barney Glaser and Anselm Strauss. 2000. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Routledge, USA.

[38] Marcos Román González. 2015. Computational thinking test: Design guidelines and content validation. In *Proceedings of EDULEARN15 conference*. 2436–2444.

[39] Mordechai Gordon. 2009. Toward a pragmatic discourse of constructivism: Reflections on lessons from practice. *Educational studies* 45, 1 (2009), 39–58.

[40] Samuel Greiff, Daniel V Holt, and Joachim Funke. 2013. Perspectives on problem solving in educational assessment: Analytical, interactive, and collaborative problem solving. *Journal of Problem Solving* 5, 2 (2013).

[41] Wolfgang Greller and Hendrik Drachsler. 2012. Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society* 15, 3 (2012), 42–57.

[42] MP Jacob Habgood, SE Ainsworth, and Steve Benford. 2005. Endogenous fantasy and learning in digital games. *Simulation & Gaming* 36, 4 (2005), 483–498.

[43] Halfbrick. 2010. *Fruit Ninja*. Game [iOS]. Halfbrick, Brisbane, Australia.

[44] Seungyeon Han and Kakali Bhattacharya. 2001. Constructionism, learning by design, and project based learning. *Emerging perspectives on learning, teaching, and technology. Retrieved April* 29 (2001), 2007.

[45] Casper Harteveld. 2011. *Triadic Game Design: Balancing Reality, Meaning and Play* (1st ed.). Springer Publishing Company, Incorporated.

[46] Stephanie Heintz and Effie Lai-Chong Law. 2015. The Game Genre Map: A Revised Game Classification. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (London, United Kingdom) *(CHI PLAY '15)*. ACM, New York, NY, USA, 175–184. https://doi.org/10.1145/2793107.2793123

[47] Garry Francis Hoban and Garry F Hoban. 2002. *Teacher learning for educational change: A systems thinking approach*. Open University Press Buckingham.

[48] Amy K. Hoover, Jackie Barnes, Borna Fatehi, Jesús Moreno-León, Gillian Puttick, Eli Tucker-Raymond, and Casper Harteveld. 2016. Assessing Computational Thinking in Students' Game Designs. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts (CHI PLAY Companion '16)*. ACM, New York, NY, USA, 173–179. https://doi.org/10.1145/2968120.2987750

[49] Ioanna Iacovides and Anna L. Cox. 2015. Moving Beyond Fun: Evaluating Serious Experience in Digital Games. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, NY, USA, 2245–2254. https://doi.org/10.1145/2702123.2702204

[50] Katherine Isbister, Mary Flanagan, and Chelsea Hash. 2010. Designing games for learning: insights from conversations with designers. In *Proceedings of the*

SIGCHI Conference on Human Factors in Computing Systems. ACM, 2041–2044.

[51] Yasmin B. Kafai. 1996. *Constructionism in Practice: Designing, Thinking, and Learning in a Digital World*. Routledge.

[52] Yasmin B. Kafai. 2005. The Classroom as "Living Laboratory": Design-Based Research for Understanding, Comparing, and Evaluating Learning Science Through Design. *Educational Technology* 45, 1 (2005), 28–34. http://www.jstor.org/stable/44429186

[53] Yasmin B. Kafai. 2006. Playing and Making Games for Learning: Instructionist and Constructionist Perspectives for Game Studies. *Games and Culture* 1, 1 (Jan. 2006), 36–40. https://doi.org/10.1177/1555412005281767

[54] Yasmin B. Kafai and Quinn Burke. 2015. Constructionist Gaming: Understanding the Benefits of Making Games for Learning. *Educational Psychologist* 50, 4 (Oct. 2015), 313–334. https://doi.org/10.1080/00461520.2015.1124022

[55] Yasmin B. Kafai and Yasmin Bettina Kafai. 1995. *Minds in Play: Computer Game Design as a Context for Children's Learning*. Routledge. Google-Books-ID: Ocyllxa8ZjkC.

[56] Marjaana Kangas. 2010. Creative and playful learning: Learning through game co-creation and games in a playful learning environment. *Thinking skills and Creativity* 5, 1 (2010), 1–15.

[57] Fengfeng Ke. 2008. A case study of computer gaming for math: Engaged learning from gameplay? *Computers & Education* 51, 4 (Dec. 2008), 1609–1620. https://doi.org/10.1016/j.compedu.2008.03.003

[58] NASA Climate Kids. 2014. OFFSET! https://climatekids.nasa.gov/offset/. NASA Climate Kids, California, USA.

[59] NASA Climate Kids. 2014. Power Up! https://climatekids.nasa.gov/power-up/. NASA Climate Kids, California, USA.

[60] John Kirriemuir and Angela Mcfarlane. 2004. Literature Review in Games and Learning. https://telearn.archives-ouvertes.fr/hal-00190453 A NESTA Futurelab Research report - report 8.

[61] Praveen Kuruvada, Daniel Asamoah, Nikunj Dalal, and Subhash Kak. 2010. The Use of Rapid Digital Game Creation to Learn Computational Thinking. *arXiv:1011.4093 [cs]* (Nov. 2010). http://arxiv.org/abs/1011.4093 arXiv: 1011.4093.

[62] Jack C Landy, Jens K Ehn, and David G Barber. 2015. Albedo feedback enhanced by smoother Arctic sea ice. *Geophysical Research Letters* 42, 24 (2015), 10–714.

[63] William J Matthews. 2003. Constructivism in the classroom: Epistemology, history, and empirical evidence. *Teacher Education Quarterly* 30, 3 (2003), 51–64.

[64] Merrilea J. Mayo. 2007. Games for Science and Engineering Education. *Commun. ACM* 50, 7 (July 2007), 30–35. https://doi.org/10.1145/1272516.1272536

[65] Camille M McCue. 2011. Learning middle school mathematics through student designed and constructed video games. (2011).

[66] David R. Michael and Sandra L. Chen. 2005. *Serious Games: Games That Educate, Train, and Inform.* Muska & Lipman/Premier-Trade.

[67] Matthew B. Miles and A. Michael Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook.* SAGE.

[68] Omid Mohaddesi, Yifan Sun, Rana Azghandi, Rozhin Doroudi, Sam Snodgrass, Ozlem Ergun, Jacqueline Griffin, David Kaeli, Stacy Marsella, and Casper Harteveld. 2020. Introducing Gamettes: A Playful Approach for Capturing Decision-Making for Informing Behavioral Models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[69] Jesús Moreno-León, Gregorio Robles, and Marcos Román-González. 2017. Towards Data-Driven Learning Paths to Develop Computational Thinking with Scratch. *IEEE Transactions on Emerging Topics in Computing* (2017).

[70] Jesús Moreno-León and Gregorio Robles. 2015. Dr. Scratch: A Web Tool to Automatically Evaluate Scratch Projects. In *Proceedings of the Workshop in Primary and Secondary Computing Education (WiPSCE '15)*. ACM, New York, NY, USA, 132–133. https://doi.org/10.1145/2818314.2818338

[71] J. Moreno-León, G. Robles, and M. Román-González. 2016. Comparing computational thinking development assessment scores with software complexity metrics. In *2016 IEEE Global Engineering Education Conference (EDUCON)*. 1040–1045. https://doi.org/10.1109/EDUCON.2016.7474681

[72] Jesús Moreno-León, Marcos Román-González, Casper Harteveld, and Gregorio Robles. 2017. On the Automatic Assessment of Computational Thinking Skills: A Comparison with Human Experts. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 2788–2795. https://doi.org/10.1145/3027063.3053216

[73] NAMCO. 1981. *Galaga*. Game [Arcade]. NAMCO, Tokyo, Japan.

[74] Nintendo. 1981. *Donkey Kong*. Game [Arcade]. Nintendo, Kyoto, Japan.

[75] Nintendo. 1983. *Mario Bros.* Game [Arcade]. Nintendo, Kyoto, Japan.

[76] Nintendo. 1985. *Super Mario Bros.* Game [NES]. Nintendo, Kyoto, Japan.

[77] Zühal Okan. 2003. Edutainment: Is learning at risk? *British Journal of Educational Technology* 34, 3 (2003), 255–264.

[78] M. Overmars. 2004. Teaching computer science through game design. *Computer* 37, 4 (April 2004), 81–83. https://doi.org/10.1109/MC.2004.1297314

[79] Marina Papastergiou. 2009. Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation. *Computers & Education* 52, 1 (Jan. 2009), 1–12. https://doi.org/10.1016/j.compedu.2008.06.004

[80] Marina Papastergiou. 2009. Exploring the potential of computer and video games for health and physical education: A literature review. *Computers & Education* 53, 3 (Nov. 2009), 603–622. https://doi.org/10.1016/j.compedu.2009.04.001

[81] Seymour Papert. 1980. *Mindstorms: Children, Computers, and Powerful Ideas*. Basic Books, Inc., New York, NY, USA.

[82] Seymour Papert. 1993. *The Children's Machine: Rethinking School in the Age of the Computer*. BasicBooks, 10 East 53rd St.

[83] Wei Peng, Jih-Hsuan Lin, and Julia Crouse. 2011. Is playing exergames really exercising? A meta-analysis of energy expenditure in active video games. *Cyberpsychology, Behavior, and Social Networking* 14, 11 (2011), 681–688.

[84] Giani Petri et al. 2018. A Method for the Evaluation of the Quality of Games for Computing Education. (2018).

[85] Vicki Phillips and Zoran Popović. 2012. More than child's play: Games have potential learning and assessment tools. *Phi Delta Kappan* 94, 2 (2012), 26–30.

[86] Jean Piaget. 1972. *Psychology and Epistemology: Towards a Theory of Knowledge*. Allen Lane. Google-Books-ID: 7DkdAQAAMAAJ.

[87] Davide Piffer. 2012. Can creativity be measured? An attempt to clarify the notion of creativity and general directions for future research. *Thinking Skills and Creativity* 7, 3 (2012), 258–264.

[88] Linda G Polin. 2018. A constructivist perspective on games in education. In *Constructivist education in an age of accountability*. Springer, 163–188.

[89] Marc Prensky. 2008. Students as designers and creators of educational computer games: Who else? *British Journal of Educational Technology* 39, 6 (Nov. 2008), 1004–1019. https://doi.org/10.1111/j.1467-8535.2008.00823_2.x

[90] Gillian Puttick, Jackie Barnes, Giovanni Maria Troiano, Casper Harteveld, Eli Tucker-Raymond, Mike Cassidy, and Gillian Smith. 2018. Exploring how student designers model climate system complexity in computer games. In *Proceedings of the Summit on Connected Learning, CLS'18*. ETC Press, 196–204.

[91] Gillian Puttick and Eli Tucker-Raymond. 2018. Building Systems from Scratch: an Exploratory Study of Students Learning About Climate Change. *Journal of Science Education and Technology* 27, 4 (Aug. 2018), 306–321. https://doi.org/10.1007/s10956-017-9725-x

[92] Joost Raessens and Jeffrey Goldstein. 2011. *Handbook of Computer Game Studies*. The MIT Press.

[93] Mitchel Resnick. 1990. *LEGO/Logo–learning Through and about Design*. Epistemology and Learning Group, MIT Media Laboratory. Google-Books-ID: AEmytgAACAAJ.

[94] Mitchel Resnick. 1996. Distributed Constructionism. In *Proceedings of the 1996 International Conference on Learning Sciences (ICLS '96)*. International Society of the Learning Sciences, Evanston, Illinois, 280–284. http://dl.acm.org/citation.cfm?id=1161135.1161173

[95] Mitchel Resnick, John Maloney, Andrés Monroy-Hernández, Natalie Rusk, Evelyn Eastmond, Karen Brennan, Amon Millner, Eric Rosenbaum, Jay Silver, Brian Silverman, and others. 2009. Scratch: programming for all. *Commun. ACM* 52, 11 (2009), 60–67.

[96] K Andrew R Richards and Michael A Hemphill. 2018. A practical guide to collaborative qualitative data analysis. *Journal of Teaching in Physical Education* 37, 2 (2018), 225–231.

[97] Lloyd P. Rieber. 1996. Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research and Development* 44, 2 (June 1996), 43–58. https://doi.org/10.1007/BF02300540

[98] Judy Robertson and Cathrin Howells. 2008. Computer game design: Opportunities for successful learning. *Computers & Education* 50, 2 (Feb. 2008), 559–578. https://doi.org/10.1016/j.compedu.2007.09.020

[99] Andrew Rollings and Ernest Adams. 2003. *Andrew Rollings and Ernest Adams on game design.* New Riders.

[100] Ricarose Roque, Natalie Rusk, and Mitchel Resnick. 2016. Supporting Diverse and Creative Collaboration in the Scratch Online Community. In *Mass Collaboration and Education*. Springer, Cham, 241–256. https://doi.org/10.1007/978-3-319-13536-6_12

[101] Connie Roser-Renouf, Edward W Maibach, Anthony Leiserowitz, and Xiaoquan Zhao. 2014. The genesis of climate change activism: From key beliefs to political action. *Climatic change* 125, 2 (2014), 163–178.

[102] Elizabeth Rowe, Jodi Asbell-Clarke, Santiago Gasca, and Kathryn Cunningham. 2017. Assessing implicit computational thinking in zoombinis gameplay. In *Proceedings of the 12th International Conference on the Foundations of Digital Games*. ACM, 45.

[103] José-Manuel Sáez-López, Marcos Román-González, and Esteban Vázquez-Cano. 2016. Visual programming languages integrated across the curriculum in elementary school: A two year case study using "Scratch" in five schools. *Computers & Education* 97 (2016), 129–141.

[104] Leon F Seltzer. 1986. *Paradoxical strategies in psychotherapy: A comprehensive overview and guidebook*. John Wiley & Sons.

[105] David Williamson Shaffer. 2006. *How Computer Games Help Children Learn*. Macmillan. Google-Books-ID: ZQqOmEdlSL8C.

[106] Disha Shrivastava, Saneem Ahmed CG, Anirban Laha, and Karthik Sankaranarayanan. 2017. A machine learning approach for evaluating creative artifacts.

arXiv preprint arXiv:1707.05499 (2017).

[107] Stephen Sterling. 2001. *Sustainable Education: Re-Visioning Learning and Change. Schumacher Briefings.* ERIC.

[108] Tarja Susi, Mikael Johannesson, and Per Backlund. 2007. *Serious Games : An Overview.* Institutionen för kommunikation och information. http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-1279

[109] Taito. 1978. *Space Invaders.* Game [Arcade]. Taito, Tokyo, Japan.

[110] Tecmo. 1992. *Aero Fighters.* Game [Arcade]. Tecmo, Tokyo, japan.

[111] Katie Salen Tekinbas, Melissa Gresalfi, Kylie Peppler, Rafi Santo, and James Paul Gee. 2014. *Gaming the System: Designing with Gamestar Mechanic.* MIT Press.

[112] Greta Thunberg. 2019. *No one is too small to make a difference.* Penguin.

[113] Sigmund Tobias, J. Dexter Fletcher, and Alexander P. Wind. 2014. Game-Based Learning. In *Handbook of Research on Educational Communications and Technology.* Springer, New York, NY, 485–503. https://doi.org/10.1007/978-1-4614-3185-5_38

[114] Donald J Treffinger, Grover C Young, Edwin C Selby, and Cindy Shepardson. 2002. Assessing Creativity: A Guide for Educators. *National Research Center on the Gifted and Talented* (2002).

[115] Giovanni Maria Troiano, Qinyu Chen, Angela Vargas Alba, Gregorio Robles, Gillian Smith, Michael Cassidy, Eli Tucker-Raymond, Gillian Puttick, and Casper Harteveld. 2020. Exploring How Game Genre in Student-Designed Games Influences Computational Thinking Development. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3313831.3376755

[116] Giovanni Maria Troiano, Sam Snodgrass, Erinç Arg, Gregorio Robles, Gillian Smith, Michael Cassidy, Eli Tucker-Raymond, Gillian Puttick, and Casper Harteveld. 2019. Is My Game OK Dr. Scratch?: Exploring Programming and Computational Thinking Development via Metrics in Student-Designed Serious Games for STEM. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children* (Boise, ID, USA) *(IDC '19)*. ACM, New York, NY, USA, 208–219. https://doi.org/10.1145/3311927.3323152

[117] David Weintrop, Elham Beheshti, Michael Horn, Kai Orton, Kemi Jona, Laura Trouille, and Uri Wilensky. 2016. Defining Computational Thinking for Mathematics and Science Classrooms. *Journal of Science Education and Technology* 25, 1 (Feb. 2016), 127–147. https://doi.org/10.1007/s10956-015-9581-5

[118] Charlotte Lærke Weitze. 2014. Learning, Education and Games. ETC Press, Pittsburgh, PA, USA, 225–249. http://dl.acm.org/citation.cfm?id=2811147.2811160

[119] Yorick Wilks. 2004. Artificial companions. In *International Workshop on Machine Learning for Multimodal Interaction.* Springer, 36–45.

[120] Mark Windschitl. 2002. Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of educational research* 72, 2 (2002), 131–175.

[121] Will Wright and Ian Bogost. 2007. *Persuasive games: The expressive power of videogames.* Mit Press.

[122] Min Lun Wu and Kari Richards. 2011. Facilitating Computational Thinking Through Game Design. In *Proceedings of the 6th International Conference on E-learning and Games, Edutainment Technologies* (Taipei, Taiwan) *(Edutainment'11)*. Springer-Verlag, Berlin, Heidelberg, 220–227. http://dl.acm.org/citation.cfm?id=2040452.2040499

[123] Susan A Yoon. 2008. An evolutionary approach to harnessing complex systems thinking in the science and technology classroom. *International Journal of Science Education* 30, 1 (2008), 1–32.

[124] Yulia and R. Adipranata. 2010. Teaching object oriented programming course using cooperative learning method based on game design and visual object oriented environment. In *2010 2nd International Conference on Education Technology and Computer*, Vol. 2. V2–355–V2–359. https://doi.org/10.1109/ICETC.2010.5529368