Computers & Education 174 (2021) 104312

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

First inspire, then instruct to improve students' creativity

Seyedahmad Rahimi [a,*], Valerie J. Shute [b]

[a] University of Florida, Gainesville, FL, USA
[b] Florida State University, Tallahassee, FL, USA

ARTICLE INFO

Keywords:
Creativity
Video games
Assessment
Creativity support
Inspiration
Instruction

ABSTRACT

Creativity is one of the most essential skills for success in life in our dynamic, complex world. For instance, we are currently facing major problems with the COVID-19 pandemic, which requires creative thinking for solutions. To increase the pool of creative thinkers, we need tools that can assess and support people's creativity. With advances in technologies, as well as in computer and learning sciences, we can create such support tools. In this study, we investigated the effectiveness of a creativity-support system that we developed in the level editor of an educational game called *Physics Playground*. Our goal was to improve college students' creativity. Participants (n = 114) were randomly assigned to one of four conditions and instructed to create as many creative game levels as possible in about two hours. The four conditions included: (1) Inspirational – with supports that provided access to a website with example levels, a brainstorming tool, and a remote association activity; (2) Instructional – with supports that provided specific instructions to first design as many levels as possible, then pick four of the levels, and enhance them using a tool called SCAMPER; (3) Both – with both inspirational and instructional supports; and (4) No Support, which did not include any creativity supports. The major finding from this research was that the Both condition was significantly more effective than the other conditions in improving students' creativity measured by the creativity of the game levels created by the students. Implications of the findings, limitations, and future research are discussed.

1. Introduction

Throughout history, creativity has been a crucial factor in life-changing inventions propelling civilizations forward (Glăveanu et al., 2020; Hennessey & Amabile, 2010). Recently, the Partnershipfor 21st Century Learning (2019) included creativity as one of the essential skills for success in life in the 21st-century, and the World Economic Forum (Gray, 2016) has listed creativity as the third most important skill (among nine other skills) that people need to be successful in the fourth industrial revolution. In this complex, interconnected world that we live in, societies face new problems that need creative solutions (e.g., dealing with the COVID-19 pandemic). Moreover, in the near future many jobs will disappear (i.e., get automated using AI) except for jobs that need human creativity (a skill that no AI algorithm can replace—at least in the near future) (Belsky, 2020).

The type of creativity we are referring to is not the commonly known type of creativity (i.e., artistic) or a type of grandiose creativity reserved for geniuses such as Leonardo da Vinci. Rather, we are referring to a type of creativity that is related to all aspects of our lives and can be shown by the majority of human beings (Richards, 2010). Such creativity should be valued and fostered by our educational systems to make individuals capable of producing creative ideas, solutions, and products as needed.

* Corresponding author. 2821 Norman Hall, School of Teaching and Learning, College of Education, University of Florida, Gainesville, FL, USA.
E-mail address: srahimi@ufl.edu (S. Rahimi).

https://doi.org/10.1016/j.compedu.2021.104312
Received 14 January 2021; Received in revised form 22 August 2021; Accepted 24 August 2021
Available online 26 August 2021
0360-1315/Published by Elsevier Ltd.

In education, creativity is placed at the highest level of learning in Bloom's taxonomy (Bloom, 1956; Krathwohl, 2002), where "create" comes above other aspects of learning—namely, remember, understand, apply, analyze, and evaluate. However, despite creativity's high level of importance, our educational systems do not actively support creativity, and even worse, on many occasions, creativity is discouraged rather than encouraged (Kaufman & Sternberg, 2007). For example, with the emphasis on standardized tests, both teachers and students could be pushed to a rigid way of teaching and learning rather than trying creative alternatives. When creative teachers and students decide to do something creative, they can be seen as "idealistic and missing the big picture" (Kaufman & Sternberg, 2007, p. 55).

In recent years, researchers have been working to create new methods for assessing and improving creativity (e.g., Barbot, 2011; Kim & Shute, 2015; Plucker & Makel, 2010; Shute & Wang, 2016; Sternberg, 2012). For example, one new and promising areas of research on creativity is using video games (Bowman et al., 2015). Video games are prevalent among children and adults around the world. For example, according to a report from the Entertainment Software Association (2019), 75% of Americans have at least one gamer in their household. Thus, the growing video game industry has much potential for innovative researchers who want to use games as assessment and learning machines (Gee, 2005). However, not all video games can be used for these purposes; rather, just well-designed ones.

Well-designed video games are so engaging that players can spend many hours playing them. Moreover, there are various reasons why well-designed games keep students so thoroughly engaged. For instance, Gee (2005) suggests that well-designed video games are internally motivating to children (and even adults); players can make mistakes and learn from them without being afraid of negative consequences. Moreover, well-designed video games provide various affordances for people to show and improve their creativity. Hall et al. (2021) discussed eight affordances that video games provide for improving creativity: (1) flexibility (i.e., autonomy over the play trajectory), (2) opportunities for creative engagement with the game's narrative (e.g., Bopp et al., 2018), (3) variety of tools and mechanics, (4) the extent to which players can explore the environment, (5) possibilities for creating content (e.g., in-game objects, levels and maps, in addition to customizing the game interface and implementing modifications), (6) the possibility of being creative when it comes to designing an avatar (e.g., Ward, 2015), (7) progression which indicates a match between a player's ability and the game's challenges, and (8) the possibility of replay ability and trying different solutions for a problem at hand. We can categories these affordances into two main categories: (1) creativity in gameplay (e.g., finding a creative solution for a challenging game level or using a creative strategy to avoid losing resources in a game), and (2) creativity in creating game content (e.g., game levels) (Bowman et al., 2015; Gee, 2005). Furthermore, the challenges in well-designed video games incrementally ratchet up difficulty, which helps students build the required skills as they progress through the games (Gee, 2005).

The progression in well-designed video games that lead to the improvement of players' skills can keep players motivated and help them experience the state of *flow* (Csikszentmihalyi, 1997). When one is experiencing flow, he or she loses track of time, becomes fully immersed in the task environment, and performs at his or her best while enjoying the experience. The state of flow can also facilitate creativity whereby ongoing involvement in creative work can enhance one's affective state (Cseh et al., 2015). This occurs via a sense of progress and productivity (Amabile, 1983) which in turn can facilitate and maintain the state of flow (Cseh et al., 2015) which can then help generate even more creative work. Therefore, it is not clear which comes first—creativity or flow. The bottom line is the cyclical relationship between creativity and flow can go on for a long time. Therefore, if students play well-designed video games, they can experience the flow state more often, and if the nature of the game supports creativity, the chances are that students can enhance their creativity by playing those well-designed video games. According to the literature, some video games (e.g., *Minecraft* or *Portal 2*) can improve creativity (e.g., Blanco-Herrera, Gentile, & Rokkum, 2019; Inchamnan, Wyeth, & Johnson, 2013; Rahimi & Shute, 2021). However, none of these games have any clear creativity support in place. This study aims to address this deficit.

The purpose of our study was to investigate the effectiveness of two different kinds of supports on creativity—i.e., inspirational and instructional. Specifically, we aimed to see if inspiring college students using inspirational supports is more effective in improving their creativity compared to guiding them through the creativity process using instructional supports. We also investigated if combining inspirational and instructional creativity supports would be more effective in improving students' creativity compared to just providing one type of support. To our knowledge, no other studies have compared the effectiveness of these two creativity support types. To address the purpose of this study, we proposed the following research questions:

1) Are the creativity supports in PP's level editor effective in improving college students' creativity?
2) If they are effective, which creativity support is the most effective one?

The underlying hypotheses for the research questions include the following: (1) Compared to no creativity supports, creativity supports will be effective in general; and (2) Providing both inspirational and instructional supports will be the most effective creativity-support system followed by inspirational supports followed by instructional supports which will be more effective than no supports.

To our knowledge, no other studies have investigated the effectiveness of any explicit creativity supports in digital games. The studies we discuss in the next section have mainly investigated the effects of video games—without creativity supports present—on creativity in general. Additionally, because our particular creativity supports (i.e., inspirational and instructional—discussed in more detail later) have been used in other creativity training programs, the findings of this study can provide additional evidence regarding the effectiveness of these types of supports in general. Finally, the effectiveness of the creativity-support system developed and examined in this study can encourage educators to use such games with creativity-support tools to both help students learn (e.g., physics in this case) and enhance their creativity at the same time.

2. Background

In this section, we define and discuss the definition of creativity based on relevant literature from well-known researchers of creativity. Then, we discuss ways that creativity can be supported. Next, we briefly review current research on creativity in video games and we address the gap in the research on creativity and video games that this study addresses. Finally, we introduce the game we used in our study—*Physics Playground*.

2.1. The definition and nature of creativity

One of the most frequently used and agreed-upon definitions of creativity is that creativity refers to any product (e.g., idea, solution, artwork, story) that is both *novel* and *appropriate* (Amabile, 1983; Kaufman & Beghetto, 2009; Kaufman & Sternberg, 2007; Sternberg, 2006). A novel product is new or original—something that nearly no one else has thought of or done before. An appropriate product is logical, functional, practical, and valuable. Based on this definition, a creative product combines both novelty and appropriateness.

Jackson and Messick (1965) identified three criteria for a product to be called creative: *unusualness*, *appropriateness*, and *transformation*. According to Jackson and Messick, these three criteria should generate three corresponding responses of *surprise*, *satisfaction*, and *stimulation*. The surprise response occurs when a product "catches our eyes" and is unusual. The surprise reaction can occur multiple times when encountering a creative product as more unusual aspects of the product are discovered by the viewer. The satisfaction reaction happens when the product is appropriate in the context at hand. The amount of satisfaction depends on how well the product meets the expectations or demands for appropriateness. Furthermore, stimulation is a reaction to the level of transformation in the product (i.e., new forms, new ways of thinking, and transforming from conventionality). Stimulation accumulates as the viewer understands the level of products' transformation through the product's continuous freshness. Such a product makes the viewer say, "Wow! That is creative!"

Other aspects of a creative product include aesthetics and high quality (Sternberg, 2006). Cropley and Cropley (2011) also indicated that creativity includes aesthetic properties but aims to go beyond those properties. Creativity exceeds ordinary beauty through novelty, unusualness, and appropriateness (Cropley & Cropley, 2011). Similarly, Runco (2003) indicates that originality alone is not enough because truly creative ideas (or products) should have both originality and aesthetic appeal. Moreover, Cropley and Cropley (2011) stated that Immanuel Kant pointed out that beauty is in truth and order regardless of the domain. In this sense, a well written mathematical formula, a well-engineered machine, and a well-written poem are beautiful as they show truth and order.

Other than focusing on the final product creativity researchers have also focused on creative processes and creative thinking that lead to those products. As twenty scholars of creativity asserted in their recent manifesto that "seeing creativity as a form of doing or making [a creative product] does not deny the role played by creative thinking" (Glăveanu et al., 2020, p. 742). In the same vein, Walia (2019) defines creativity as "… an act arising out of a perception of the environment that acknowledges a certain disequilibrium, resulting in productive activity that challenges patterned thought processes and norms, and gives rise to something new in the form of a physical object or even a mental or an emotional construct" (p. 242). This definition shows how creativity is related to the environment in which one lives in and how one can perceive an opportunity for being creative through careful observations. In this vein, Glăveanu (2013) asserts that not only does the environment provide abundant affordances for creative people show their creativity, but creative people can also generate new affordances, useful for themselves and others in society.

Various researchers have operationalized creativity into its facets or processes to be able to assess and improve it accurately. For example, Guilford (1956) operationalized creativity as *divergent thinking* with four sub-facets: *fluency* (the ability to produce a large number of relevant ideas), *flexibility* (e.g., the ability to come up with relevant ideas from different categories or themes), *originality* (the ability to produce ideas that are statistically rare), and *elaboration* (the ability to implement an idea in detail and high quality). Moreover, research on creativity indicates that creative thinking made of both divergent and convergent thinking (Cropley, 2006). Convergent thinking refers to the ability to focus on one or few ideas to embellish and complete them and to come up with the single most correct solution (Cropley, 2006). Such multi-faceted models and theories can help researchers develop accurate creativity assessments and target creativity sub-facets to enhance creativity (e.g., Shute & Wang, 2016; Torrance, 1974).

One of the most famous tests of creativity is the Torrance Test of Creative Thinking (TTCT; Torrance, 1974). TTCT encompasses both verbal and figural tests of creative thinking. The verbal test consists of five activities: ask-and-guess, product improvement, unusual uses, unusual questions, and just suppose. Participants respond to a picture in writing for the verbal test. The figural test consists of three activities: picture construction, picture completion, and repeated figures of lines or circles. For example, in picture completion, participants are required to complete 10 incomplete figures and create a picture or a meaningful object. The TTCT comes in two parallel forms and is usually scored for fluency, flexibility, originality, and elaboration of participants' responses (see Kim, 2006 for a review of TTCT).

Other researchers have divided creativity into various levels. For example, Kaufman and Beghetto (2009) introduced a 4-C model of creativity including *mini-c* (i.e., personally meaningful creativity experiences when one is learning a new concept or skill and has a new insight about that learning experience), *little-c* (i.e., everyday creativity when one deals with common life problems via creative solutions that no one or just a few people think of), *Pro-c* (i.e., great breakthroughs that can change a field), and *Big-C* (i.e., the work of giants whose work still is being recognized many years after their passing). Our research aims to assess and enhance little-c creativity. Kaufman and Beghetto (2009) indicated that a consistent enhancement of creativity on the lower levels (i.e., mini-c and little-c) may eventually lead to creativity on the higher levels (i.e., Pro-c and possibly Big-c). What most of researchers of creativity focus on to assess and improve is the little-c or everyday creativity (Richards, 2010). Amabile (2017) also asserts that the literature of creativity still lacks sufficient answers to this question: "When ordinary people undertake creative endeavors in their work or their non-work lives, what is the nature of their everyday psychological experiences, and how do those experiences affect creative outcomes?" (p. 336)—by focusing on little-c creativity, this study aims to address this gap. It is worth mentioning that in the current study, the participants were asked to create as many creative game levels as possible, which can be seen as an example of little-c creativity. Little-c creativity, however, encompasses many other aspects of everyday life that is outside the scope of our study.

The creativity we refer to in this study comes from our comprehensive review of creativity definitions and categorizations. First, we are referring to everyday (Richards, 2010) or little-c creativity (Kaufman & Beghetto, 2009). Second, we use the divergent thinking definition of creativity (Guilford, 1956) for our external measure of creativity (i.e., fluency, flexibility, originality, and elaboration). Third, we include some aspects of a creative product that were mentioned in the literature to score the game levels that students created (i.e., humor and surprise and aesthetics) (Cropley, 2006; Jackson & Messick, 1965). We also include relevance (i.e., appropriateness), originality (Amabile, 1983; Kaufman & Beghetto, 2009; Kaufman & Sternberg, 2007; Sternberg, 2006), and elaboration (Guilford, 1956) to score the game levels that students created. We will discuss our measures in more detail in the Method section.

In summary, creativity is a complex, hard-to-measure construct. However, according to the literature, we know that creativity can be measured and possibly enhanced. Moreover, we know the essential factors that can influence creativity—both positively and negatively. With advancements in technologies and the learning sciences, new environments can be used to assess, support, and enhance creativity—for example, video games. Before we discuss the effects of video games on creativity, we review the relevant literature about supporting creativity, writ large.

2.2. *Supporting creativity*

Shneiderman (2007) suggested that the extensive literature on creativity and innovation can be seen in three schools of thought: *Inspirationism*, *Structurism*, and *Situationism*. Inspirationalists believe that creative ideas or solutions can emerge from temporarily getting away from conventional and familiar structures to get inspired by unusual thoughts, connections, associations, and even unrelated problems. To be more creative, inspirationalists encourage people to meditate, walk in scenic locations, and engage with some other unrelated problems. Inspirationalists also suggest methods such as brainstorming, reviewing existing work in the domain at hand, sketching to explore all possibilities, mapping out concepts to realize new relationships, and visualizing activities to aid in understanding the big picture.

Structuralists believe that people can be creative if they follow specific steps in an orderly manner. For example, Amabile's componential model of creativity (Amabile, 1983; Amabile & Pratt, 2016) includes five stages: *Stage 1*: task presentation or problem formulation, *Stage 2*: preparation, *Stage 3*: idea generation, *Stage 4*: idea validation, and *Stage 5*: outcome assessment. Amabile argues that Stages 1 and 3 can help to enhance the number and novelty of ideas, and Stages 2 and 4 can help with the usefulness or appropriateness of ideas.

Finally, situationalists see creativity as a social activity. Life-long development of individuals' creativity—from when we start playing with others as a child until we rely on others in producing creative ideas and products—is not possible without social relations among people (Glăveanu et al., 2020). That is, this school of thought focuses on people's backgrounds (e.g., family, social class, mentors, and friends), as well as their motivation (i.e., the internal or external drive that makes people creative). Shneiderman (2007, 2009) further suggests that one can develop creativity-support tools that include ideas from one, two, or all the three schools of thought. In this study, we focus on inspirationalist and structuralist views—the situationalists' views are not in the scope of this study. Next, we discuss the common creativity-support strategies and techniques for improving creativity relevant to this study.

2.3. *Creativity-support strategies and techniques*

Over the past seventy years, creativity researchers have come up with strategies and techniques that have shown positive effects on enhancing people's creativity (Csikszentmihalyi, 1997; Sternberg, 1988; VanGundy, 1982). In this section, we briefly review some of the common strategies and techniques for improving creativity that are relevant to our study.

2.3.1. *Analogy making*

Analogical reasoning is a cognitive process that can support the acquisition of new knowledge and facilitate learning—especially learning the unknown via the known (Hofstadter, 2001). An example includes connecting real-life experiences, like playing on a seesaw, to concepts in physics, like torque and equilibrium. Koestler (1975) explains creativity as "the sudden interlocking of two previously unrelated skills, or matrices of thought" (p. 121). This interlocking can happen when one is trying to make an analogy connecting two seemingly unrelated things (i.e., remote association). Making analogies can help generate unique, unexpected ideas, and they have been the source of inspiration in many architectural designs (e.g., designing the Museum of Tokyo like a snail shell) (Casakin, 2004; Casakin & van Timmeren, 2014). Analogical reasoning can be used in the early stages of the creativity process—i.e., the idea generation phase (Bonnardel, 1999). One can start with comparing similar objects (e.g., a bird and an airplane) and then compare objects/things that are not similar at all (e.g., a tree and a car). A simple practice for enhancing people's creativity using analogy making is to provide two randomly selected words and ask the trainee to come up with a creative idea using those two words. This practice can continue for several iterations using different words. Alternatively, a creativity support system can include a list of remotely associated words shown to the participants to help them come up with creative ideas from combining those words (see section 3.3.3 in this paper for one example).

2.3.2. Divergent thinking strategies

These strategies enhance students' ability to think of as many ideas/solutions as possible in a given situation (Runco & Acar, 2012). Divergent Thinking (DT) includes fluency, flexibility, originality, and elaboration. Fluency can be seen as the foundation of DT activities. For example, instructing people to think of as many ideas as they can—rather than spending a lot of time working on the first idea that comes to mind—in the beginning of a creativity process enhances people's fluency. Solving a problem in different ways enhances flexibility (e.g., requiring student to solve a math problem in two different ways), and thinking of things that no one else has thought of improves originality (e.g., providing the common solutions to a problem and asking students to come up with solutions other than those provided). Finally, asking someone how they can improve upon an idea enhances elaboration (e.g., instructing students to elaborate their ideas by providing more details). Many of the creativity training programs over the past decades were designed based on improving people's DT skills—which showed positive results (Baer, 1994, 1996; Runco & Acar, 2012)—see sections 3.4 and 3.3.2 to see how we used instructions to help enhance students' DT skills.

2.3.3. Brainstorming

The most common strategy for generating ideas is brainstorming (Osborn, 1953). Osborn's four rules for brainstorming include: (1) generating as many ideas as possible, (2) building new ideas upon previous ones, (3) not judging the generated ideas, and (4) feeling free to contribute seemingly impractical or wild ideas. This method has shown to be effective for generating a large number of original ideas both in group- and individual-brainstorming sessions (Litchfield, 2008). However, there is some evidence that individual brainstorming leads to a greater number of ideas with higher quality compared to group brainstorming (e.g., Putman & Paulus, 2009). In a literature review of 93 studies of creativity-support tools, Frich et al. (2019) indicated that 45% of the tools used a type of brainstorming support (28 out of 93 tools were aimed for designers including game designers). Creativity support tools can include brainstorming tools and platforms that demand creativity can promote brainstorming to enhance people's creativity. For example, if one wants to create a logo for a company using *Adobe Photoshop*, the software could offer a space for the artist to brainstorm some ideas first or the artist might choose to do brainstorming on a piece of paper first before getting started with the design process—see section 3.3.2 which shows details about the brainstorming tool we included in our study.

2.3.4. SCAMPER

SCAMPER is the name of a method for using idea-spurring questions to help students generate diverse ideas, solutions, or products (Eberle, 1972). SCAMPER stands for ***S**ubstitute* (e.g., replacing a material in a product), ***C**ombine* (e.g., getting some pieces from other ideas and creating a new idea), ***A**dapt* (i.e., changing something known), ***M**odification* (i.e., improving previous ideas using tools), ***P**ut* (i.e., using objects for other uses than they were planned to be used for), ***E**liminate* (i.e., removing things to solve a problem creatively), and ***R**earranging* or ***R**eversing* (e.g., left-handed scissors). Any or all parts of SCAMPER can be used when students need to generate new ideas or when they need to enhance their previous ideas. Students can use SCAMPER questions to bring new ideas to mind—rather than waiting for the ideas to form. SCAMPER can raise curiosity, stimulate engagement, and provide strategies for developing people's imagination (Eberle, 1972; Islim & Karataş, 2016). See Appendix A for the SCAMPER question we used in this study and see section 3.4. for detailed explanation about how we presented the SCAMPER questions to our participants. Alternatively, SCAMPER questions be printed out and provided to people as an aid when they want to create a creative product. There are other techniques and strategies for improving creativity (e.g., concept), which outside the scope of this study.

Research from the human-computer interaction field (HCI) shows that creativity-support tools (systems) facilitating the techniques and tools described above can effectively enhance creativity (e.g., Massetti, 1996; Remy et al., 2020; Resnick et al., 2005; Ristow, 1988; Sengewald & Roth, 2020; Shneiderman, 2009). However, more research is needed to (a) understand how creativity-support tools affect the creativity processes and products, and (b) find ways to enhance the effectiveness of creativity-support tools in various settings and contexts. Toward that end, HCI research on creativity-support tools should be conjoined with creativity research in other fields such as learning sciences and psychology (Frich et al., 2018). The current study is a response to this call. Next, we discuss the impact of video games on creativity.

2.4. Creativity and video games

The findings of correlational studies investigating the effects of video games on people's creativity (usually measured by TTCT) are conflicted. For example, Hamlen (2009) found no significant relationships between the time students spent playing video games and students' creativity—holding students' gender and grade constant. In contrast, Jackson et al. (2012) found a positive and significant correlation between students' video games playing and their creativity. That is, students who played more video games tended to have higher creativity scores. However, one should not mistake correlation for causation. The results also showed that creativity scores were not significantly correlated with the use of other technologies.

These studies do not provide a reliable answer about whether or not playing video games can improve creativity. Moreover, the results of these types of studies are questionable for two reasons. First, putting all types of video games into one category when studying the effects of video games on creativity is too general. Second, the self-report nature of some of the data collected in these studies (e.g., gaming background) can lead to misleading results. Several other qualitative studies using observations and participants' perception data (e.g., Cipollone et al., 2014; Inchamnan et al., 2014; Karsenti & Bugmann, 2017) have shown some positive evidence about the potential of particular video game genres (e.g., puzzle and sandbox games, like *Minecraft* and *Portal 2)*.

In an experimental study, Moffat et al. (2017) conducted an experimental study comparing the effects of three games (*Minecraft*, *Portal 2*, and *Serious Sam*) on 21 undergraduate students' creativity (measured by TTCT, in both the pretest and posttest). Results showed that playing these three video games did not have an effect on participants' originality and fluency scores, but two of the games showed a significant effect on pretest-to-posttest improvement in flexibility—*Serious Sam*, and *Portal 2*. Focusing on *Portal 2*, Gallagher and Grimm (2018) investigated if game-making (i.e., designing *Portal 2* game levels) can improve people's creativity (i.e., divergent thinking) and spatial skills. Comparing the control group, participants who created game-levels in *Portal 2* showed more creativity and spatial skills on the posttest controlling for the pretest.

Similarly, Checa-Romero and Gómez (2018) focused on *Minecraft* and investigated its effectiveness on students' creativity in an eight-week workshop in a technology course. Students had to create a machinima (i.e., a video made of screen recording from *Minecraft*) of their "dream house" in *Minecraft*. Students completed a pretest and a posttest of creativity—the CREA test of creativity (Corbalán et al., 2003), which is designed to measure creativity through searching and solving problems (i.e., students received three illustrations had to come up with as many questions as possible related to those illustrations). Results showed a significant improvement in students' creativity from pretest to posttest ($t = -6.11$, $p < .05$, Cohen's $d = 0.45$).

Additionally, Blanco-Herrera et al. (2019), in a more recent and larger study ($n = 352$ undergraduates from a Midwestern university), compared the impact of playing Minecraft—with and without instructions (i.e., students who played *Minecraft* and received instructions were told to be creative), playing a racing game called NASCAR, and watching a TV show on students' creativity. Participants were randomly assigned to one of the four groups and played their assigned game or watched an engaging TV show (in the control group) for 45 min. They then completed three posttest creativity scales (i.e., creative production alien drawing, divergent thinking scale of alternative uses, and convergent thinking remote association). Results showed no significant effect of condition on the alternative uses scale or remote association scale. However, the convergent thinking alien drawing scale showed a significant condition effect (F (3, 29) $= 7.74$, $p < .01$) controlling for the participants' GPA. Participants in the *Minecraft* group with no instructions scored significantly higher on the convergent thinking task than those in the other groups. As researchers expected, students in both Minecraft groups scored higher than the other two groups (i.e., NASCAR and the TV-watching group) on all of the creativity measures. However, the researchers hypothesized that the Minecraft group with instructions would do better on the posttest than the one with no instructions. This unexpected finding, again, shows that the conditions under which people play these complex games are critical for the positive effects of such games on participants' creativity.

Based on the findings from the studies discussed above, we can conclude that playing some video games (e.g., *Minecraft*, *Portal 2*) can have a positive impact on students' creativity. Moreover, game genres such as sandbox and puzzle games seem to have a higher potential for improving students' creativity than other game genres (e.g., shooting games) as they make it possible for players to create their own levels or worlds (e.g., Blanco-Herrera et al., 2019; Cipollone et al., 2014; Inchamnan et al., 2014; Karsenti & Bugmann, 2017; Moffat et al., 2017). Additionally, by including the players in the development of the game environment, the game can live for as long as the players want to create, maintain, and play new levels (Gee, 2005). To read a more detailed review of the effects of video games on creativity see Rahimi and Shute (2021).

Now the question is, "how can we enhance students' creativity via video games?" Games with high potential for developing creativity (e.g., Minecraft and Portal 2) do not have any clear support systems in them for engendering creativity. Moreover, little is even known about how creativity supports influence people's creativity in video games. As discussed earlier, there are two areas related to improving creativity via video games: (1) creativity in gameplay, and (2) creativity in creating game content (e.g., game levels). In this study, we are focusing on the latter—designing game levels rather than playing them. Therefore, the aim of the current study is to design, develop, and test a creativity-support system with two main support types (i.e., inspirational and instructional), embedded in a puzzle game called *Physics Playground*.

2.5. *Physics Playground*

Physics Playground (*PP*;; Shute et al., 2019) is a 2-dimensional game designed to teach conceptual physics to 8th and 9th graders. The goal across all levels in this game is simple—hit a red balloon with a green ball by drawing objects and creating simple physics machines (i.e., ramp, springboard, pendulum, and lever) on the screen. Fig. 1 shows an example of a *PP* game-level called "*Little Mermaid,*" where the player has drawn a spoon-shaped springboard (in red) and a weight (in blue), which will drop on the springboard when released and propel the ball to hit the balloon.

PP includes a level editor whereby non-technical users (e.g., students and teachers) can create their own levels by drawing objects

Fig. 1. Example of a game-level in PP called Little Mermaid.

(e.g., lines, shapes, or other objects) on the screen (Fig. 2 shows the current version of *PP*'s level editor—discussed in the Method section). Starting with an empty stage, students can place the ball and balloon anywhere on the screen, and draw any number of obstacles between them. There are endless opportunities for students to show their creativity in this environment. Our goal, in this study, was to help students successfully design creative levels. Hypothetically, each time students successfully designed a creative level using the creativity-support system, their creativity gets stronger than before through practice, and can potentially be transferred to other related situations (e.g., the posttest creativity tasks).

3. Method

3.1. Research design and participants

This study's research design is an experimental, pretest-posttest design (Shadish et al., 2002). We recruited 124 undergraduate and graduate students from twelve colleges at a university in Florida. Ten students dropped out of the study and our final sample included 114 students ($M_{Age} = 26.25$, $SD = 8.06$; Min = 17, Max = 51; Females = 54%, Males = 46%; undergraduate students = 47%; graduate students = 53%, from various ethnicities with the majority of them as White (47%), Asian (16%), and Hispanic (15%). Frich et al. (2019) indicated in their literature review of creativity support tools that most of the tools they reviewed were created for experts and not for novices. We decided to focus on novices. That is, our target audience did not have any experience with the game we used and none of them were game designers. Based on the demographic data, the majority of the participants were from colleges of Arts and Science (30%), Education (28%), Business (11%), Social Sciences (11%), and Communication and Information (10%). Upon completion of all steps of the study, students received a $10 gift card. To further encourage students to participate in this study, we gave four extra $25 gift cards to the most creative students per condition.

We randomly assigned participants into one of the following four conditions: (1) the Inspirational condition ($n = 29$) where students received inspirational supports (i.e., brainstorming, accessing *PP*'s existing game levels for inspiration, and analogy-making or remote-association support); (2) the Instructional condition ($n = 28$) where students received structural supports (i.e., students were instructed to follow three steps to create their levels—discussed later in detail); (3) the Both condition ($n = 28$) who received inspirational and structural supports; and (4) the No Support condition ($n = 29$) who received no creativity supports and only used *PP*'s level editor to create game levels. Each condition used a different version of *PP*'s level editor to create game levels.

Students in all four conditions received a brief version of the scoring rubric (see Appendix C) that was used to assess the creativity of their game levels. They all were also instructed to "be creative and create as many creative levels as you can," to "think of ideas that you think no one else can think of," and "make sure that the levels you create can be solved." Additionally, a researcher was available throughout the process to rectify any technical issues or answer students' questions about how to use the level editor.

3.2. What all the conditions Share—PP's level editor

The level editor used in this study is shown in Fig. 2. We created four different versions of the level editor for each condition. *PP*'s level editor includes various tools which allow students to create their own game levels (e.g., drawing tools, loading an existing level, saving and playing a level, and changing the behavior of objects in the level—such as making them static or dynamic). Next, we discuss the details of *PP*'s level editor per condition starting with the Inspiration condition.

Fig. 2. PP's level editor.

3.3. Inspirational condition

3.3.1. Existing PP levels' website

Clicking the "Example Levels" button opened a new tab showing a website of various *PP* levels (Fig. 3). Using this website, students could access the images and the solution videos of 30 existing *PP* levels. These levels were categorized as *low*, *medium*, and *high* relative to creativity (ten levels per category). A button called *Rubric* was also available for the students to let them know how their levels will be judged. During the study, students in the Inspirational condition were instructed to spend 10 min exploring the levels and record their thoughts in the brainstorming spreadsheet.

Providing examples can sometimes lead to fixation (i.e., replication of examples) (Sio et al., 2015). To prevent this issue, when explaining the *originality* aspect of the scoring rubric, we mentioned that "the levels you will create should be different that the levels you see on this website; if a game level you create is identical to any of these levels on the website, you will receive 0 for originality; if a game level is somewhat similar to a game level on the website, you will receive 1 for originality, and if it different than these levels, you will receive 2 for originality." Sio et al. (2015) indicated that looking at examples may lead to copying and combining some aspects of the examples which can lead to greater novelty of products.

3.3.2. Brainstorming

The brainstorming support included two main parts: (1) *Instructions* that reflect Osborn's four rules (i.e., produce as many ideas as possible, withhold judgment of the ideas, build on previous ideas, and freely suggest wild ideas); and (2) the *Brainstorming spreadsheet* – a simple Google spreadsheet for each participant (see Fig. 4) where students can record their ideas, elaborate them, rate them (from 1—not creative at all, to 5—very creative), and access them online whenever they need them. Students were instructed to spend 5 min writing as many ideas as they could think of.

3.3.3. Remote associations

Before designing a level, a popup menu would appear on the screen showing six words in two columns (see Fig. 5). The words in one column were randomly selected from a pool of intradomain words (related to physics—e.g., seesaw), and the words in the other column were randomly selected from a pool of interdomain words (e.g., "spider"). Students could click "OK" whenever they felt they were ready to start designing their level. Note that students were only asked to read the words and try to make a connection between

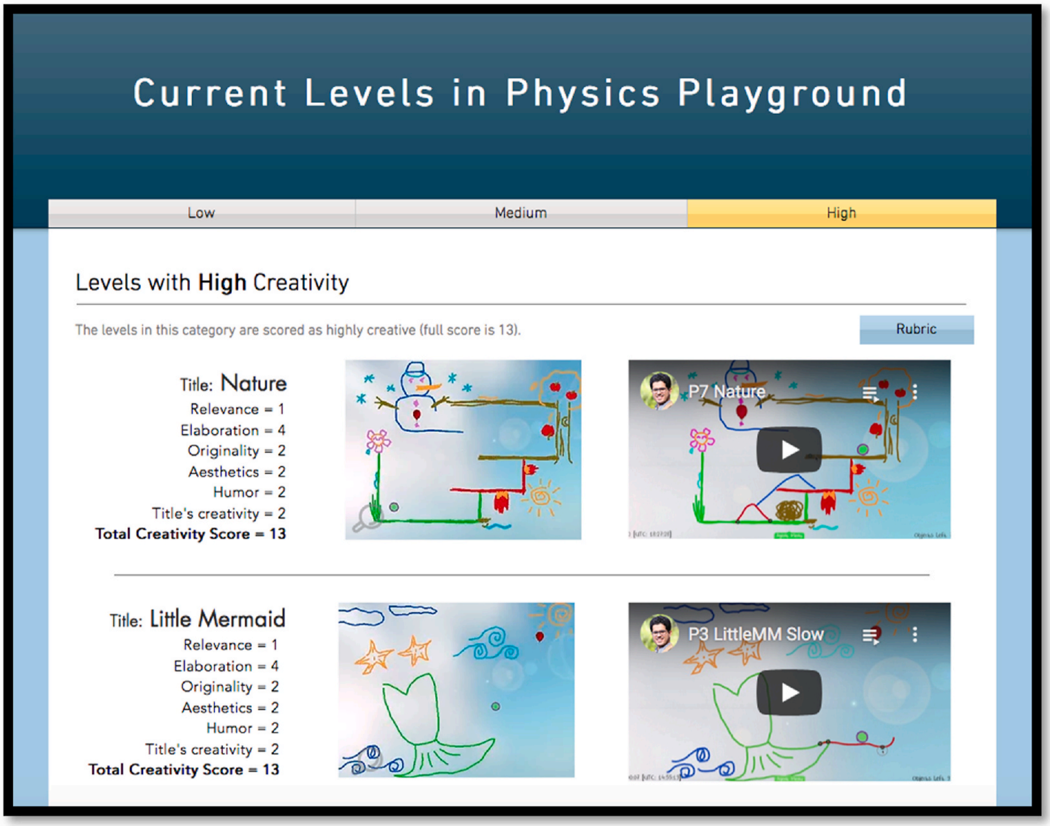

Fig. 3. The existing *PP* levels' website.

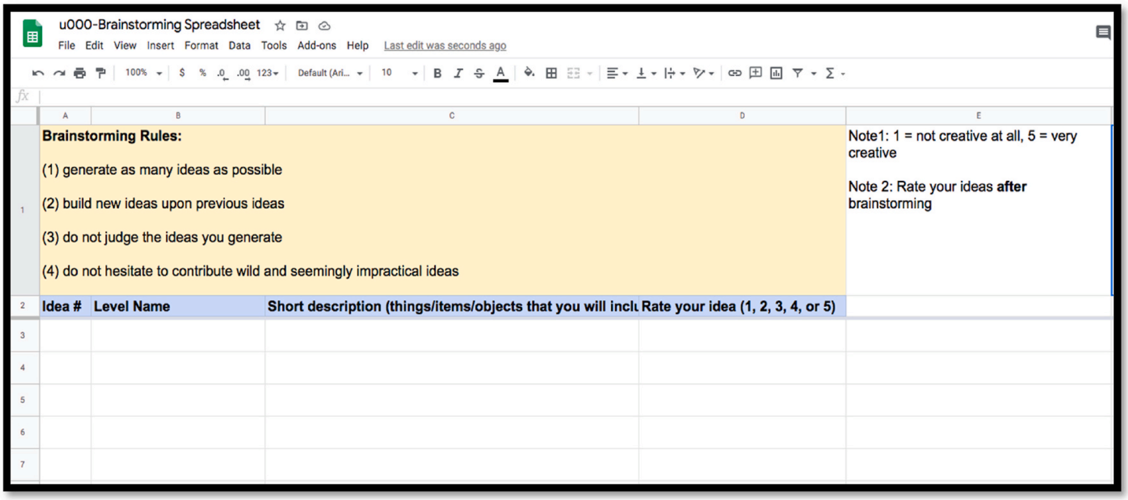

Fig. 4. Brainstorming Google spreadsheet with brainstorming rules.

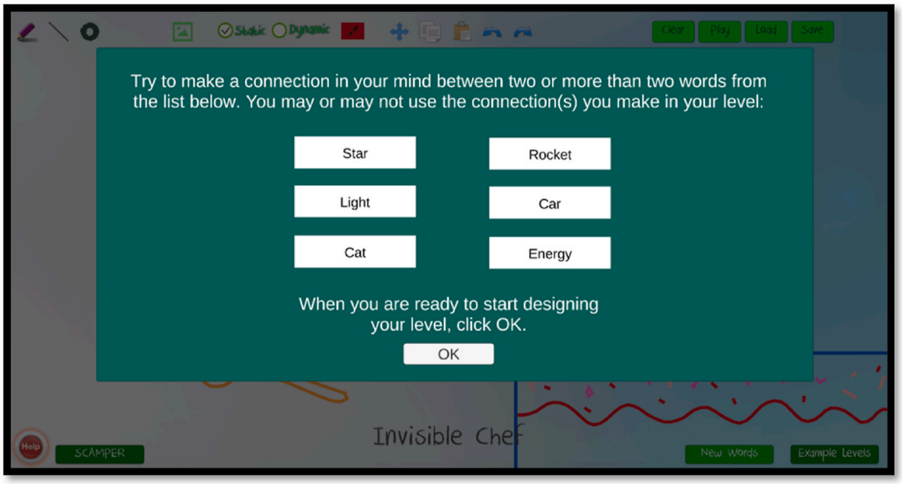

Fig. 5. Remote-association popup menu.

two or more words in their minds—they were not asked to write a sentence using the two words.

Creativity-support tools should facilitate thinking by free (or remote) associations (Shneiderman, 1999). When students try to make a connection between two remotely related words that they see on the screen, they could possibly use the same association in their level design experience to increase the creativity of their game levels. For example, when one connects “seesaw” and “spider” in a meaningful way in his or her mind, they may get inspired by this association and use it as a point of departure for designing a creative level. Having words instead of pictures (as inspirational supports) and having both interdomain and intradomain words can be beneficial to students’ creativity (Bonnardel, 1999).

Students were instructed that they could refer to these supports as many times as they wanted during the game design process.

3.4. Instructional condition

Students in the Instructional condition received supports that were based on the Structurism school of thought which aimed to get students to follow a predefined, orderly process. Students were instructed to design their game levels in three steps. *Step 1*: design as many levels as possible without paying attention to details; *Step 2*: choose four levels you think you can enhance; and *Step 3*: use SCAMPER to enhance those four levels’ creativity. Students could access SCAMPER via a button in the level editor. By clicking the SCAMPER button, a popup menu would show up (Fig. 6), which included seven buttons representing the words that SCAMPER stands for.

When students clicked any of the buttons on the popup menu, they could see the questions related to that word they selected. To see

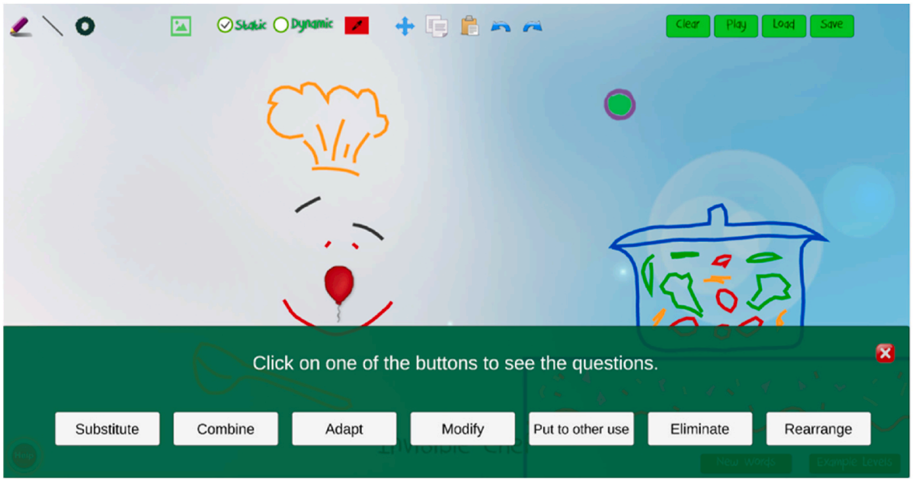

Fig. 6. SCAMPER popup menu in *PP*'s level editor.

a more complete list of SCAMPER messages and questions, see Appendix A.

Students were instructed that they could choose one letter from SCAMPER at a time in any order they wanted (e.g., M for Modify), read the questions under that word, and try to think and act on that question. Eberle (1972) suggested that SCAMPER questions can be used in any order. That is, one might need to get some ideas only by modifying related questions and not substituting related ones at a given moment. Note that using SCAMPER or enhancing game levels is, in general, one step of the instructions provided to students—the other steps were (1) create as many game levels as possible, (2) choose the four most creative game levels. Moreover, student could choose to apply all or some of the SCAMPER questions, or they could freely, without referring to the SCAMPER questions, enhance their game levels.

3.5. Both condition

Students in the Both condition received inspirational and instructional creativity supports. The same supports that were available for the Inspirational condition (i.e., brainstorming spreadsheet, previous levels website, and remote association exercise) and the Instructional condition (i.e., guidelines for designing levels in three steps, and SCAMPER questions) were available. Students went through the following steps that were timed and facilitated by a researcher: (1) reviewing the existing levels' website for 10 min, (2) brainstorming for 5 min, (3) seeing and making a remote association before creating a level, (4) creating as many levels as they could by referring to their brainstorming spreadsheet, (5) selecting four levels to enhance, and (6) using SCAMPER to elaborate those four levels for the remaining time.

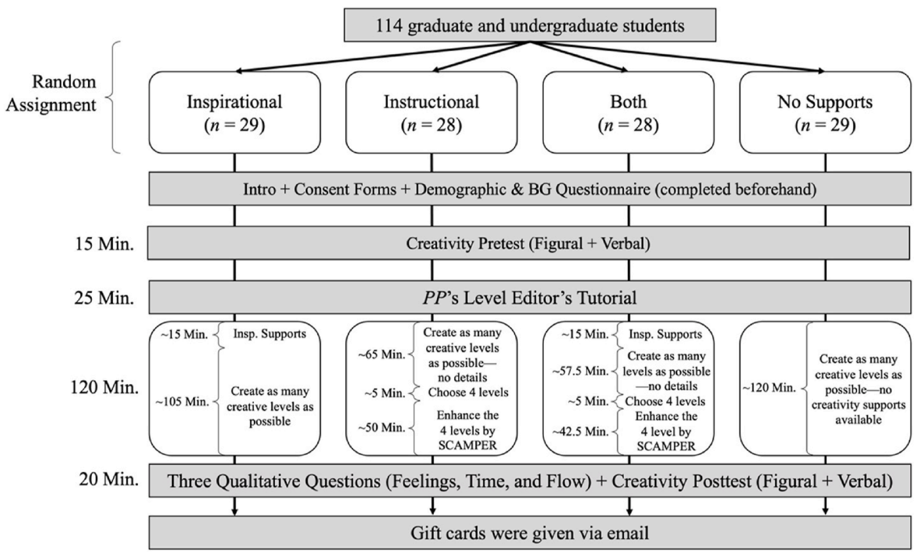

Fig. 7. Procedure per condition. *Note.* Insp. = Inspirational, BG = Background.

We did not counterbalance the delivery order of supports in the Both condition because the inspirational supports, in our view, logically should precede particular instructions (i.e., first get inspired and then plan and create something following instructions).

3.6. No support condition

Students in the No Support condition received instructions on how to use the level editor and then used the level editor with no creativity supports.

3.7. Procedure

Due to the COVID-19 pandemic, this study was conducted fully online via the Zoom platform in one session that lasted about 3 h (see Fig. 7). Collecting data from all 114 students took about three months (from May 2020 until August 2020 across 67 sessions). Students received the materials they needed for the study before the study began (i.e., the consent form, the demographic questionnaire, a Zoom link, and a user ID), electronically signed the consent forms, and completed a demographic questionnaire. On the study day, each student first completed a creativity pretest, then received a game and level-editor tutorial, as well as information about how their levels would be judged for creativity. After the instructions, students were told to load and play one easy level to experience playing *PP* before creating a level.

When the level-editor tutorial was completed, students started designing their own game levels (for 120 min). Students assigned to conditions with creativity supports received condition-specific instructions regarding how to use the creativity supports and what steps to follow. Data collection was done one student in one condition at a time. At the end of the game-level design period, students in all conditions were asked to select their four most creative levels. Then, students responded to three open-ended questions about their experience via a private chat message in Zoom and their responses were recorded. Afterward, students completed a posttest of creativity. Finally, students sent their game levels and their responses to the creativity tests (they were required to draw on a paper and take a picture using their phone) to a researcher in the session via email.

3.8. Measures

3.8.1. Demographic and background questionnaire

The first part of this questionnaire included demographic questions such as participants' gender, age, and current academic level (i.e., Bachelor, Master, or Doctoral). The second part of this questionnaire included gaming background questions (e.g., a 1 to 5 scale question "how often do you play video games?" with 1 being never and 5 being almost every day) and a 1 to 5 scale question about participants' perceived creativity (i.e., "how creative do you think you are?") with 1 being not creative at all and 5 being very creative.

3.8.2. Figural test

To measure participants' level of creativity before and after the study, we designed, developed, and pilot-tested an assessment of creative thinking inspired by the figural Torrance Test of Creative Thinking (Torrance, 1974) with two isomorphic forms (Form A and Form B). We counterbalanced forms so that those who received Form A on the pretest received Form B on the posttest, and vice versa. Each form contained two items. *Item 1*: students were asked to create interesting objects using basic shapes like circles and diamonds. Specifically, students were given 30 circles and they were asked to complete as many circles as they could (e.g., make a smiley face out of a circle) and then title their drawings (i.e., we used circles for item 1 in Form A and diamonds for item 1 in Form B). *Item 2*: students were asked to create as many meaningful objects as they could using three ordinary shapes only (e.g., a rectangle, a circle, and a diamond) and then state what they created. Students had 4 min to complete each task—8 minutes in total. Both pretest and posttest were administered using pen and paper.

Each item per form was scored on *fluency*, *originality, flexibility*, and *elaboration*. *Elaboration* and *flexibility* were subjective measures and needed to be rated by two trained raters. We trained two raters to rate students' figural responses for these two measures independently. For elaboration the raters scored each drawing using a scale from 1 to 5 using two rubrics (one rubric per item) with 1 being not elaborated at all and 5 being highly elaborated—see Appendix B. The two raters went through all students' responses and gave a 1-to-5 score to each drawing. Then, we averaged those scores to have one elaboration score per item per student.

For flexibility, we provided the common categories (generated from students' responses when we scored for fluency and originality) to the raters. The raters counted the number of unique categories in the responses based on the common categories. For example, a student who drew three different balls (e.g., a beach ball, a basketball, and a tennis ball) would receive a score of 1 for flexibility for the circles item (all came from the same category). But a student who drew three responses that fit into three categories (e.g., a smiley face, a basketball, and a flower) would receive a score of 3 for flexibility for the circles item. In cases where the raters could not fit a drawing into one of the given categories, they were told to create their own category for that particular drawing. The interrater reliability of the ratings for each figural item ranged from ($r = 0.86, p < .001$) to ($r = 0.96, p < .001$). We averaged the ratings of the two raters to get one single score for elaboration and flexibility per item per form per student.

Fluency and *originality* were two variables that could be scored objectively. To score fluency, we recorded the number of unique drawings each student created, per item. To score originality, we examined the uniqueness of the objects drawn based on its frequency within the sample. That is, an object's frequency was computed as a ratio (e.g., 95 out of 114 students drew a smiley face—thus it was a common/frequent response). Then the originality per object was defined as a number between 0 (common) to 1 (unique) computed by the following formula:

$$Each\ Drawing's\ Originality = 1 - \frac{Frequency\ of\ the\ drawing}{Number\ Participants}$$

We first averaged the originality scores for each drawing per student. In this method, we included scores with low or medium originality into the originality scores. This method led to an originality score that highly correlated with fluency which could be problematic (i.e., we needed an originality score which was only moderately correlating with other facets). Therefore, we defined the originality score as the amount of *originality* produced by each student. In this definition, the originality of drawings such as a smiley face would not be included because such drawings had a low originality score. To compute the *raw originality score* per student we summed the originality scores of the drawings that were original (i.e., very few people thought of them). For example, as shown in Table 1, the student with u114 user ID drew three drawings two of which were original (i.e., Elephant and Traffic Light). Therefore, the raw originality score for that student is computed as the sum of 0.97 and 0.98 which is 1.95—excluding 0.54.

But to include and exclude the originality scores we needed a cutoff score. The originality cutoffs per item differed, and were selected after several iterations (i.e., we selected cutoffs per item that would yield originality scores that were moderately correlating with fluency). For example, we chose 0.91 as the cutoff score for item 1 on Form A (i.e., only summed the originality scores above 0.91). Next, we correlated students' raw originality scores with their fluency scores for item 1 on Form A, and found a correlation of 0.72—which was too high. Therefore, we had to increase the originality cutoff for that particular item and reexamine the correlation of originality with the fluency score. This method allowed us to differentiate originality from fluency. Note that in this way students' responses should be very original to be included in their originality score.

The cutoffs we selected for figural items per form ranged from 0.95 to 0.98, with correlations ranging from 0.54 to 0.58 between students' raw originality scores and the corresponding students' fluency scores for each item per form. One last step to finalize the originality scores was to weight them based on the elaboration scores for each student per item. We used this elaboration score to weight the raw originality scores for each item per form and per student using the following formula:

$$Weighted\ Originality\ of\ Item_i = Raw\ Originality\ of\ Item_i \times \frac{Elaboration\ of\ Item_i}{5}$$

The reason for including elaboration in computing originality was to get a measure for originality which can be more discriminating than the raw originality scores. For example, if a student drew only one highly original item (e.g., with originality of .98) with average elaboration score of 2, the weighted originality score of that student for that particular item = .39. In a similar situation, another student with only one original drawing (e.g., originality of 0.98) with an average elaboration score of 5 will retain his or her originality score as 0.98. The weighted originality scores showed lower and more reasonable correlations with fluency scores than the raw originality scores, and generated a small to medium correlation with elaboration scores per test item. Therefore, we believe that including the elaboration weight into computing students' originality scores was reasonable.

We calculated the reliability of the scores per form. The Cronbach's α was 0.73 for the figural Form A, and 0.78 for the figural Form B, which were reasonable. Finally, to have one single score per facet for each form (i.e., four scores per form), we summed the scores of each facet in each form. To create a single score for creativity for the pretest and the posttest per student, we transferred all four equated pretest and posttest scores on a *z* scale. Finally, we averaged fluency and flexibility (as they highly correlated; $r_{\text{fluency * flexibility on the pretest}} = .91$, $p < .001$; $r_{\text{fluency * flexibility on the posttest}} = .91$, $p < .001$) to avoid redundant information, and then summed the standardized originality, elaboration, and the average of fluency and flexibility to compute a single overall figural creativity score for the pretest and the posttest per student.

3.8.3. Level-design measures of creativity

3.8.3.1. Number of levels created—fluency. We counted all game levels that students created during their game design time by counting each solvable level created per student—a proxy for fluency.

3.8.3.2. Creativity assessment of the four most creative levels. As mentioned earlier we asked students to choose their four most creative levels at the end of the game design session. We modified the creativity scoring rubric created by Shute and Wang (2016; see Appendix C) to score those four most creative levels selected by the students. Four main modifications were applied to the rubric mentioned above: (1) changing the components for elaboration from a focus on difficulty of the game level to details of the level in terms of lines and meaningful objects, (2) including color in considering aesthetics, (3) including surprise in the humor facet, and (4) adding the creativity of the title as its own facet. This method of creativity assessment is called the consensual technique for creativity assessment, first introduced by Amabile (1996). In this technique, at least two experts on the topic on which the products are judged

Table 1
Raw originality scores per student for one test item.

UID	Smiley Face	Elephant	…	Traffic Light	*Raw Originality*
u001	.54	.97	…	0	.97
u002	.54	0	…	.98	.98
…	…	…	…	…	…
u114	.54	.97	…	.98	1.95

rate the products on various predefined features. If both raters agree that a product is creative, then we must accept that product as creative.

We trained two raters to score each student's four levels. Both raters had experience working with *Physics Playground* over the past several years and have used *PP*'s level editor to design multiple game levels. Each level received seven scores for relevance, line elaboration, meaningful objects elaboration, originality, aesthetics, humor or surprise, and level's title's creativity (see Appendix C for the full rubric). The sum of the seven scores created the overall creativity score per created level (maximum possible score = 13). Next, we averaged the four single scores per student from each rater and created one single score per student, then computed the interrater reliability ($r = 0.94$, $p < .001$) which was high. The Cronbach's α for the creativity scores was 0.92 suggesting that the measure was also reliable—see Rahimi (2020) for a full report on the validity and reliability of this rubric.

3.8.4. Three open-ended qualitative questions

To get a more in-depth sense about the usability and effectiveness of the creativity supports in this study, at the end of the 3 h, we asked three open-ended questions about participants' overall affect, the time sufficiency, and whether or not they lost sense of time (as a proxy for experiencing flow):

1. How did you feel when designing your game levels (frustrated, bored, happy, excited, or any other feelings?) Why?
2. How was the time for you (enough, not enough, mixed)?
3. Did you lose track of time while designing your game levels or you were frequently checking the time to see when the session would be over?

Student responses to these questions were coded by two raters independently using the three codes per question—see Appendix D. For example, a response such as, "the time was perfect" was coded as 3 for *time*. However, a response such as, "for the first part I had enough time, but for the second part to enhance my levels I needed more time" was coded as 2 for time. Generally, the mixed code was given to responses where students provided two or more than two opposite responses about their feelings, the time, and their flow experience. The percentages of agreement between the ratings of the two raters were 81% for *feelings*, 92% for *time*, and 81% for *flow*. The two raters met and discussed disagreements and reached 100% agreement on all the three variables.

3.9. Statistical analyses

To address the research questions, we conducted multiple analyses of covariance (ANCOVA) for the overall figural creativity scores and for each facet of the figural posttest scores (i.e., fluency, originality, flexibility, and elaboration) as the dependent variables. Condition (i.e., Inspirational, Instructional, Both, and No Supports) served as the independent variable, and pretest scores for each of the four creativity tests served as covariates. Additionally, we conducted an ANCOVA using the average creativity score (computed from the four most creative levels created by students) in each condition as the dependent variable, and the figural pretest score of creativity (i.e., sum of the z scores of the four facets) as the covariate. Our final ANCOVA used the number of levels students created as the dependent variable to test if there was any condition effect (with the figural and verbal fluency scores as covariates).

Huberty and Morris (1989) suggested that multiple ANCOVAs can be selected over a MANCOVA when there are a few number of outcome variables in the study, and when the research questions are univariate (i.e., targeting the effects of the independent variable on each of the dependent variables separately). The goal of our analyses was not to find any significant p value from these tests and claim that the treatments worked. If that was the case, we had to correct for the alpha level and divide our selected alpha level (.05) by the number of tests we conducted. Instead, our goal was to report the effect of the treatments on each outcome variable regardless of the effect being significant or not. In this case, correction for alpha is not relevant.

Before running each ANCOVA, we checked the relevant assumptions:

(1) All the dependent variables and the covariates were continuous.
(2) The independent variable consisted of one categorical variable with four independent categories (i.e., conditions).
(3) Students were randomly assigned to the four conditions; therefore, the observations were independent from each other.
(4) The significant outliers were removed from these figural analyses: overall creativity, originality, and flexibility.
(5) All of the Levene's tests of equality of error variances showed non-significant results indicating that the assumption of homogeneity of variance was met for each ANCOVA.
(6) The residuals of each category of the independent variable were tested for normality and all of them were normally distributed.
(7) The dependent variables in each ANCOVA significantly correlated and have a linear relationship with the corresponding covariate used.
(8) No significant interaction effect between the independent variable and the covariate was detected.

Finally, we conducted Pearson's Chi-Square tests of association using the students' responses on the three open-ended questions and the condition.

4. Results

4.1. Descriptive data

The random assignment of participants into the four conditions resulted in a balanced distribution in terms of the number of students, gender, and self-reported gaming and creativity backgrounds, per condition (see Table 2). In the demographic survey, students responded to three questions: Q1 "*How often do you play video games?*" scaled from 1 (never) to 5 (almost every day); Q2 "*How good do you think you are at playing video games?*" scaled from 1 (very poor) to 5 (very good); and Q3 "*How creative do you think you are?*" scaled from 1 (not creative at all) to 5 (very creative). Table 2 shows the means and standard deviations of participants' responses to these three questions broken down by condition and gender.

Results from an analysis of variance (ANOVA) showed that students' responses were not significantly different on Q1 [F (3, 110) = 0.58, $p = .63$, $\eta^2 = 0.02$], Q2 [F (3, 110) = 0.13, $p = .95$, $\eta^2 = 0.003$], and Q3 [F (3, 110) = 0.27, $p = .85$, $\eta^2 = 0.007$] which suggest that the conditions were comparable.

4.2. Creativity supports' effectiveness

4.2.1. Figural test

To address our research question concerning the effectiveness of the creativity supports, we conducted separate analyses of covariance (ANCOVA). Students in the four conditions scored statistically the same on the figural creativity posttest (overall score) controlling for the pretest (see Table 3). Results of the ANCOVAs using the facet-level variables showed that the difference among conditions was not significant for fluency, originality, and elaboration, but was significant for flexibility [F (3,108) = 2.88, $p = .04$, $\eta^2 = 0.07$]. The parameter estimates showed that the mean difference between the Both condition and the No Support condition was statistically significant on flexibility ($\beta = 1.68$, $SE = 0.75$, $t = 2.26$, $p = .03$, 95% CI [0.20, 3.16], $\eta^2 = 0.05$). That is, students in the Both condition scored significantly higher on the flexibility posttest (*Adjusted M* = 13.38, $SE = 0.53$) than students in the No Support condition (*Adjusted M* = 11.69, $SE = 0.52$) with Cohen's $d = 0.39$ controlling for the pretest. The Inspirational and Instructional conditions were not significantly different on the posttest flexibility compared to the No Support condition controlling for the pretest.

Results from other comparisons showed that students in the Both condition scored significantly higher on the posttest flexibility (*Adjusted M* = 13.38, $SE = 0.53$) than students in the Inspirational condition (*Adjusted M* = 11.66, $SE = 0.52$) (Mean Difference = 1.72, $SE = 0.75$, $p = .02$, 95% CI [0.23, 3.21], Cohen's $d = 0.48$) and students in the Instructional condition (*Adjusted M* = 11.38, $SE = 0.54$) (Mean Difference = 1.99, $SE = 0.69$, $p = .01$, 95% CI [0.48, 3.50], Cohen's $d = 0.61$) controlling for the pretest. The other planned comparisons between the conditions did not yield significant results. Overall, as Fig. 8 shows, the five ANCOVAs using the figural outcome variables show a pattern in the hypothesized direction. That is, the students assigned to the Both condition performed better than the students assigned to the other conditions relative to their creativity gains (on overall creativity, fluency, originality, and flexibility), but not for elaboration.

4.2.2. Level-design measures—four most creative levels

As discussed earlier in the Method section and elaborated in Appendix C, two raters independently scored students' four most creative (selected by each student) for relevance (0 or 1), line elaboration, meaningful object elaboration, originality, aesthetics, humor or surprise, and level's title's creativity (each variable, after relevance, was scored as 0, 1, or 2 depending on the level's creativity). Then, these seven scores were summed to compute the score of each level (maximum score possible = 13). Finally, after

Table 2
Descriptive statistics for the four conditions (n = 114).

Condition	Gender	*n*	Gaming Freq.	Gamer	Creativity
			M (*SD*)	*M* (*SD*)	*M* (*SD*)
Inspirational	Male	12	3.00 (1.54)	3.25 (1.14)	3.75 (0.75)
	Female	17	2.53 (1.51)	2.94 (1.03)	3.71 (1.05)
	Total	29	2.72 (1.51)	3.07 (1.07)	3.72 (0.92)
Instructional	Male	14	3.57 (1.28)	3.79 (0.80)	4.00 (0.88)
	Female	14	2.50 (1.61)	2.64 (1.15)	3.43 (0.85)
	Total	28	3.04 (1.53)	3.21 (1.13)	3.71 (0.90)
Both	Male	11	3.82 (0.87)	3.73 (0.91)	3.73 (0.91)
	Female	17	2.71 (1.69)	2.82 (0.95)	3.94 (0.75)
	Total	28	3.14 (1.51)	3.18 (1.02)	3.86 (0.80)
No Supports	Male	15	3.53 (1.36)	3.80 (1.32)	3.73 (0.88)
	Female	14	1.86 (1.17)	2.29 (0.73)	3.57 (0.94)
	Total	29	2.72 (1.51)	3.07 (1.31)	3.66 (0.90)
Total	Male	52	3.48 (1.29)	3.65 (1.06)	3.81 (0.84)
	Female	62	2.42 (1.51)	2.69 (0.99)	3.68 (0.90)
	Total	114	2.90 (1.51)	3.13 (1.13)	3.74 (0.87)

Table 3
Means, adjusted means, overall ANCOVAs for figural outcome variables.

Outcome V.	Condition	*n*	*Pre. M (SD)*	*Post. M (SD)*	*Adj. M (SE)*	95% CI		*F (df1, df2)*	*p*	η^2
						UB	LB			
Creativity	Inspirational	28	−.66 (1.53)	−.59 (2.1)	−.14 (.27)	−.68	.40	.51 (3, 108)	.68	.01
	Instructional	28	−.06 (1.77)	−.16 (1.93)	−.15 (.27)	−.69	.39			
	Both	28	.46 (1.96)	.61 (2.08)	.23 (.27)	−.31	.78			
	No Support	29	.04 (2.27)	−.10 (1.90)	−.17 (.27)	−.70	.36			
Fluency	Inspirational	29	16.64 (4.60)	16.25 (4.90)	15.57 (.70)	14.18	16.96	2.04 (3, 109)	.11	.05
	Instructional	28	16.01 (4.56)	15.93 (4.65)	15.70 (.71)	14.29	17.11			
	Both	28	14.84 (5.32)	16.91 (5.95)	17.52 (.71)	16.11	18.94			
	No Support	29	15.25 (6.78)	14.98 (5.91)	15.29 (.70)	13.90	16.68			
Originality	Inspirational	28	3.06 (1.47)	2.84 (1.76)	3.10 (.3)	2.50	3.69	1.48 (3, 106)	.22	.04
	Instructional	27	3.42 (1.92)	3.25 (1.99)	3.34 (.3)	2.74	3.94			
	Both	28	4.14 (2.54)	4.13 (1.82)	3.92 (.3)	3.32	4.51			
	No Support	28	3.93 (2.66)	3.32 (1.76)	3.20 (.3)	2.61	3.79			
Flexibility	Inspirational	29	12.64 (3.23)	12.05 (3.25)	11.66 (.52)	10.62	12.70	2.88 (3, 108)	.04*	.07
	Instructional	27	12.41 (3.04)	11.62 (2.54)	11.38 (.54)	10.31	12.46			
	Both	28	11.36 (3.51)	12.95 (3.89)	13.38 (.53)	12.32	14.43			
	No Support	29	11.69 (4.94)	11.48 (4.66)	11.69 (.52)	10.66	12.73			
Elaboration	Inspirational	29	5.94 (1.17)	5.94 (1.26)	6.34 (.18)	5.98	6.70	.76 (3, 109)	.52	.02
	Instructional	28	6.42 (1.13)	6.09 (1.12)	6.15 (.18)	5.80	6.51			
	Both	28	7.17 (1.38)	6.40 (1.45)	5.96 (.19)	5.59	6.33			
	No Support	29	6.57 (1.31)	6.29 (1.23)	6.26 (.18)	5.90	6.61			

Note. *$p < .05$; The F tests are for the condition effect on the outcome variables.

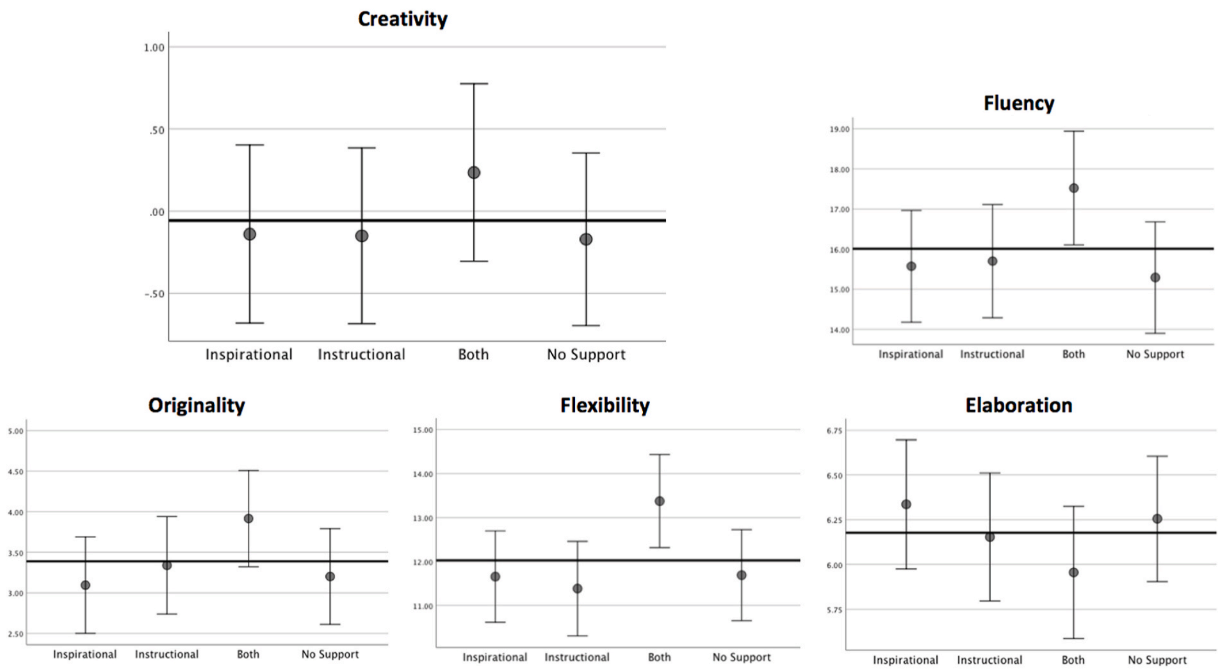

Fig. 8. Estimated marginal means of posttest figural scores (the grand means are marked as bold black line; Error bars: 95% CI).

averaging the two scores of each of the four levels from the raters per student, we averaged the scores of those four levels to get a single score of level-design creativity per student. We used the single score for our analyses.

To examine if the game levels that students created (see Fig. 9 for some examples) differed in terms of creativity, we conducted an ANCOVA with the average creativity score of the four most creative game levels as the dependent variable, condition as the independent variable, and the overall figural creativity score as the covariate. We chose the overall figural creativity score as the covariate because (a) it significantly correlated with the level-design measure of creativity, and (b) the nature of creating game levels was close to the figural tests—based on drawing. Results showed an overall significant effect of the condition [F (3, 109) = 19.13, $p < .001$, $\eta^2 =$ 0.36] relative to the creativity of the student-made levels. Table 4 shows the means and adjusted means of the four most creative game levels for each condition.

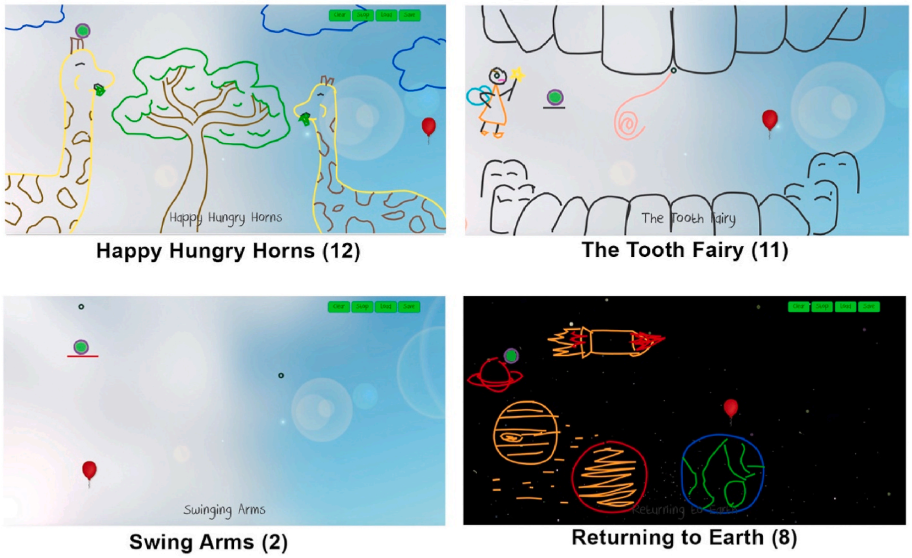

Fig. 9. Example student-made game levels with their creativity scores out of 13 possible.

Table 4
Means and adjusted means of the four most creative levels (n = 114).

Condition	*N*	*M* (*SD*)	*Adj. M* (*SE*)	95% *CI*	
				LB	UB
Inspirational	29	9.49 (1.45)	9.63 (.37)	8.89	10.36
Instructional	28	6.78 (2.32)	6.80 (.38)	6.06	7.54
Both	28	9.15 (2.11)	9.01 (.38)	8.26	9.75
No Support	29	6.25 (2.29)	6.24 (.37)	5.51	6.97

Note. Maximum possible score for game levels was 13.

Students in the Inspirational condition designed significantly more creative levels than students in the Instructional condition and the No Support condition controlling for their pretest overall figural creativity. There was no significant difference between the Both and Inspirational conditions. Similarly, there was no significant difference between the creativity scores of the four most creative levels of the Instructional condition and the No Support condition (Fig. 10). Overall, the results from the level-design measure of creativity indicate that most of the creativity shown in students' game levels come from the inspirational supports. Table 5 shows the pairwise comparisons between the conditions.

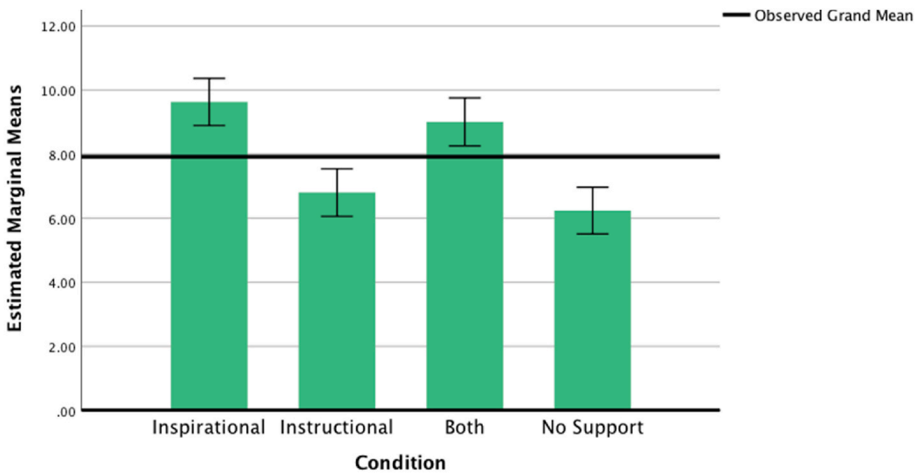

Fig. 10. Estimated marginal means of four most creative levels (Error bars: 95% CI).

Table 5
Pairwise comparisons—four most creative levels (n = 114).

Condition (I)	Condition (J)	*MD* (I-J)	*d*	*SE*	95% CI for Difference [b]	
					LB	UB
Inspirational	Instructional	2.83*	1.46	0.53	1.42	4.24
	Both	0.62	0.34	0.53	−0.80	2.05
	No Support	3.39*	1.77	0.52	1.99	4.79
Instructional	Inspirational	−2.83*	1.46	0.53	−4.24	−1.42
	Both	−2.21*	1.00	0.53	−3.63	−0.78
	No Support	0.56	0.24	0.53	−0.85	1.97
Both	Inspirational	−0.62	0.34	0.53	−2.05	0.80
	Instructional	2.21*	1.00	0.53	0.78	3.63
	No Support	2.77*	1.26	0.53	1.36	4.18
No Support	Inspirational	−3.39*	1.77	0.52	−4.79	−1.99
	Instructional	−0.56	0.24	0.53	−1.97	0.85
	Both	−2.77*	1.26	0.53	−4.18	−1.36

Note. Based on estimated marginal means; *$p < .001$; *b* Adjustment for multiple comparisons: Sidak. MD = Mean Difference.

4.2.3. *Level-design measures—the number of levels created (fluency)*

To find a good covariate to be used in the ANCOVA for examining the difference between the number of levels students created in the four conditions, we used the pretest figural fluency scores. The number of levels created by the students significantly and positively correlated with the pretest fluency scores ($r = 0.26$, $p < .01$) indicating that the pretest fluency scores could be used as a covariate for this ANCOVA. Results of the ANCOVA (number of levels as the dependent variable, condition as the independent variable, and pretest fluency score as the covariate) showed an overall significant difference by condition [F (3, 109) = 2.79, $p = .04$, $\eta^2 = 0.07$].

Students in the Instructional condition ($M = 7.82$, $SE = 0.75$) created significantly more levels than students in the Inspirational condition ($M = 4.95$, $SE = 0.74$) controlling for their pretest fluency (Mean Difference = 2.86, $p = .01$, 95% *CI* [0.78, 4.96], $\eta^2 = 0.04$, Cohen's $d = 0.74$)—see Fig. 11. The mean difference between the Both condition ($M = 6.97$, $SE = 0.75$) and the Inspirational condition ($M = 4.95$, $SE = 0.74$) approached significance (Mean Difference = 2.02, $p = .06$, 95% *CI* [−0.06, 4.11], Cohen's $d = 0.77$).

Moreover, students in the No Support condition ($M = 7.19$, $SE = 0.74$) created significantly more levels than students in the Inspirational condition ($M = 4.95$, $SE = 0.74$) controlling for their pretest fluency (Mean Difference = 2.25, $p = .03$, 95% *CI* [0.23, 4.27], $\eta^2 = 0.04$, Cohen's $d = 0.57$). The rest of the comparisons were not significant. These findings show that by providing proper instructions and guidelines, the quantity of the products also can increase.

4.2.4. *Three open-ended questions*

To test the relationships among students' feelings, satisfaction of time, flow experience, and their assigned condition, we computed three Pearson's Chi-Square tests of association (Fig. 12). There was no statistically significant relationship between how students *felt* during game design and the condition they were in (χ^2 (6) = 3.85, $p = .70$). That is, students in the four conditions reported the same in terms of how they felt during the game design process.

There appear to be a marginally significant relationship between *time* and condition (χ^2 (6) = 11.13, $p = .08$, Cramer's $V = 0.22$). Specifically, most students in the two conditions that received less instructions and had more time to create game levels (i.e.,

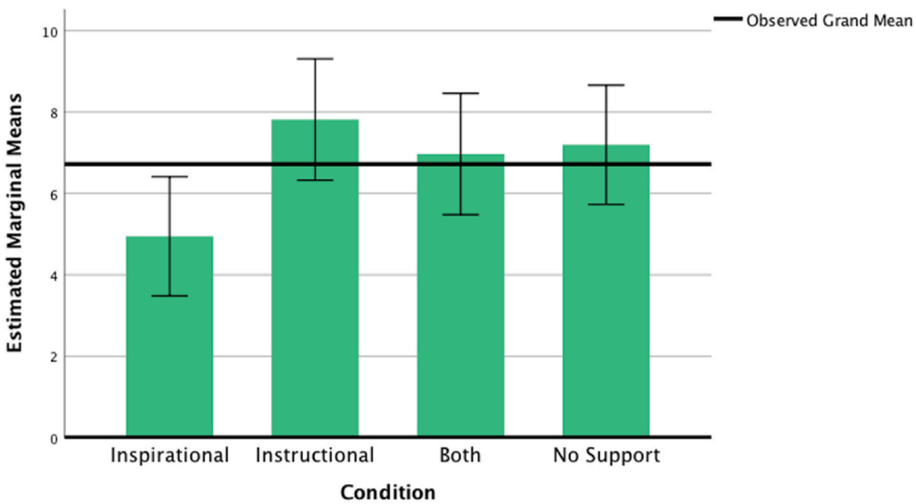

Fig. 11. Estimated marginal means of the number of levels created (Error bars: 95% CI).

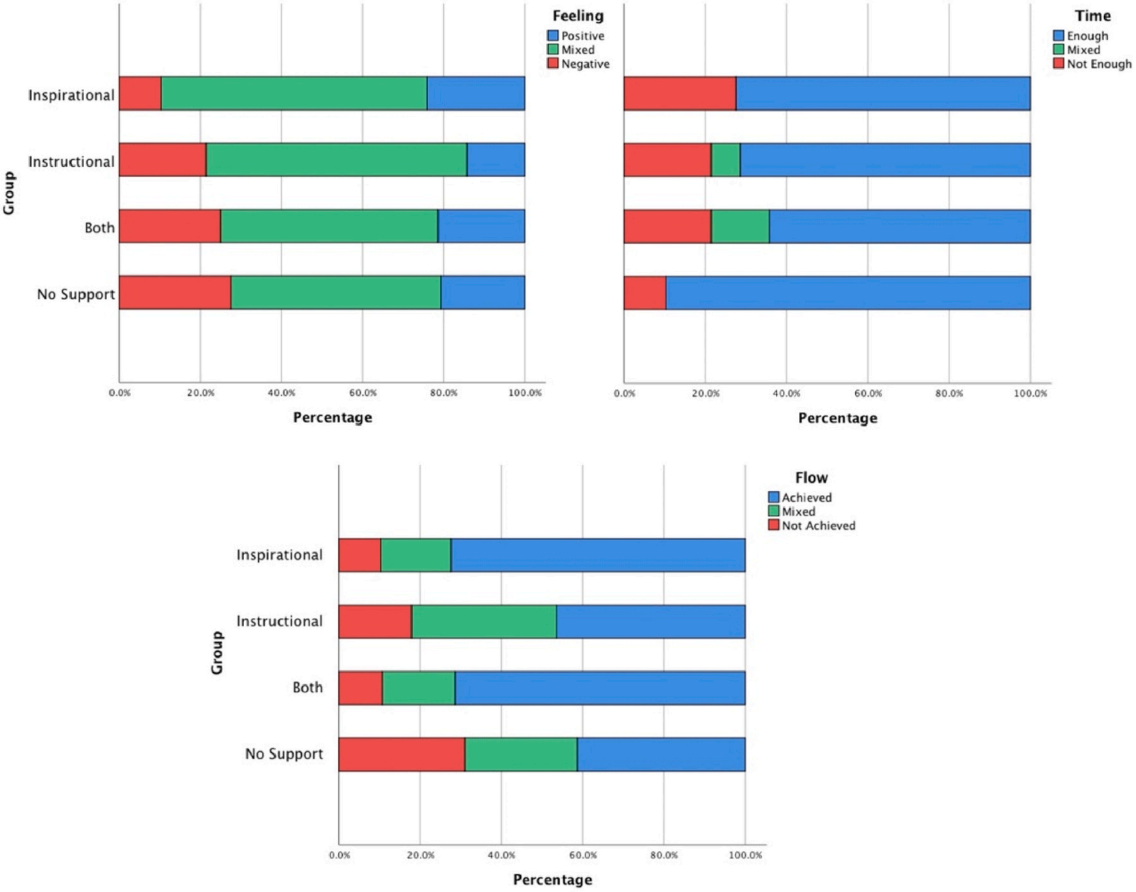

Fig. 12. Bar charts for students' responses to the three open-ended questions.

Inspirational and No Support) reported that time was enough for them (i.e., 72% in Inspirational and 90% in No Support condition). The percentage of students who reported the time was enough was lower in the other two conditions (i.e., 64% in Both and 71% in Instructional condition). Generally, most of the students (75%) felt the time was enough for creating their game levels. We can conclude that time was not a limitation for the students.

Similarly, there appears to be a relationship between experiencing *flow* (i.e., losing track of time) and condition (χ^2 (6) = 11.23, p = .08, Cramer's V = 0.22). Specifically, more students in conditions with inspirational supports reported that they lost track of time (i.e., 72% in the Inspirational and 71% in the Both conditions) than those who did not receive inspirational supports (31% in the No Support and 18% in the Instructional conditions). These findings are similar to the findings for the level-design measure of creativity (i.e., the four most creative game levels). The Inspirational and the Both conditions, and the Instructional and the No Support conditions scored the same on the level-design measure of creativity. Next, we discuss the findings, limitations, and future steps of this research.

5. Discussion

In this study, we investigated the effectiveness of a creativity-support system that we developed using *Physics Playground's* level editor on improving college students' creativity. The creativity supports used in this study were created based on two schools of thought—Inspirationism and Structurism (Shneiderman, 2007, 2009). In this section, we discuss the findings of this study for each research question.

5.1. Creativity supports' effectiveness

5.1.1. External measures of creativity

Based on the analyses using the external measures of creativity, the Both condition showed promising results. On all of the posttest figural measures (except elaboration) students in the Both condition scored higher than those in the other conditions, controlling for pretest scores. However, not all of the mean differences between the conditions were statistically significant. Specifically, the Both condition was significantly more effective on improving students posttest *flexibility* than the other conditions, controlling for the corresponding pretest. This finding aligns with what Moffat et al. (2017) found—a treatment effect only for flexibility. The other treatment conditions (i.e., Inspirational and Instructional) did not significantly differ from the No Support condition on the external measures (overall, and at the facet level). It is important to note that in this study we compared four versions of a game's level editor (with and without creativity supports). Thus, we could not compare the results of this study with experimental studies that investigated the effects of playing games on participants' creativity compared to a control condition that did not play a game (e.g., Blanco-Herrera et al., 2019; Gallagher & Grimm, 2018; Inchamnan et al., 2014). However, the findings of our study highlight the need for including creativity-supports in games with high potential for improving creativity (e.g., Minecraft, Portal 2) (see Blanco-Herrera et al., 2019; Checa-Romero & Gómez, 2018; Inchamnan et al., 2014).

These results do not support the hypothesis that participants in conditions with supports would perform better on the posttest compared to the participants in the No Support condition. Generally, research in the human-computer interaction field has shown that creativity-support systems can help people enhance their creativity (Frich et al., 2018; Massetti, 1996; Remy et al., 2020; Resnick et al., 2005; Ristow, 1988; Shneiderman, 2007). However, those studies differ from the current study as this study required participants to design game levels while participants in the other studies only had to only generate ideas. Based on our results from the external measures of creativity in this study, providing only one type of creativity support (inspirational or instructional) is about the same as providing no supports. However, providing *both* supports together can potentially help improve participants' creativity. This claim needs to be tested further in studies with longer exposure to the treatment (e.g., one week vs. one 3-h session).

These results, however, partially support our other hypothesis (i.e., that the Both condition would be the most effective, followed by the Inspirational, then Instructional, and then the No Support condition). This finding suggests that supporting students' creativity in game design environments via both inspiring and instructing them through a step-by-step process can improve their creativity. The results only partially support the hypothesis because (a) not all of the statistical tests showed that the Both condition was significantly superior to the other conditions, and (b) the ordering of the supports' effectiveness was not as expected. In general, with regard to the external measures, the pattern of the Both condition showing promising results in the expected direction and the other three conditions being comparable emerged.

There are some mechanisms that likely helped students in the Both condition transfer more creativity on the posttest controlling for the pretest compared to the students in the other three conditions. To explain those mechanisms, we discuss the findings from the level-design measures of creativity and the results from the three open-ended questions next.

5.1.2. Level-design measures of creativity

We compared the creativity of the four most creative game levels that students created in the four conditions. Results showed that participants in the Inspirational and the Both conditions designed levels that were more creative than levels designed by those in the Instructional and No Support conditions, controlling for figural creativity pretest scores. There was no significant difference between students' creativity scores in the Inspirational and the Both conditions, nor between those in the Instructional and No Support conditions. This finding indicates that most of the creativity in the Inspirational and the Both conditions come from what they shared—*inspirational* supports rather than the *instructional* supports. These findings provide more empirical evidence for the effectiveness of creativity support techniques such as brainstorming (e.g., Frich et al., 2018; Litchfield, 2008; Putman & Paulus, 2009), analogy-making and remote association techniques (e.g., Casakin, 2004; Casakin & van Timmeren, 2014), and reviewing previous examples for inspiration (e.g., Resnick et al., 2005; Sio et al., 2015).

One reason for the similar creativity scores for the Inspirational and Both conditions is that almost the same percentage of students in these conditions reported experiencing flow (i.e., losing themselves in the experience and feeling that time has gone by quickly). As we indicated in the introduction, flow and creativity can have a cyclical relationship. That is, ongoing involvement in creative work can enhance one's affective state (Cseh et al., 2015) via a sense of progress and productivity (Amabile, 1983) which in turn can facilitate and maintain the state of flow (Cseh et al., 2015) which can help generate more creative work. One reason that the students of these two conditions (Inspirational and Both) experienced the state of flow was that they were inspired by viewing previous level examples, thus getting a clear goal in mind. This facilitated progress which rendered good feelings (Amabile, 1983).

The results from the number of game levels students created (i.e., a proxy for fluency), however, showed that students in the Instructional and the No Support conditions created significantly *more* game levels than students in the Inspirational condition but not significantly more levels than those in the Both condition, again controlling for their pretest figural fluency. The mean difference between the Both and Inspirational conditions' number of game levels created approached significance (with the Both condition creating more game levels than the Inspirational condition). Students in all four conditions had about 2 h to create game levels and they were all told to "create as many creative levels as you can." Therefore, it is reasonable to expect that students in the No Support condition would create more levels during the 2 h than the other conditions because they did not have to attend to any supports. As such, they may have focused more on the "many" levels part of the instructions, and less on the creativity of their levels.

The most similar condition to the No Support condition was the Instructional condition where students were told to design as many levels as possible without paying attention to details, pick four levels, and enhance them. Students in this condition focused more on quantity rather than quality of levels because they did not have any inspirational supports. It is important to note that the instructional supports were useful in increasing the number of game levels students created for the Instructional condition as the students in this condition only had about 65 min to increase the number of their game levels compared to the students in the No Supports condition who had about 120 min to do the same. Yet, the two conditions created similar number of game levels on average. This finding shows that the instructional supports alone could improve some aspects of creativity (e.g., fluency) but are not enough for overall creativity (as discussed earlier).

Despite having less time, students in the Both condition created the same number of game levels compared to the students in the

Instructional and the No Support conditions. Students in the Both condition knew what they wanted to create (they had their ideas written down on the brainstorming spreadsheet) and they also had a plan (instructions) to create as many levels as possible. However, students in the Inspirational condition, who created the fewest game levels compared to the other conditions, only focused on increasing the quality of their game levels and did not focus on the quantity of them—they lacked instructional supports.

When we considered the reason for this finding (i.e., that students in the Both condition created significantly more levels than the Inspirational condition and a similar number of levels compared to the other two conditions), it seemed to relate to the instructional supports received. Research has shown how important it is to provide clear instructions when the goal is to enhance people's creativity (e.g., Chua & Iyengar, 2008; Crutchfield & Covington, 2009; Jaben, 1985; Niu & Liu, 2009). For instance, students in the Both condition were told to "create as many levels as you can in the first 40 min" However, these instructions were given to all four conditions. Additional instructions given to students in the Both condition likely caused the difference we saw in terms of levels solved. That is, they were told to, "divide your time intro three parts: (a) create as many levels as you can without paying attention to details, (b) choose four or them, and then (c) enhance those four levels". These instructions provided structure for students that guided them through the creativity process (from idea generation to elaboration and enhancement of their final product). This type of instruction is what Shneiderman (2007) refers to as structural supports and we refer to in this study as instructional supports. At first one might think that providing structure for a creative process is limiting and paradoxical. However, our results and previous research (e.g., Crutchfield & Covington, 2009) have shown that it is beneficial to provide such structure for improving people's creativity.

When we speculated the reasons for this finding (students in the Both condition created significantly more levels than the Inspirational condition and a similar number of levels compared to the other two conditions), we thought that the main reason was the instructional supports they received. Research showed how important it is to provide clear instructions when the goal is to enhance people's creativity (e.g., Chua & Iyengar, 2008; Crutchfield & Covington, 2009; Jaben, 1985; Niu & Liu, 2009). For instance, students in the Both condition received instructions such as "be creative and created as many creative levels as you can." However, the instruction we just mentioned was given to students in the four conditions. It was the further instructions that students in the Both condition received but the students Inspirational conditions did not which made the difference (i.e., "divide your time intro three parts: first created as many levels as you can without paying attention to details, choose four or them, and then enhance those four levels". These instructions provided some structure for students and guided them through the creativity process (from idea generation to elaboration and enhancement of their final product). This type of instruction is what Shneiderman (2007) refers to as structural supports and we referred to, in this study, as instructional supports. At first, one might think that providing any structure for a creative process is limiting and paradoxical. However, our results, and previous research (e.g., Crutchfield & Covington, 2009), showed that it is beneficial to provide such structure for improving people's creativity.

The findings about the number of levels students created partially support our hypothesis that students in all of the conditions with supports (Inspirational, Instructional, and Both) would perform better on the posttests compared to the No Support condition. That is, we found that: (a) the students in two of the treatment conditions (i.e., Inspirational and Both) showed more creativity than the ones in the No Support condition for the four most creative levels, and (b) participants in none of the treatment conditions created significantly more game levels than those in the No Support condition. Considering the two level-design measures of creativity (i.e., four most creative levels' creativity scores and the number of levels created by each condition), students in the Both condition scored higher on quality (i.e., four most creative game levels) compared to students in the No Support condition while staying competitive in terms of quantity (i.e., number of levels they created). This finding, therefore, supports our hypothesis suggesting that the Both condition would be the most effective one based on both of the level-design measure of creativity—similar to the findings from on the external measures of creativity.

Considering the results from the open-ended question about flow and time, we see that student in the Inspirational and the Both conditions reported similar ratings of flow. Losing track of time could be interpreted as students being fully engaged and not thinking about time rather than not having enough time (the majority of the students in the four conditions reported they had enough time for their tasks). Therefore, experiencing flow could not be the reason for the relatively low number of levels the students in the Inspirational condition created. Rather, the lack of instructional supports seems to be the reason for the low number of levels, relative to the other conditions, created by the students in the Inspirational condition.

In sum, students in the Both condition enhanced their game levels' creativity by using the inspirational supports and could also increase the number of their levels using the instructional supports. These findings support what Shneiderman (2009) has proposed—i.e., that one can develop creativity-support tools that include ideas from two, or all the three schools of thought. Moreover, the findings of this study support the fact that creative thinking is about both divergent and convergent thinking (Cropley, 2006). Use of both inspirational and instructional supports appear to optimally foster these two types of thinking for students in the Both condition which we explain in more detail next.

In general, six main reasons for the effectiveness of the Both condition include the following:

(1) *Practice*: Students in the Both condition received more practice being creative than those in the other conditions. For instance, they created more game levels than the students in the Inspirational condition which could enhance their fluency and flexibility skills, and they created levels with higher levels of creativity which could affect their originality skill. Creativity, like a muscle, can be trained, exercised, and strengthened by practice (Kaufman & Beghetto, 2009; Richards, 2010; Sternberg & Lubart, 1996). The purpose of providing both inspirational and structural supports in the Both condition was to help students exercise both divergent thinking (i.e., designing as many levels as possible using instructional supports) and convergent thinking (i.e., designing as detailed and original levels as possible using inspirational supports) skills. Therefore, through more practicing of

being creative than the other conditions, students in the Both condition increased both their in-game and external scores of creativity.

(2) *Focus on stages 1 through 4 of the creativity process:* Amabile's componential model of creativity (Amabile, 1983; Amabile & Pratt, 2016) (Amabile, 1983; Amabile & Pratt, 2016) includes five stages (as we discussed earlier). Amabile argues that Stages 1 (task presentation or problem formulation) and 3 (idea generation) can help to enhance the number and novelty of ideas, and Stages 2 (preparation) and 4 (idea validation) can help with the usefulness or appropriateness of ideas. The supports that students in the Both condition received thus facilitated these four stages leading to an increase in both quality and quantity of the levels that students created in the Both condition.

(3) *Familiarity with the possibilities and going beyond them*: Students in the Both condition were likely inspired by viewing existing game levels, thus allowing them to see what was possible in *PP*, and possibly go further. Theories of creativity have pointed out that to be creative, one must know what has been done to be able to go beyond it (Csikszentmihalyi, 1997; Kaufman & Beghetto, 2009). Moreover, students in the Both condition had access to existing *PP* levels with low, medium, and high levels of creativity. These three categories of game levels provided students an opportunity to understand what is and what is *not* creative. Therefore, they could avoid creating game levels that were scored low on creativity. Also, viewing the existing levels also helped the students in the Both condition be prepared for what they were about to do—designing game levels for Physics Playground. Preparation is one of the most important steps in any creativity process (Amabile & Pratt, 2016; Sadler-Smith, 2015).

(4) *Reduced cognitive load because of the instructional supports*: Compared to the Inspirational condition, participants in the Both condition learned how to use the level editor easier which could have led to less cognitive load—cognitive load is detrimental to creativity (Shneiderman, 2007); especially for novices (e.g., Sun & Yao, 2012)—cognitive load is detrimental to creativity; especially for novices. Specifically, students in the Both condition had to divide their time into three parts. First, they had to lay the foundations of their game levels without paying attention to details. This stage probably played an important role in reducing students' cognitive load which could increase if students aimed to complete their game levels initially.

(5) *Boost in their affective state*: As students completed the game levels listed in their brainstorming spreadsheet, they experienced a feeling of making progress which according to Amabile (2017) research, can positively affect their affective state, which likely led to more creativity on the posttests.

(6) *Flow experience*: 71% of the student in the Both condition reported having lost track of time (the same as the Inspirational condition). The reasons mentioned above probably helped the majority of students in the Both condition reach the state of flow (Csikszentmihalyi, 1997). It is at this stage that a positive feedback loop can occur—being in the state of flow can lead to more creativity; continuous production of creative work can lead to maintaining the flow which can lead to more creativity and so on—as we discussed earlier in the introduction, flow and creativity can be linked through affect (Amabile, 1983, 2017; Cseh, Phillips, & Pearson, 2015).

Lack of all or some of these reasons could have led to the other conditions' ineffectiveness.

5.2. Implications

The findings of this study can inform the development of computer- and web-based creativity-support tools that can help people think and act more creatively in different situations (Shneiderman, 2009). For instance, virtual reality (VR) environments (e.g., VR painting applications such as Tilt Brush by Google where people can create 3D paintings) provide numerous affordances for assessing and improving creativity. We used a game level-editor in this study as a tool; however, the findings of this study can be useful for other computer-based applications that can enhance people's creativity (e.g., Frich et al., 2018; Remy et al., 2020). Moreover, the results can provide useful information about the relative effectiveness of inspirational and instructional creativity supports in video games that demand creativity through similar level-design activities (e.g., *Minecraft, Portal 2, Crayon Physics Deluxe*). We also see the results of our study as useful for research involving computer-based learning environments.

However, the results from our study should be generalized to other contexts with caution. First, the participants of our study were college students. Testing the creativity supports in studies using younger students could increase the generalizability of our findings. Second, participants of this study could be affected by the COVID-19 pandemic's side effects (e.g., anxiety, low motivation, stress or other mental-health-related issues) which could affect their creativity in unpredictable ways. Additionally, students participated from home using various computers, environments, and internet connections that could affect their performance and creativity. Replicating this study in a normal situation in a controlled laboratory setting after the pandemic could increase the generalizability of the results of our study. However, we can argue that the random assignment of the participants to the four conditions has controlled the possible negative environmental effects on the validity of the results in our study to a reasonable extent. This study's findings could be more deterministic and generalizable (especially regarding the external measures of creativity) without the limitations we faced during this study—discussed next.

6. Limitations and future research

Due to the COIVD-19 pandemic, we had to reduce the number of test items per test form (from 4 to 2 for figural), the duration of the study from 6 h across two days to 3 h in one day, and from face-to-face to online administration of the study. These changes may have added some noise to the data and weakened the strength of the treatments. Another limitation of this study could be the sample size. Despite the power analysis we conducted showing 100 students were enough, it appears that the dosage of the treatment was rather low, thus a larger sample size could make the difference between the conditions more apparent than what was found. Finally, students reported some technical issues (e.g., their system being slow, their browser crashing due to memory overflow, and some technical issues with the level editor) that could make students frustrated. If the study was done face-to-face, we were planning to provide the same personal computer for each student which could reduce these technical issues considerably.

The first direction for future research would be to replicate this study without the limitations mentioned above. For example, one could conduct a replication study with more test items per test form, longer duration (over a week), larger sample size, and fewer technical issues. Alternatively, conducting the same study on a younger population (e.g., middle and secondary school students), and using other games (with high potential for creativity) such as *Minecraft* or *Portal 2* could test if the results were replicated, useful for generalizability purposes.

The second direction for future research could be toward an automated assessment method of creativity. For example, studies may be conducted that use log data to investigate how students actually use the supports (e.g., recording the number times that a student uses the creativity supports which provide a deeper understanding about how useful the supports are). Moreover, using assessment method such as stealth assessment of creativity (e.g., Shute & Rahimi, 2020) within *PP*'s level editor could result in an automated, diagnostic assessment of creativity in real time. For example, for elaboration, identifying and counting the number of meaningful objects (e.g., a flower, a car) as students create them could be used as one variable. These assessment methods are objective, just in time, and could be used to provide individualized supports for students during the game design process.

The last direction for future research involves using the best creativity support system from the current study (i.e., both inspirational and instructional supports) in the context of the third school of thought for enhancing creativity, which Shneiderman (2007, 2009) refers to as *situationism*. This suggests that creativity may be improved in a social environment where students interact with each other. For example, one can create an online community where students can create their game levels using *PP*'s level editor, share their levels with others, and get feedback from them. In such an online community, assessment of creative game levels can be outsourced. For example, students can rate other students' game levels' creativity, and in a short amount of time, the most creative game levels could be identified. If 90% of the students in the online community think a game level is highly creative (given the right rubric), we can accept that game level is creative. Csikszentmihalyi (1997) in his system theory of creativity suggests that it is the field (i.e., people who are working in the domain at hand) who decides if a product is creative. The effectiveness of this social vs. individual enhancement method of creativity could be compared in an experimental study.

7. Conclusion

Creativity is one of the most important skills needed for success in life in the 21st-century. Consider the current situation with the COVID-19 pandemic. People need to be creative to keep their businesses open, to stay socially connected yet physically distanced, to continue teaching and learning, to find or generate hope for themselves and their communities, and in short, creativity can help them overcome their challenges in these difficult times. With this in mind, we as researchers and educators have an obligation to create tools that can assess and foster people's creativity in the coming years. Excuses such as "creativity is complex, hard to define, measure, and enhance" are no longer acceptable. With the advances in psychology, learning sciences, technologies, and computer sciences, we can create the tools we need to meet this obligation—we just need to be creative.

Author contribution

Seyedahmad Rahimi: Conceptualization, Data curation, Funding acquisition, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draftand Writing – review & editing. **Valerie Shute:** Supervision, Resources, and Writing – review & editing.

Declaration of competing interest

We wish to confirm that there are no known conflicts of interest associated with this manuscript.

Acknowledgments

This work was supported by Florida State University's Graduate School. We also want to thank Russell Almond, Fengfeng Ke, Walter Boot, and James Kaufman for their feedback on this paper. Finally, we want to thank Xiaotong Yang, Renata Kuba, Yunkai Xu, and Maedeh Agharazi for helping us with the scoring process of the data.

Appendix A

Some examples of SCAMPER questions/suggestions that we wrote for this study include:

1. Substitute

- Pick one object you included in this level and ask yourself, "what else can be used instead of [that item] to make the level more creative?"

- If you are illustrating a place at this level (e.g., a beach), ask yourself, "what other places can be shown instead that no one else would think of?"

2. *Combine*

- Do you have more than one object at this level? What if you combine two or more objects in this level to create a new object?

3. *Adapt*

- What about using some of your previous ideas in previous levels (if you created some levels already) for this level?
- What can you use or get ideas from when you look around you?
- What other objects in real-life can you think of to get ideas from for this level?

4. *Modify*

- What can you add? Magnify? Minify?
- Can you redraw any of the objects to increase the quality of your level?
- What color(s) to change?
- What objects can you move to a new location to make the level harder or more interesting?

5. *Put to other uses*

- Can you change the context of the level? For example, if the level you designed has a scary theme, can you think of a new level with a fun theme? What about creating a level with a scientific theme (i.e., use scientific objects like lab equipment)?
- What other themes can you think of?

6. *Eliminate*

- What is unnecessary or redundant? Remember, "less is more." Try to see if you can remove any objects/parts from this level to make it clearer or more interesting.

7. *Rearrange*

- Sometimes, if you change your perspective, you may see new and interesting things. For example, you can redraw an object with a different angle and create a new object. Can you redraw any object with a new angle (e.g., upside-down) to make this level more interesting?
- What other arrangements of objects can you use at this level?

Appendix B

For elaboration for circles/diamonds:

0 not at all (only 1 or 2 lines; it is hard to understand what the drawing is and you need to read the title)
1 a little (more than 2 lines less than 5 lines used; and it is that easy to understand what the drawing is and you still need to read the title)
2 elaborated (more than 5 lines less than 10 lines; it is easy to understand what the drawing is without reading the title)
3 a lot (more than 10 lines; but no artistic details like shadows; clearly you can understand what the drawing is)
4 Beyond-the-norm (a lot of details, artistic details like shadows, clearly you can understand what the drawing is)

For elaboration for circles/diamonds:

0 not at all (only 1 or 2 shapes are used; it is hard to understand what the drawing is and you need to read the title)
1 a little (only 3 shapes used; and it is not that easy to understand what the drawing is and you still need to read the title)
2 elaborated (more than 3 shapes less than 5 shapes are used; it is easy to understand what the drawing is without reading the title)
3 a lot (more than 5 shapes less than 7; clearly you can understand what the drawing is)
4 Beyond-the-norm (more than 7 shapes are used, artistic details like shadows, clearly you can understand what the drawing is)

Appendix C

Scoring rubric for levels' creativity (modified from Shute and Wang, 2016).

Subscale	Scoring rules
Relevance	Can it be solved? (This is a screening criterion) o Unsolvable, then do not score other variables = 0 o Solvable, continue scoring other variables = 1
Lines Elaboration	*Is it well elaborated?* (Possible scores: 0, 1, 2) • Lines: o If 0 lines are on the screen = 0 o If 1 to 5 lines = 1 o If more than 5 lines = 2
Meaningful Objects Elaboration	*Is it well elaborated?* (Possible scores: 0, 1, 2) • Meaningful Objects (a simple line is not meaningful; a flower is): o If no objects are present = 0 o If 1 to 5 meaningful objects are present = 1 o If more than 5 objects are present = 2
Originality	*Is it original relative to existing levels?* (Possible scores: 0, 1, and 2) o Almost identical = 0 o Has some similarities = 1 o Very dissimilar = 2
Aesthetics	*Is it aesthetically pleasing?* (Possible scores: 0, 1, and 2) • Aesthetically unappealing with poor visual elements = 0 o Poorly drawn objects o 1 to 2 colors used o Colors are used in a disturbing way (in a way that it hurts to look at the level) • Plain with completed visual elements = 1 o Complete objects are drawn o 3 to 4 colors are used • Very pleasant with well-thought-out visual elements = 2 o Well-designed objects are drawn o More than four colors are used
Humor or Surprise	*Is it humorous or surprising* (i.e., Does it make you smile or do you recognize the level's idea in a way that it generates a sense of surprise for you)? (Possible scores: 0, 1, and 2) o Not humorous or surprising at all = 0 o Somewhat humorous or surprising = 1 o Very humorous or surprising = 2
Title's Creativity	*Is the title creative and related to the level?* (Possible scores: 0, 1, 2) o There is no connection between the level and the title = 0 o There is weak connection and the title is just descriptive = 1 o When one or more than one of the things below were true = 2 o a remote association was used (If the title is related to the level but not in an obvious way) o if an analogy was made o or if the title helps the reader understand the level that was not possible without the title o or if the title creates a sense of surprise o or if the title makes the reader smile/laugh

Appendix D

Codes for three open-ended questions.

Feelings	Time	Flow
Negative = 1	Not Enough = 1	Not Achieved = 1
Mixed = 2	Mixed = 2	Mixed = 2
Positive = 3	Enough = 3	Achieved = 3

Appendix E

Means adjusted means of number of levels created (n = 114)

Condition	*N*	*M* (*SD*)	*Adj. M* (*SE*)	95% *CI*	
				LB	UB

(*continued on next page*)

(*continued*)

Means adjusted means of number of levels created (n = 114)					
Condition	*N*	*M* (*SD*)	*Adj. M* (*SE*)	95% *CI*	
				LB	UB
Inspirational	29	4.83 (2.05)	4.95 (.74)	3.48	6.41
Instructional	28	7.96 (5.07)	7.82 (.75)	6.32	9.31
Both	28	6.82 (3.10)	6.97 (.75)	5.48	8.46
No Support	29	7.31 (5.17)	7.19 (.74)	5.73	8.66

References

Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology, 45*(2), 357. https://doi.org/10.1037/0022-3514.45.2.357

Amabile, T. M. (1996). *Creativity in context: Update to the social psychology of creativity* (New edition). Routledge.

Amabile, T. M. (2017). In pursuit of everyday creativity. *The Journal of Creative Behavior, 51*(4), 335–337. https://doi.org/10.1002/jocb.200

Amabile, T. M., & Pratt, M. G. (2016). The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior, 36*, 157–183. https://doi.org/10.1016/j.riob.2016.10.001

Baer, J. (1994). Divergent thinking is not a general trait: A multidomain training experiment. *Creativity Research Journal, 7*(1), 35–46. https://doi.org/10.1080/10400419409534507

Baer, J. (1996). The effects of task-specific divergent-thinking training. *The Journal of Creative Behavior, 30*(3), 183–187. https://doi.org/10.1002/j.2162-6057.1996.tb00767.x

Barbot, B. (2011). Assessing creativity in the classroom. *The Open Education Journal, 4*(1), 58–66. https://doi.org/10.2174/1874920801104010058

Belsky, S. (2020). *Creativity will be key to competing against AI in the future workforce—here's how*. World Economic Forum. https://www.weforum.org/agenda/2020/11/ai-automation-creativity-workforce-skill-fute-of-work/.

Blanco-Herrera, J. A., Gentile, D. A., & Rokkum, J. N. (2019). Video games can increase creativity, but with caveats. *Creativity Research Journal, 31*(2), 119–131. https://doi.org/10.1080/10400419.2019.1594524

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives, Handbook 1: Cognitive domain* (2nd ed.). Addison-Wesley Longman Ltd.

Bonnardel, N. (1999). Creativity in design activities: The role of analogies in a constrained cognitive environment. *Proceedings of the 3rd Conference on Creativity & Cognition*, 158–165.

Bopp, J. A., Opwis, K., & Mekler, E. D. (2018). "An odd kind of pleasure": Differentiating emotional challenge in digital games. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1–12). https://doi.org/10.1145/3173574.3173615

Bowman, N. D., Kowert, R., & Ferguson, C. J. (2015). The impact of video game play on human (and Orc) creativity. In J. C. Kaufman, & G. P. Green (Eds.), *Video games and creativity* (pp. 39–61). Elsevier. https://doi.org/10.1016/B978-0-12-801462-2.00002-3.

Casakin, H. (2004). Visual analogy as a cognitive strategy in the design process: Expert versus novice performance. *Journal of Design Research, 4*(2). https://doi.org/10.1504/JDR.2004.009846, 0.

Casakin, H., & van Timmeren, A. (2014). Analogies as creative inspiration sources in the design studio: The teamwork. *Athens Journal of Architecture, 1*(1), 51–64. https://doi.org/10.30958/aja.1-1-4

Checa-Romero, M., & Gómez, I. P. (2018). Minecraft and machinima in action: Development of creativity in the classroom. *Technology. Pedagogy and Education, 27*(5), 625–637. https://doi.org/10.1080/1475939X.2018.1537933

Chua, R. Y.-J., & Iyengar, S. S. (2008). Creativity as a matter of choice: Prior experience and task instruction as boundary conditions for the positive effect of choice on creativity. *The Journal of Creative Behavior, 42*(3), 164–180. https://doi.org/10.1002/j.2162-6057.2008.tb01293.x

Cipollone, M., Schifter, C. C., & Moffat, R. A. (2014). Minecraft as a creative tool: A case study. *International Journal of Game-Based Learning, 4*(2), 1–14. https://doi.org/10.4018/ijgbl.2014040101

Corbalán, J., Martínez, F., Donolo, D., Alonso, C., Tejerina, M., & Limiñana, M. R. (2003). *CREA. Creative intelligence, a cognitive measure of creativity*. TEA Editions.

Cropley, A. (2006). In praise of convergent thinking. *Creativity Research Journal, 18*(3), 391–404. https://doi.org/10.1207/s15326934crj1803_13

Cropley, D. H., & Cropley, A. J. (2011). Aesthetics and creativity. In M. A. Runco, & S. R. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 24–28). Academic Press/Elsevier. https://doi.org/10.1016/B978-0-12-375038-9.00004-2.

Crutchfield, R. S., & Covington, M. V. (2009). Programed instruction and creativity. Theory into Practice. https://doi.org/10.1080/00405846609542022.

Cseh, G. M., Phillips, L. H., & Pearson, D. G. (2015). Flow, affect and visual creativity. *Cognition & Emotion, 29*(2), 281–291. https://doi.org/10.1080/02699931.2014.913553

Csikszentmihalyi, M. (1997). *Creativity: Flow and the psychology of discovery and invention*. Basic Books.

Eberle, R. F. (1972). Developing imagination through scamper. *The Journal of Creative Behavior, 6*(3), 199–203. https://doi.org/10.1002/j.2162-6057.1972.tb00929.x

Entertainment Software Association. (2019). Essential facts about the computer and video game industry. https://www.theesa.com/wp-content/uploads/2019/05/2019-Essential-Facts-About-the-Computer-and-Video-Game-Industry.pdf.

Frich, J., Biskjaer, M. M., & Dalsgaard, P. (2018). Why HCI and creativity research must collaborate to develop new creativity support tools. *Proceedings of the Technology, Mind, and Society, 1–6*. https://doi.org/10.1145/3183654.3183678

Frich, J., MacDonald Vermeulen, L., Remy, C., Biskjaer, M. M., & Dalsgaard, P. (2019). Mapping the landscape of creativity support tools in HCI. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–18). https://doi.org/10.1145/3290605.3300619

Gallagher, D., & Grimm, L. R. (2018). Making an impact: The effects of game making on creativity and spatial processing. *Thinking Skills and Creativity, 28*, 138–149. https://doi.org/10.1016/j.tsc.2018.05.001

Gee, J. P. (2005). Learning by design: Good video games as learning machines. *E-Learning and Digital Media, 2*(1), 5–16.

Glăveanu, V. P. (2013). Rewriting the language of creativity: The five A's framework. *Review of General Psychology, 17*(1), 69–81. https://doi.org/10.1037/a0029528

Glăveanu, V. P., Hanson, M. H., Baer, J., Barbot, B., Clapp, E. P., Corazza, G. E., Hennessey, B., Kaufman, J. C., Lebuda, I., Lubart, T., Montuori, A., Ness, I. J., Plucker, J., Reiter-Palmon, R., Sierra, Z., Simonton, D. K., Neves-Pereira, M. S., & Sternberg, R. J. (2020). Advancing creativity theory and research: A socio-cultural manifesto. *The Journal of Creative Behavior, 54*(3), 741–745. https://doi.org/10.1002/jocb.395

Gray, A. (2016). *The 10 skills you need to thrive in the Fourth Industrial Revolution*. World Economic Forum. https://www.weforum.org/agenda/2016/01/the-10-skills-you-need-to-thrive-in-the-fourth-industrial-revolution/.

Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin, 53*(4), 267. https://doi.org/10.1037/h0040755

Hall, J., Stickler, U., Herodotou, C., & Iacovides, I. (2021). Using reflexive photography to investigate design affordances for creativity in digital entertainment games. *International Journal of Human-Computer Interaction, 37*(9), 867–883. https://doi.org/10.1080/10447318.2020.1848162

Hamlen, K. R. (2009). Relationships between computer and video game play and creativity among upper elementary school students. *Journal of Educational Computing Research, 40*(1), 1–21. https://doi.org/10.2190/EC.40.1.a

Hennessey, B. A., & Amabile, T. M. (2010). Creativity. *Annual Review of Psychology, 61*(1), 569–598. https://doi.org/10.1146/annurev.psych.093008.100416

Hofstadter, D. R. (2001). Epilogue: Analogy as the core of cognition. In D. Gentner, K. L. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 499–538). MIT Press.

Huberty, C. J., & Morris, J. D. (1989). Multivariate analysis cersus multiple univariate analyses. *Psychological Bulletin, 105*(2), 302–308.

Inchamnan, W., Wyeth, P., & Johnson, D. (2013). Does activity in computer game play have an impact on creative behaviour? *2013. IEEE International Games Innovation Conference (IGIC)*, 77–84. https://doi.org/10.1109/IGIC.2013.6659169

Inchamnan, W., Wyeth, P., & Johnson, D. (2014). Design for creative activity: A framework for analyzing the creative potential of computer games. In C. Salinesi, M. C. Norrie, & Ó. Pastor (Eds.), *Advanced information systems engineering* (Vol. 7908, pp. 19–26). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-45212-7_3.

Islim, O. F., & Karataş, S. (2016). Using the Scamper technique in an ICT course to enhance creative problem solving skills: An experimental study. *Turkish Online Journal of Educational Technology*, 1291–1296.

Jaben, T. H. (1985). Effect of instruction for creativity on learning disabled students' drawings. *Perceptual & Motor Skills, 61*(3), 895–898. https://doi.org/10.2466/pms.1985.61.3.895

Jackson, P. W., & Messick, S. (1965). The person, the product, and the response: Conceptual problems in the assessment of creativity. *Journal of Personality, 33*(3), 309–329. https://doi.org/10.1111/j.1467-6494.1965.tb01389.x

Jackson, L. A., Witt, E. A., Games, A. I., Fitzgerald, H. E., von Eye, A., & Zhao, Y. (2012). Information technology use and creativity: Findings from the children and technology Project. *Computers in Human Behavior, 28*(2), 370–376. https://doi.org/10.1016/j.chb.2011.10.006

Karsenti, T., & Bugmann, J. (2017). Exploring the educational potential of Minecraft: The case of 118 elementary-school students. *International Association for Development of the Information Society*, 175–179.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four-c model of creativity. *Review of General Psychology, 13*(1), 1–12. https://doi.org/10.1037/a0013688

Kaufman, J. C., & Sternberg, R. J. (2007). Resource review: Creativity. *Change, 39*(4), 55–58 (JSTOR).

Kim, K. H. (2006). Can we trust creativity tests? A review of the torrance tests of creative thinking (TTCT). *Creativity Research Journal, 18*(1), 3–14. https://doi.org/10.1207/s15326934crj1801_2

Kim, Y. J., & Shute, V. J. (2015). Opportunities and challenges in assessing and supporting creativity in video games. In J. Green, & J. Kaufman (Eds.), *Research frontiers in creativity* (pp. 100–121). Academic Press.

Koestler, A. (1975). *The act of creation*. Picador.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice, 41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Litchfield, R. C. (2008). Brainstorming reconsidered: A goal-based view. *The Academy of Management Review, 33*(3), 649–668. https://doi.org/10.2307/20159429

Massetti, B. (1996). An empirical examination of the value of creativity support systems on idea generation. *MIS Quarterly, 20*(1), 83. https://doi.org/10.2307/249543

Moffat, D. C., Crombie, W., & Shabalina, O. (2017). Some video games can increase the player's creativity. *International Journal of Game-Based Learning, 7*(2), 35–46. https://doi.org/10.4018/IJGBL.2017040103

Niu, W., & Liu, D. (2009). Enhancing creativity: A comparison between effects of an indicative instruction "to be creative" and a more elaborate heuristic instruction on Chinese student creativity. *Psychology of Aesthetics, Creativity, and the Arts, 3*(2), 93–98. https://doi.org/10.1037/a0013660

Osborn, A. F. (1953). In *Applied imagination: Principles and procedures of creative thinking*. New York: Scribner.

Partnership for 21st Century Learning. (2019). Framework for 21st-century learning. http://www.nea.org/home/34888.htm.

Plucker, J. A., & Makel, M. C. (2010). Assessment of creativity. In J. C. Kaufman, & R. J. Sternberg (Eds.), *The cambridge handbook of creativity* (pp. 48–73). Cambridge University Press. https://doi.org/10.1017/CBO9780511763205.005.

Putman, V. L., & Paulus, P. B. (2009). Brainstorming, brainstorming rules and decision making. *The Journal of Creative Behavior, 43*(1), 29–40. https://doi.org/10.1002/j.2162-6057.2009.tb01304.x

Rahimi, S. (2020). *Inspire, Instruct, or both? Game-based and support of creativity*. Florida State University [Doctoral Dissertation].

Rahimi, S., & Shute, V. J. (2021). The Effects of video games on creativity: A systematic review. In S. W. Russ, J. D. Hoffmann, & J. C. Kaufman (Eds.), *Handbook of lifespan development of creativity*. Cambridge University Press.

Remy, C., MacDonald Vermeulen, L., Frich, J., Biskjaer, M. M., & Dalsgaard, P. (2020). Evaluating creativity support tools in HCI research. In *Proceedings of the 2020 ACM designing interactive systems conference* (pp. 457–476). https://doi.org/10.1145/3357236.3395474

Resnick, M., Myers, B., Nakakoji, K., Shneiderman, B., Pausch, R., Selker, T., & Eisenberg, M. (2005). Design principles for tools to support creative thinking. In B. Schneiderman, G. Fischer, M. Czerwinski, B. Myers, & M. Resnick (Eds.), *Creativity support tools: A workshop sponsored by the national science foundation* (pp. 25–35).

Richards, R. (2010). Everyday creativity. In J. C. Kaufman, & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (pp. 189–215). Cambridge University Press. https://doi.org/10.1017/CBO9780511763205.013.

Ristow, R. S. (1988). The teaching of thinking skills: Does it improve creativity? *Gifted Child Today Magazine, 11*(2), 44–46. https://doi.org/10.1177/107621758801100219

Runco, M. A. (2003). Education for creative potential. *Scandinavian Journal of Educational Research, 47*(3), 317–324. https://doi.org/10.1080/00313830308598

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal, 24*(1), 66–75. https://doi.org/10.1080/10400419.2012.652929

Sadler-Smith, E. (2015). Wallas' four-stage model of the creative process: More than meets the eye? *Creativity Research Journal, 27*(4), 342–352. https://doi.org/10.1080/10400419.2015.1087277

Sengewald, T., & Roth, A. (2020). User experience of creativity support tools – a Literature review in a management context. In H. Moreen, K. Poustcchi, & H. Krasnova (Eds.), *WI2020 Zentrale tracks* (pp. 1857–1869). GITO Verlag. https://doi.org/10.30844/wi_2020_r14-sengewald.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Company.

Shneiderman, B. (1999). User interfaces for creativity support tools. *Proceedings of the Third Conference on Creativity & Cognition-C&C, '99*, 15–22. https://doi.org/10.1145/317561.317565

Shneiderman, B. (2007). Creativity support tools: Accelerating discovery and innovation. *Communications of the ACM, 50*(12), 20–32. https://doi.org/10.1145/1323688.1323689

Shneiderman, B. (2009). Creativity support tools: A grand challenge for HCI researchers. In M. Redondo, C. Bravo, & M. Ortega (Eds.), *Engineering the user interface* (pp. 1–9). Springer London. https://doi.org/10.1007/978-1-84800-136-7_1.

Shute, V., Almond, R., & Rahimi, S. (2019). Physics Playground (1.3) [Computer software] https://pluto.coe.fsu.edu/ppteam/pp-links/.

Shute, V. J., & Rahimi, S. (2020). Stealth assessment of creativity in a physics video game. *Computers in Human Behavior, 116*, 1–13. https://doi.org/10.1016/j.chb.2020.106647

Shute, V. J., & Wang, L. (2016). Assessing and supporting hard-to-measure constructs in video games. In A. A. Rupp, & J. P. Leighton (Eds.), *The handbook of cognition and assessment* (pp. 535–562). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118956588.ch22.

Sio, U. N., Kotovsky, K., & Cagan, J. (2015). Fixation or inspiration? A meta-analytic review of the role of examples on design processes. *Design Studies, 39*, 70–99. https://doi.org/10.1016/j.destud.2015.04.004

Sternberg, R. (Ed.). (1988). *The nature of creativity: Contemporary psychological perspectives*. CUP Archive.

Sternberg, R. J. (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence, 34*(4), 321–350. https://doi.org/10.1016/j.intell.2006.01.002

Sternberg, R. J. (2012). The assessment of creativity: An investment-based approach. *Creativity Research Journal, 24*(1), 3–12. https://doi.org/10.1080/10400419.2012.652925

Sternberg, R. J., & Lubart, T. I. (1996). Investing in creativity. *American Psychologist, 51*(7), 677. https://doi.org/10.1037/0003-066X.51.7.677
Sun, G., & Yao, S. (2012). Investigating the relation between cognitive load and creativity in the conceptual design process. *Proceedings of the Human Factors and Ergonomics Society - Annual Meeting, 56*(1), 308–312. https://doi.org/10.1177/1071181312561072
Torrance, E. P. (1974). *Torrance tests of creative thinking: Norms-technical manual.* Scholastic Testing Service.
VanGundy, A. B. (1982). *Training your creative mind.* Prentice Hall.
Walia, C. (2019). A dynamic definition of creativity. *Creativity Research Journal, 31*(3), 237–247. https://doi.org/10.1080/10400419.2019.1641787
Ward, T. B. (2015). Content, collaboration, and creativity in virtual worlds. In G. P. Green, & J. C. Kaufman (Eds.), *Video Games and creativity* (pp. 119–136). Academic Press. https://doi.org/10.1016/B978-0-12-801462-2.00006-0.