International Journal of Child-Computer Interaction 33 (2022) 100503

Contents lists available at ScienceDirect

International Journal of Child-Computer Interaction

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

Redesigning learning games for different learning contexts: Applying a serious game design framework to redesign *Stop* & *Think*[☆]

Andrea Gauthier[a,*], Kaśka Porayska-Pomsta[a], Sveta Mayer[b], Iroise Dumontheil[c], Emily K. Farran[d], Derek Bell[e], Denis Mareschal[c], and the UnLocke Team

[a] UCL Knowledge Lab, Department of Culture, Communication and Media, IOE, UCL's Faculty of Education & Society, University College London, 23-29 Emerald Street, London, WC1N 3QS, UK

[b] IOE, UCL's Faculty of Education & Society, University College London, 20 Bedford Way, London WC1H 0AL, UK

[c] Department of Psychological Science, Birkbeck College University of London, Malet Street, Bloomsbury, London, WC1E 7HX, UK

[d] School of Psychology, University of Surrey, Elizabeth Fry Building, Stag Hill, Guildford, Surrey, GU2 7XH, UK

[e] Learnus®, Calder & Co, 30 Orange Street, London, WC2H 7HF, UK

A R T I C L E I N F O

Article history:
Received 3 August 2021
Received in revised form 18 March 2022
Accepted 24 May 2022
Available online 27 May 2022

Keywords:
Learning games
Design framework
Activity theory
Independent vs collaborative learning
Inhibitory control

A B S T R A C T

The Activity Theory-based Model of Serious Games (ATMSG) provides a visual framework through which designers and researchers can explicitly map the gaming, learning, and instructional design of their learning game mechanics and game flow. Here, we use the ATMSG to redesign an existing learning game, *Stop* & *Think* (S&T), which was created to train children to apply their inhibitory control skills when solving counterintuitive mathematics and science problems. S&T was previously found to be effective at increasing science and mathematics achievement when the activity was led by a teacher in the classroom. However, we sought to modify its design for use by children in an independent learning scenario (e.g., homeschooling). This work contributes to the literature by demonstrating how the ATMSG was used iteratively during the redesign of S&T for use in a child-led context. We found the ATMSG useful for (i) identifying design gaps created by removing the teacher from the gaming activity, thereby outlining areas of the game requiring modification, (ii) ideation to facilitate discussion about how different design ideas would impact the structure of the game and the feasibility of the approach, (iii) negotiating design decisions between team members, communicating proposed changes in the design amongst stakeholders, seeking approval from project leaders, and serving as a design document for developers, and (iv) cataloguing changes made to the game throughout the *re*design process, thereby archiving versions of the game which can be used to reflect upon how each version might impact counterintuitive reasoning. Yet, we also found some challenges in using the ATMSG, including its lack of ability to represent non-structural design decisions (e.g., visual strategies, adaptivity), its impractical format for representing more complex games, and its time-consuming nature.

1. Introduction

Learning games can capitalise on challenge, narrative, rewards, user autonomy, interactivity, and just-in-time feedback to foster mastery of specific knowledge and skills, to change behaviour, and/or to sustain learning motivation (Baptista & Oliveira, 2019; Byun & Joung, 2018; Clark et al., 2016; Garris et al., 2002). A number of game design and assessment frameworks have been developed over recent years in an attempt to understand how learning design principles can be embedded in game interactions and play contexts (see, e.g., Arnab et al., 2015; Carvalho et al., 2015; Kelle et al., 2011). For instance, the Activity Theory-based Model of Serious Games (ATMSG) outlines a concrete way to describe how the mechanics of gaming, learning, and instruction interrelate and to visualise the flow of the interactive learning experience (Carvalho et al., 2015). Such frameworks have been used in research dissemination and game evaluation contexts (e.g., Atmaja et al., 2020; Callaghan et al., 2016, 2018; Gauthier & Jenkinson, 2018), but it remains unclear how they are used during the game design process itself. We argue here that the ATMSG can be a valuable tool in the *re*design of learning games, for instance, to transform the game to fit different learning contexts. To the best of our knowledge, the use of the ATMSG framework within a game redesign and development pipeline has not yet been documented.

☆ This research was funded in part by the Wellcome Trust [105466/Z/14/Z] and the Education Endowment Foundation [N-1343].

* Corresponding author.
E-mail address: andrea.gauthier@ucl.ac.uk (A. Gauthier).

https://doi.org/10.1016/j.ijcci.2022.100503

As part of the Education Endowment Foundation- and Wellcome Trust-funded UnLocke educational neuroscience project, we developed a learning game, called *Stop & Think*, aimed at improving children's use of inhibitory control (specifically by stopping and thinking) when solving counterintuitive science and mathematics problems. Children played the game as a whole class for 15 min, three times per week, for 10 weeks, with the activity facilitated by the teacher. Two evaluation studies have demonstrated the effectiveness of the teacher-led game at improving counterintuitive reasoning and academic achievement in this context (Roy et al., 2019; Wilkinson et al., 2019). However, the onset of the Covid-19 pandemic emphasised the importance and potential benefits of adaptive, independent learning technologies that can support children outside the classroom, without the presence of a teacher or parent to facilitate the learning activity (Fitzgerald et al., 2018; Meyer et al., 2008; Proulx et al., 2017). We undertook a three-phase redesign process to transform *Stop & Think* into a child-led, independent learning experience, appropriate for homeschooling, utilising the ATMSG during each phase.

Our aim is to demonstrate how the ATMSG framework was iteratively applied to facilitate the redesign of the teacher-led *Stop & Think* game to fit a child-led learning context. Our contribution is the exemplification of the advantages and challenges of using the ATMSG framework in this process through a concrete redesign case study. In doing so, we highlight its role in considering modifications to the *intrinsic* instructional elements of the game (i.e., instruction integrated within the game itself, like feedback, scaffolding, and adaptivity) to compensate for the removal of *extrinsic* instruction (i.e., instruction provided by the teacher, parent, peer, or facilitator, like verbal feedback and personalised guidance and support) (Carvalho et al., 2015; Johnson et al., 2017; Jonassen & Rohrer-Murphy, 1999). We also propose an extension of the ATMSG visual map, discuss how the ATMSG may be valuable in both user-centred and participatory design approaches, and recommend the framework's further exploration in future learning game research, (re)design, and development.

2. Background

2.1. Game-based learning contexts

There are documented benefits to whole-class teaching, small-group collaboration, and independent learning with technology. Whole-class learning activities and collaboration in small groups can foster diffusion of knowledge between children and helps children develop a sense of community and shared understanding (Elbers & Streefland, 2000; Wood & O'Malley, 1996). During whole-class activities, teachers take charge of the learning, can respond to children's questions in real-time, and support their engagement (Elbers & Streefland, 2000), effectively delivering *extrinsic instruction* to support children's learning. Children can also provide extrinsic instruction to their peers; children learning in small groups with computer-supported collaborative learning (CSCL) technologies have been found to learn science and mathematics more effectively than children learning with technology independently (Gallardo-Virgen & de Villar, 2011; Gijlers et al., 2013; Jackson et al., 2013; Tsuei, 2012). However, due to the lack of reporting in the designs of these technologies, it is difficult to ascertain whether these tools were designed appropriately to support independent learning, without extrinsic support from peers or teachers, or whether they provided adequate *intrinsic instruction*, e.g., by adapting to support the unique needs of individual learners. Furthermore, all these interventions took place in the classroom, where collaboration happened face-to-face. There is some suggestion from older work (e.g., Van Der Meijden & Veenman, 2005) that CSCL technologies for young children may only be effective in face-to-face contexts because children's communication skills are not sufficiently developed to collaborate effectively when communication is mediated by the computer. A recent study conducted during the 2020 COVID-19 lockdown showed that young children in the UK struggled with remote peer contact during homeschooling, but that computer-mediated peer communication increased with age (Thorell et al., 2020); this could suggest that distance CSCL may be more appropriate for older children and teens than for young children in a homeschooling scenario.

In contrast, independent learning technologies leverage *intrinsic instruction* strategies over *extrinsic instruction*. Well-designed independent learning technologies can foster increased learner motivation and offer personalised, real-time feedback and scaffolding that adapts to the changing knowledge and abilities of learners (Fitzgerald et al., 2018; Meyer et al., 2008). Importantly, personalised learning technologies—especially games—are thought to give learners a sense of control over their learning process (Fitzgerald et al., 2018; Proulx et al., 2017). These potential benefits may become particularly important when learning is not supported by extrinsic instruction (by peers, parents, or teachers), such as during homeschooling; Thorell et al. (2020) demonstrated that, during the first COVID-19 lockdown, children in the UK spent on average 4.4% of their homeschooling time in contact with a teacher. Whilst parental engagement in homeschooling was more substantial (45.2%), this was associated with increased parental stress, especially amongst those with younger children (Thorell et al., 2020). In such contexts, learning technologies that rely on *intrinsic* instruction and *do not* require *extrinsic* support from teachers, parents, or peers, may prove to be effective because they could (i) provide personalised, adaptive support to each child, (ii) allow the limited student–teacher contact time to be spent on pedagogical matters of higher priority, and (iii) alleviate parental stress in needing to facilitate the learning activity.

2.2. Learning game design frameworks

Historically, the game-based learning literature has failed to measure causal links between the designs of learning games and the significant learning outcomes that their use generates (Boyle et al., 2016; Clark et al., 2016). As such, there has been a push toward a more design-based paradigm of investigation, which requires researchers to provide robust descriptions of how the designs of their game-based interventions are intended to promote learning and engagement. Several game design frameworks attempt to tackle this challenge, including the Game Object Model (Amory, 2007), the Serious Game Lemniscate Model (Koops & Hoevenaar, 2012), and the Cognitive Behavioural Game Design Model (Starks, 2014). While all these frameworks may demonstrate unique benefits, they do not allow authors to make explicit associations between gaming, learning and instructional components of games, which may be important in planning which configuration of mechanics may be most likely to drive the desired learning outcomes.

Other frameworks have been more successful at elucidating concrete associations between learning and gaming mechanics (e.g., Arnab et al., 2015; Kelle et al., 2011; Proulx et al., 2017), but do not tackle the instructional component. Kelle et al. (2011) provide mappings of how various learning functions (e.g., related to preparation, knowledge manipulation, higher order relationships, learner regulation, and productive actions) might be associated with diverse game design patterns (e.g., goal-related, information-related, score-related, mastery-related). However, this framework lacks an understanding of overall game structure/flow. This is remedied in part by the Learning Mechanics-Game Mechanics (LM-GM) framework which provides a flowchart structure on which to map pairs of learning mechanics and game mechanics, to represent the flow of the game (Arnab et al., 2015). The framework provides a non-exhaustive list of learning and gaming mechanics from the literature and allows authors to pick and choose which mechanics get paired together to accurately describe the function/purpose of interactions within the game-flow chart. The authors further discuss how certain gaming mechanics induce motivational behaviour in relation to self-determination theory, depending on their ability to promote players' sense of autonomy, connection or relatedness, and competence (Proulx et al., 2017). For instance, they describe mechanics with selection/action task components as highly intrinsically motivating, whereas progression mechanics may either have high or low motivational value depending on how directed or linear the progression is. Yet, the LM-GM still lacks concreteness in how mechanics connect to the game's high-level educational objectives (Carvalho et al., 2015), and does consider the role of intrinsic and extrinsic instructional mechanics within learning games, which we suggest is critical when considering a redesign from a teacher-led to a child-led context.

2.3. Activity theory-based model of serious games

As suggested by its name, the Activity Theory-based Model of Serious Games (ATMSG) builds on Activity Theory, which suggests that all human activities (or interactions) comprise of *actions* that are enabled by *tools,* to achieve specific *goals*. The ATMSG postulates that, within a learning game, these activities may either be (i) game-oriented, (ii) learning-oriented, or (iii) instructionally oriented (or any combination of the three) (Carvalho et al., 2015). Importantly, the framework makes a distinction between intrinsic instruction (e.g., feedback/scaffolding given *within the game*) and extrinsic instruction (e.g., feedback/scaffolding provided *by a teacher, parent, peer, or facilitator*), which has educational relevance for learning games designed as classroom-based vs. independent activities. The framework also provides a visual mapping structure that capitalises on the Unified Modelling Language (UML, a well-known visual vocabulary to communicate sequential processes (Gomaa, 2006)) in order to enable authors to concretely describe the flow of their game. This enables mapping of gaming, learning, and instructional activities to one another in accordance with the interaction flow of the game (Fig. 1). Like the LM-GM model, the ATMSG provides a non-exhaustive taxonomy of gaming, learning, and instructional activities, but expands on this by offering potential actions, tools, and goals specified for each, with which to populate the table under the flow visualisation. ATMSG visual maps can be generated using free web-based UML visualisation software, e.g., Lucidchart (Lucid, 2022) or Cacoo (Nulab, 2022), or—for those who are more proficiently creative—professional design software, e.g., Adobe Illustrator (Adobe, 2022) or Affinity Designer (Serif Europe LTD, 2020).

The ATMSG has been applied successfully in recent years to describe the designs of several learning games for research dissemination and game evaluation purposes (see Callaghan et al., 2018; Garneli et al., 2021; Gauthier & Jenkinson, 2018). For example, Gauthier and Jenkinson (2018) used the framework to compare and contrast design differences in a pair of simulation-based interventions for molecular biology students, wherein one was considered a serious game, and one was considered a non-game intervention. In a similar way, we suggest the ATMSG is a versatile framework that might also be used to redesign a learning game to render its application in different pedagogical contexts. Specifically, we used it to help redesign *Stop & Think*, a game to train children to apply inhibitory control when solving counterintuitive science and mathematics problems, from a teacher-led context to child-led, independent learning context.

2.4. Inhibitory control and counterintuitive reasoning

Here we present some literature on inhibitory control (IC), which is required to understand the reasoning behind the design of both the teacher-led game and its redesign for a child-led context. IC is one of many executive functions known to impact learning (Coulanges et al., 2021; Mason & Zaccoletti, 2021). IC is a person's ability to inhibit or suppress prepotent behaviours and intuitive knowledge and is believed to be of particular importance to learning counterintuitive concepts (i.e., concepts that contradict our naïve and immature beliefs or theories). This is because our brain works in two distinct but parallel systems (Evans, 2003): (1) a heuristic system, which enables fast, intuitive, and reflexive decision-making, and (2) a slower analytic system, which facilitates concentrated thought and logical reasoning. Because of its fast-processing speed, the heuristic system is typically called upon first in familiar problem-solving situations, allowing us to react quickly and intuitively by drawing on embedded prior knowledge and experiences. However, relying on the heuristic system during problem-solving becomes an issue when such problems involve counterintuitive concepts or common misconceptions. In such situations, the heuristic system might call upon *seemingly* similar intuitive knowledge/experiences to find a quick solution, leading to an incorrect result. In these cases, learners need to suppress (or inhibit) the heuristic system using IC, to allow the analytic system to take over in problem solving (Diamond & Lee, 2011; Mareschal, 2016).

Three types of IC (Nigg, 2000) are particularly important when tackling counterintuitive concepts in mathematics and science:

(1) *Cognitive inhibition* involves putting aside a previous belief or learnt fact, rule, or procedure, which does not apply to the current problem. For example, in mathematics children learn positive integer numbers in sequence, i.e., 1, 2, 3, 4, 5 etc. Later, children are taught about negative numbers. Children now need to inhibit their prior knowledge that 5 is greater than 1 in order to correctly identify -1 as greater than -5 (Bofferding, 2019).
(2) *Perceptual interference control* may also be needed to ignore distracting visual or other perceptual cues that are irrelevant to the current problem. For example, children become familiar with the appearance and habitat of fish (i.e., streamlined body with fins, swims in the sea). When presented with a dolphin, the child needs to inhibit these visual cues and instead think about the properties of fish and mammals (e.g., gills vs. lungs) to recognise that a dolphin is a mammal and not a fish (Allen, 2014).
(3) Finally, *response inhibition* allows children to suppress a motor response. In the case of classroom learning, this motor response might be shouting out a first idea or immediately raising a hand to provide an answer. In game-based learning, this might be trying to interact with objects on screen. So, response inhibition allows children to withhold their immediate response or behaviour, i.e., to *stop*, so that they have more time to *think*.

2.5. About stop & think

Stop & Think (henceforth, S&T) is a learning game designed to train children (aged 7–10) to apply IC skills—i.e., to "stop and think"—when answering age-appropriate science and mathematics problems. In doing so, they suppress their heuristic system and engage their analytic system, thereby inhibiting their intuitive

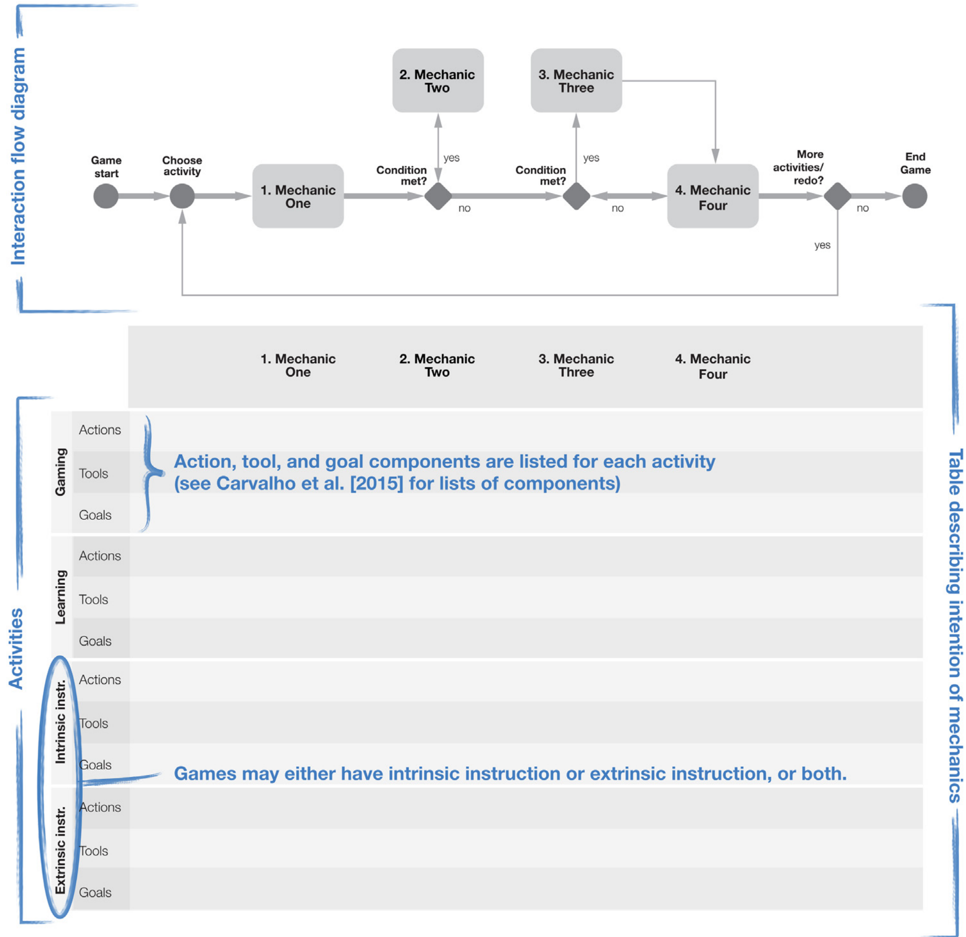


Fig. 1. Generic structure of an Activity Theory-based Model of Serious Game (ATMSG) visual map (based on Carvalho et al., 2015). Rounded rectangles represent mechanics that are then elaborated upon in the table beneath. Diamonds represent conditions that affect the interaction flow in some way. Circles represent game states.

responses and giving them time to better consider the problem at hand. Full curriculum mappings, as well as open-source software files, user manuals, videos, and more, are available on the UnLocke project's Open Science Foundation (OSF) site: https://osf.io/6er4k/?view_only=895f0d877d144241bc642ec771f4db10.

S&T is set-up like a gameshow, in which a virtual host, Andy, poses questions to three virtual non-player-character gameshow contestants and, thereby, also children in class. There are two phases to the game show: (1) the Exploratory problem, where Andy and the gameshow contestants offer different levels of support depending on the class's response, and (2) the Structured Practice problems in the "Bonus Round". Each session runs for 12 min and is split equally between mathematics and science content, presented in random order. Whilst the number of problems completed in any session depends on how quickly the class works through them, the emphasis is on practicing IC during the session, rather than trying to cover all the content. The programme is designed to run three times per week for 10 weeks, with the activity led by the teacher at the front of the classroom. Previous evaluations of the game have demonstrated that the teacher-led activity leads to improved counterintuitive reasoning and academic achievement in science and mathematics (Roy et al., 2019; Wilkinson et al., 2019).

3. Redesigning *stop & think* for a child-led context

As discussed in Section 2.1, the onset of the Covid-19 pandemic has emphasised the importance and potential benefits of adaptive, independent learning technologies that can support children *outside* the classroom, *without* support from a teacher or parent (Fitzgerald et al., 2018; Meyer et al., 2008; Proulx et al., 2017). This prompted us to consider a child-led, adaptive version of S&T to support independent IC training. Technologies that adapt to the needs of individual learners can promote sustained engagement/motivation in learning by scaffolding the activity in an individualised way (Fitzgerald et al., 2018), which would support the use of such technologies in, e.g., a homeschooling environment. However, when removing the teacher (i.e., the extrinsic instruction) from the gameplay scenario, it is worthwhile to consider the impact of this on children's ability to engage meaningfully in the S&T activity, and what design changes should be made to better support them.

Therefore, in the sections that follow, we use the ATMSG to redesign the S&T game to be used as a child-led, adaptive, and independent learning intervention. We made the decision from the outset that this redesign would rely only on *intrinsic* instruction (where learning is supported by design of the game), rather than *extrinsic* instruction (where learning is supported by a teacher, peer, or parent), for the reasons outlined in Section 2.1 and informed by Thorell et al. (2020).

To do this, we approached the redesign through three-phases that leveraged the ATMSG framework to support the redesign process:

- **Phase I**: We applied the ATMSG to the teacher-led version of the game and examined it for insights on gaps that would be created by removing the extrinsic instruction (Section 3.1);

- **Phase II**: We generated new ideas through an interdisciplinary design workshop, using the ATMSG to constrain the ideation process (Section 3.2)
- **Phase III**: We tested the new design ideas through a randomised user study before recommending further modifications to the game through the ATMSG (Section 3.3)

3.1. Phase I: Gaining insights from a teacher-led ATMSG

The teacher-led S&T game was not originally designed using the ATMSG framework. Below, we describe how the ATMSG was used to map out the design of the teacher-led game (Section 3.1.1) and then how this visual map was analysed to discover gaps created by removal of extrinsic instruction (Section 3.1.2).

3.1.1. Applying the ATMSG to visualise IC training in the teacher-led game

The redesign effort was spearheaded by the first author of this paper, who is an interaction designer and researcher brought onto the project specifically to help redesign the teacher-led game for a child-led context. She started by mapping out the teacher-led game using the ATMSG. This process involved examining any existing design documents (e.g., teacher manuals) and by repeatedly playing through the game to develop a concrete understanding of its structure. This helped her understand how S&T integrated IC training, scaffolded supports to help the children grasp counterintuitive concepts, and—importantly—leveraged the teacher to support the training activity. Throughout this preliminary research, the first author sketched out the flow and mechanics of the game (using conventions provided by the ATMSG) in a sketchbook before digitising the diagram in Affinity Designer (Serif Europe LTD, 2020). The final result of this process is Fig. 2, which illustrates the design of the teacher-led game using the ATMSG map as a visual aid. The analytical and creative process required to (a) understand how S&T promoted IC and then (b) map out the design logically was intensive and took several hours over two days.

Just as it was important for the designer to come to grips with the structure of the teacher-led game before considering its redesign, we feel that it is important for our readers to understand it, too, so we will use Fig. 2 to briefly walk through the flow of the teacher-led game. When referring to specific 'M'echanics in the map, we will use the convention "M", followed by the number of the mechanic as labelled in Fig. 2 (e.g., M1, M2).

Upon launching the game, the class first encounters the Exploratory problem, which allows multiple attempts to correctly complete the activity, with progressively greater levels of support offered each time an incorrect response is given. Fig. 3 shows screenshots of several mechanics from an exemplar Exploratory problem on tens and units (Year 3 Mathematics). For each problem, Andy presents a question (M1) and reminds children to "stop and think" about their answer, and the screen is briefly locked on a pulsing "stop and think" icon (M2); this is where the children exercise their IC skills, and the teacher is expected to demonstrate this behaviour. When the icon stops pulsing, a user can now interact with the screen in their first attempt at answering the question (M3; refer to OSF site for details on different types of interactions) – the teacher will determine how children will participate, e.g., by group vote, by choosing a child coming to the interactive whiteboard, etc. After completing the problem, the user then needs to click on a green 'check answer' button at the bottom of the screen (depicted in Fig. 3-M3).

Following a *correct* response on any Exploratory problem, children are given feedback by Andy (M4; "That's the correct answer! Let's see what the other contestants thought".) and then are shown the three contestants giving their own thoughts about the question (M5; follow the flow going *downward* in Fig. 2, after M4). An example of contestants' responses is visible in Fig. 3-M5. The class is then asked to select which contestant had the correct reasoning (M6). This is designed to consolidate the reasoning behind the correct answer, encouraging children to think things through rather than just going with their intuitive response. After selection, the correct reasoning is shown alone on the screen (M7), followed by another presentation of the correct problem solution by Andy (M9). Following this, the gameshow moves on to the Structured Practice problems (the "Bonus Round"), if there is time remaining in the 6 min allotted to the current subject.

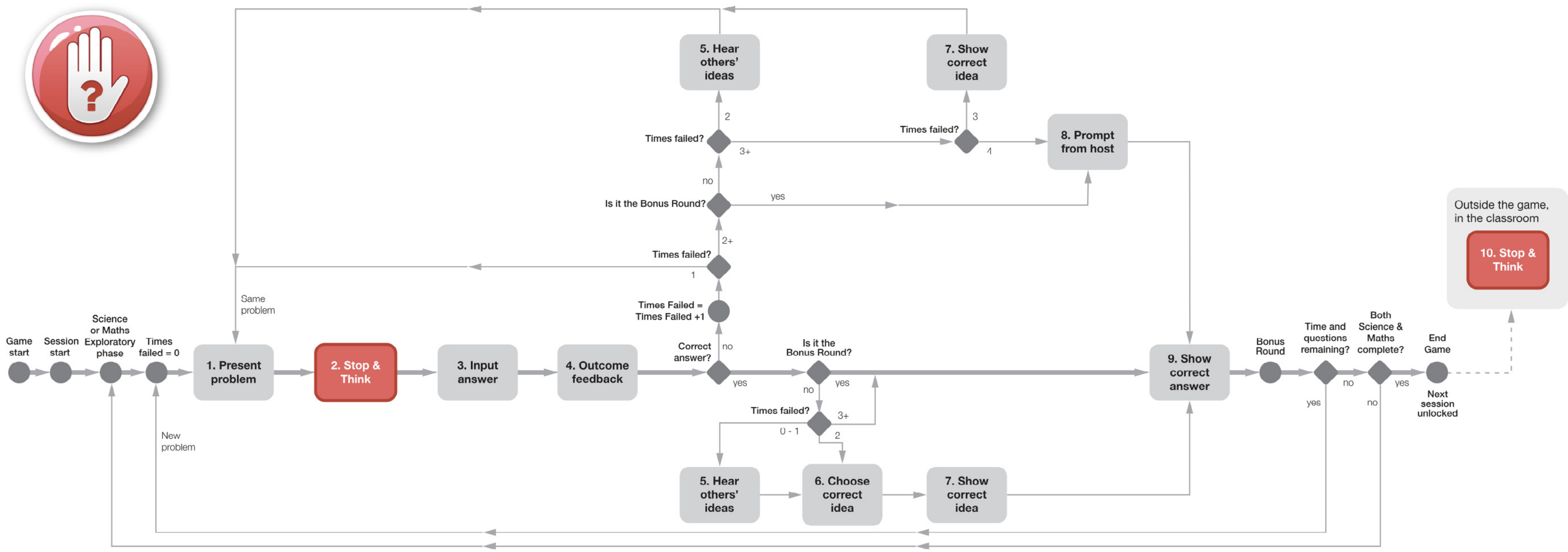
However, if the first attempt at the Exploratory problem is *incorrect*, Andy says (M4), "That's not quite right. Have another go" and sends the children back to M1 (follow the flow going *upward* in Fig. 2, after M4). If still incorrect upon a second try, children are provided with the following levels of scaffolding to support their counterintuitive reasoning:

(1) Children are shown the virtual contestants giving their thoughts about the question (M5). One contestant will have the correct reasoning, while the other two will have incorrect reasoning. This is designed to encourage children to *think* about the concept rather than just go with their impulsive response or keep guessing. The class then has another attempt at the initial question (M1).
(2) If incorrect a third time, children are shown which contestant had the correct reasoning (M7). The class now gets a final attempt at the initial question (M1).
(3) If incorrect a fourth time, Andy gives a verbal prompt, e.g., "Think about…" or "Remember what you have learnt about…" (M8). The correct answer is then revealed (M9), and the child is moved on to the Bonus Round (time-permitting). Whilst teachers are encouraged to let children make errors, at this point they might stimulate their recall of prior knowledge to help in error recovery going forward.

Once the correct response is entered, or four incorrect attempts are made, the gameshow moves on to a "Bonus Round" comprising of Structured Practice problems. Structured Practice problems provide children with the opportunity to practice more questions on the same topic, to consolidate their understanding of the mathematics or science concept that was introduced in the Exploratory phase and give children more opportunities to practice applying their IC skills. Structured Practice problems look very similar to the Exploratory phase, but do not progress through these same levels of support (M5-7). Instead, children are given two attempts with the initial "stop and think" prompt (M2) and then a prompt/hint from Andy upon the second incorrect attempt (M8), which they can apply to the next problem. Teachers are encouraged to make links between stopping and thinking in the game and applying IC in science and mathematics learning in the classroom (M10).

3.1.2. Analysing the teacher-led ATMSG to identify design gaps

After creating the teacher-led ATMSG visual map, the first author then analysed it to identify new design gaps that would be created due to the transition to a child-led context. A "gap" was identified as a possible detrimental design flaw (i.e., that would reduce the effectiveness of the game at training IC behaviours), resulting from the removal of the teacher's role in facilitating the S&T intervention. This was achieved first by scrutinising the *extrinsic instruction* rows of the ATMSG mechanics table in Fig. 2, noting the actions, goals, and tools involved in the extrinsic instruction activity, and then by thematising the gaps created by the removal of that activity. The first author then discussed and consolidated her findings with another design team member in advance of the interdisciplinary workshop in Phase II.



		1. Present problem	2. Stop & Think	3. Input answer	4. Outcome feedback	5. Hear others' ideas	6. Choose correct idea	7. Show correct idea	8. Prompt from host	9. Show correct answer	10. Stop & Think
Gaming	Actions	x	x	Answer questions/trivia	See performance evaluation	Obtain help, listen to/read information	Select	x	Obtain help	x	x
	Tools	x	x	Graphic puzzle or text entry	Acknowledgement correct/incorrect	Dialogue from non-player characters	Dialogue from non-player characters	x	Advice from guide character	x	x
	Goals	x	x	Solve puzzle/trivia question	Solve puzzle/trivia question	Solve puzzle/trivia question	Solve puzzle/trivia question	x	Solve puzzle/trivia question	x	x
Learning	Actions	x	Recognise, compare, distinguish	Choose, classify, complete, categorise	x	Find more information	Recognise, distinguish, choose	Recognise, relate	Recognise, relate	Memorise	x
	Tools	x	Delay interaction (wait)	Graphics, text input	x	Written info/arguments	Written info/arguments	Written info/arguments	Written info/arguments	Graphics/text	x
	Goals	x	Reflective observation	Application, integration, active experimentation	x	Internalise values/other perspectives	Integration	Integration, adaptation	Understand	Remember, internalise values	x
Intrinsic instr.	Actions	Present problem	Stress importance of stopping & thinking	x	Assess performance	Support recovery from errors, scaffold	Reinforce idea	Demonstrate / reinforce idea, scaffold	Support recovery from errors	Demonstrate	x
	Tools	Challenge	Stop & Think icon	x	Performance measure (correct/incorrect)	Tips and assistance, limited set of choices	Limited set of choices	Display correct idea	Tips and assistance	Display correct answer	x
	Goals	Elicit performance, stimulate recall of prior learning	Provide learning guidance	x	Provide feedback	Provide learning guidance	Enhance retention and transfer	Provide learning guidance	Provide learning guidance	Provide learning guidance	x
Extrinsic instr.	Actions	x	Stress importance, demonstrate	Present material, scaffold	x	x	x	x	x	Support recovery from errors	Stress importance of stopping & thinking
	Tools	x	Demonstration	Discussion, challenge	x	x	x	x	x	Tips and assistance	Discussion, demonstration
	Goals	x	Provide learning guidance	Stimulate recall of prior learning, elicit performance	x	x	x	x	x	Provide learning guidance	Provide learning guidance

Fig. 2. ATMSG visual map for the teacher-led version of *Stop & Think*.

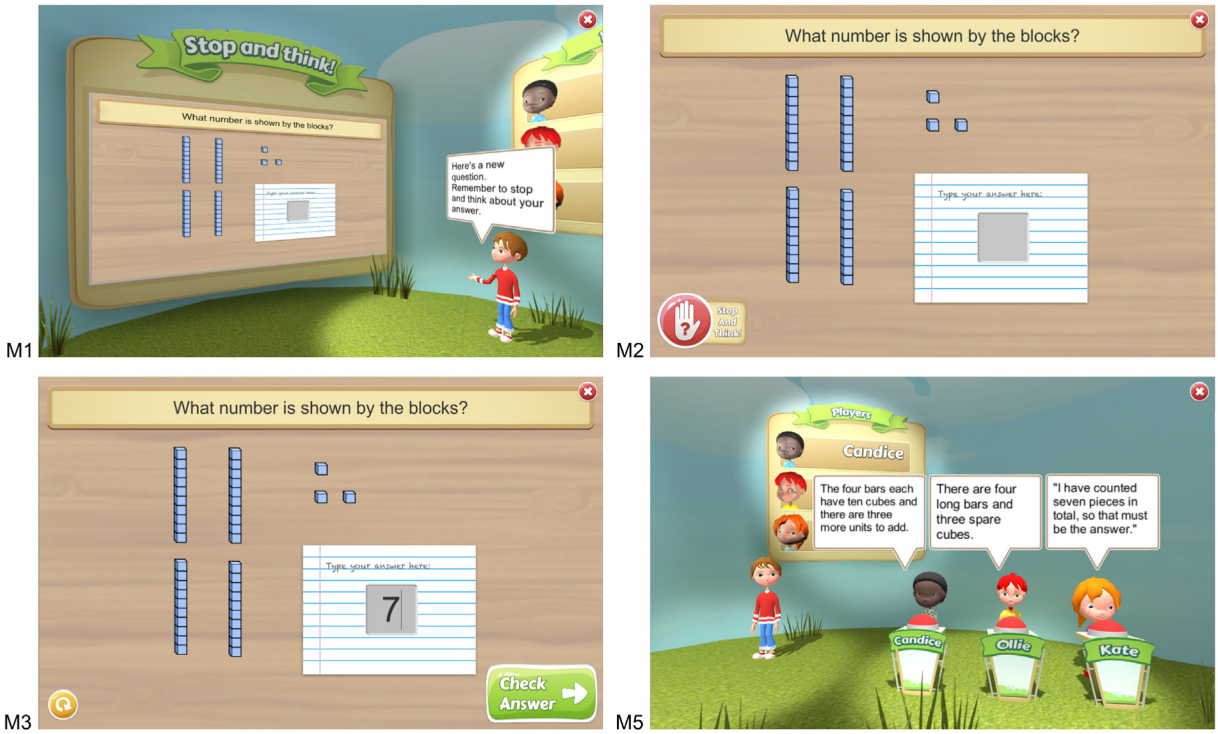


Fig. 3. Example of Year-3 mathematics problem. M1 – Andy presents a problem; M2 – S&T icon pulsates in the corner; M3 – the class inputs their answer (common misconception displayed); M5 – the class hears NPC contestants' ideas, one of whom has the correct idea (Candice), another displays the common misconception (Kate), and the third is just incorrect (Ollie), in this example. M-numbered mechanics relate back to the ATMSG visual map in Fig. 2.

		While playing the game in class			Outside of the game
		M2. Stop & Think	M3. Input answer	M9. Show correct answer	M10. Stop & Think
Extrinsic instr.	Actions	Stress importance, demonstrate	Present material, scaffold	Support recovery from errors	Stress importance of stopping & thinking
	Tools	Demonstration	Discussion, challenge	Tips and assistance	Discussion, demonstration
	Goals	Provide learning guidance	Stimulate recall of prior learning, elicit performance	Provide learning guidance	Provide learning guidance

Fig. 4. The actions, tools, and goals of the extrinsic instructional activity (i.e., the role of the teacher) in the teacher-led version of *Stop & Think*.

Fig. 4 summarises the *extrinsic instructional activities* of the teacher in the S&T intervention, previously presented in Fig. 2, which are broken down into actions, tools, and goals. These revealed four thematic design gaps that would be created by removing these components of the game mechanics. These were: **Gap 1**, stopping-and-thinking behaviour will no longer be reinforced (e.g., through demonstration) during and beyond the game (M2, M10); **Gap 2**, stimulation of prior knowledge regarding counterintuitive concepts during task completion and during error recovery will no longer be supported extrinsically (M3, M9); **Gap 3**, children may be less motivated and engaged in the learning activity without the scaffolding, discussion, and general encouragement delivered by the teacher (M3); and **Gap 4**, children's frustration, confusion or inappropriate (i.e., off-track) behaviours may go unchecked (throughout the game).

By identifying and documenting these gaps, the design team could focus the workshop ideation in Phase II on developing new, context-specific mechanics or modification of existing mechanics to fill these design gaps.

3.2. Phase II: Ideating design changes in an interdisciplinary workshop

We conducted an interdisciplinary design workshop to ideate on ways in which the design of our teacher-led game could be altered to promote stop-and-think behaviours in a child-led context by addressing gaps created by the removal of extrinsic instruction, as identified in Section 3.1. Due to limitations in budget, timing, and the strict remit of the project, the goal was to stick with the overall structure of the game but to redesign aspects of it to support children's IC training in the absence of a teacher. As such, a user-centred design approach was selected, using teachers and domain experts as informants and children as evaluators of the various outputs (e.g., through a user study, Section 3.3). Whilst participatory design approaches are generally preferred in the development of child- and learner-centred interventions (Khaled & Vasalou, 2014; Nunes et al., 2016; Soloway et al., 1994; Vasalou et al., 2021), the strict outcome-oriented nature of this redesign effort necessitated a less exploratory approach (Parsons & Cobb, 2014)—the implications of which are discussed further in Section 4.3. Overall, the workshop had four activities, most of which centred around the teacher-led ATMSG to facilitate ideation, and which was followed by the redevelopment of the ATMSG to take into account our ideated changes after the workshop.

3.2.1. Participants

The workshop involved nine adults (plus the first author who acted as the workshop facilitator), including developmental neuroscientists, artificial intelligence and human–computer interaction researchers, game designers, and educators. Six of these participants were project team members who had in-depth knowledge of S&T, covered the relevant interdisciplinary expertise, and two who were themselves teachers. One of these project members was involved on the development/programming side of the software. The other participants were game designers and primary education experts, who were not previously involved in the project and were invited to bring a fresh perspective to the workshop.

3.2.2. Activity 1: Presentation of design gaps

Activity 1 was a presentation from the facilitator to introduce new participants to the background of the project and bring everyone to the same understanding of about the design gaps and redesign objectives. This included (i) an explanation of the importance of supporting IC behaviours in children's counterintuitive reasoning, (ii) an inexhaustive summary of general strategies for guiding independent learning behaviours in games (e.g., visual interface features, adaptivity of the software, activity scaffolding), (iii) a demonstration of the teacher-led game, (iv) a presentation and explanation of the teacher-led ATMSG, highlighting the four design gaps identified upon removing the extrinsic instruction, (v) an explanation of the limitations of the project in terms of feasibility, and (vi) the objectives of the workshop (i.e., to generate design ideas to fill those gaps within the given limitations). The outcome of this activity was a shared understanding of our goals and the knowledge necessary to achieve them.

3.2.3. Activity 2: Facilitated ideation using the ATMSG

Activity 2 was a whole-group facilitated ideation session where participants generated ideas on sticky-notes about how game/interactivity design, visual interface features, adaptivity of the software, and activity scaffolding might support children's IC training, once the extrinsic instruction of the teacher was removed. The existing ATMSG was printed on a large sheet of paper and displayed at the front of the workshop room so that (1) participants could indicate how their ideas fit into or would necessitate a change to the current structure of the game, and (2) the S&T software designers/developers could explain how the feasibility of some ideas would be limited, based on how extensively the structure would need to be modified to accommodate new ideas.

This activity resulted in several ideas to fill gaps generated by removal of extrinsic instruction in S&T. Participants wrote or drew their ideas on sticky notes and displayed them around the printed ATMSG visual map, giving a verbal explanation about where their idea appeared in the flow of the game or altered the flow of the game, in addition to how they believed the idea supported children's independent use of IC. In a few instances, the developer fed back that an idea may not be feasible given the time and resources, and these were marked to keep for future consideration (see below). After initial ideation on stickie notes, the more feasible ideas were thematically grouped on the whiteboard in relation to game/interactivity design, visual interface features, adaptivity of the software, and activity scaffolding and displayed on the whiteboard (Fig. 5a–b). Idea themes included: **(#1) awarding point, coins, or tokens** for accuracy and time spent stopping-and-thinking to motivate engagement (addresses Gap 3); **(#2) integrating an "I'm ready to answer" button** that children can press to declare that they have finished stopping-and-thinking, similar to how they might raise their hand to provide an answer in the teacher-led condition, thereby scaffolding the IC activity (addresses Gap 1); **(#3) using symbolic colour** (e.g., traffic light symbols), rather than pulsating motion of the S&T icon, to investigate if this approach might be better at guiding the IC activity (addresses Gap 1); **(#4) individualised support**, through adapting the enforced "thinking" time (duration of pulsating icon) to each child's behaviour in the game and gradually remove any visual cues (e.g., pulsating icon, "I'm ready" button) as the child consistently displays the IC behaviour during play (addresses Gap 1); **(#5) scaffolding the activity with a walk-through** on how to stop-and-think in the context of science and maths problems (addresses Gaps 2 and 4); and **(#6) integrate a game progression mechanic**, allowing the child to unlock new levels as they progress and to retry previously completed content (addresses Gaps 2 and 3); this is intended to help children develop a sense of progression and achievement in the game, reinforce science/mathematics concepts and IC behaviours, and increase learning motivation through goal setting (e.g., improve a previous score) (Proulx et al., 2017).

We believe that, because of the use of the ATMSG to visualise the structure of the game and constrain people's thinking, most ideas generated were very feasible, but a few were not. For example, one idea that was abandoned involved implementing two gameplay modes, e.g., "Home Mode", which would not limit their play time, and "School-Use", which would limit their play to 12 min to accommodate classroom learning restrictions (addresses Gap 3), similar to the teacher-led S&T. The programmers argued that the extra effort required to create and integrate two separate game-flow structures would not be feasible given the funding limitations—but this idea could be considered in the future, as it would create continuity between home-based and school-based learning. Other ideas were also abandoned based on their relevance to addressing the gaps; these are not pertinent to our reflection on the use of the ATMSG, so will not be reported further here.

3.2.4. Activity 3: Breakout group idea development

Activity 3 was a breakout-group ideation activity, where interdisciplinary sub-groups took ideas from the sticky notes and developed them into more detailed mock-ups using printed templates. Prior to going into breakout groups, the workshop participants all discussed and decided on ideas that were the most feasible to develop and that were most likely to improve the child-led learning experience in S&T. Of the six core ideas presented in Section 3.2.3, ideas #1, #2, #3, and #5 were selected for elaboration and mocking-up within breakout groups.

Participants divided into three multidisciplinary breakout groups, each taking one of the selected ideas to work on. Participants all agreed that idea #2 about the "readiness indication" mechanic (e.g., through an "I'm ready to answer" button), was a feature that they wanted to be included in the redesign and should be tested for efficacy in future design experiments. This was based on the theory of planned behaviour, which postulates that goals (in this case, engaging in stopping-and-thinking) are achieved through a series of more-or-less well-planned actions (Ajzen, 1985, 1991); participants in the workshop made a connection between this idea of planned behaviour during playing S&T in the classroom, through hand-raising before answering, as a mechanism to achieve the goal of stopping-and-thinking (i.e., engaging in IC). As such, it was felt that the "readiness" mechanic should be tested for its efficacy in supporting this type

Fig. 5. Examples of design workshop outcomes: (a) whole-group whiteboard ideation, with sticky notes made around ATMSG visual map, (b) detail of a, showing thematic grouping of initial ideas, (c) and (d) breakout group ideation mock-up describing an idea about the use of symbolic colour to guide the IC activity.

of behaviour planning in S&T, and all breakout groups would integrate this mechanic into their mock-ups where relevant.

For idea #1 about rewards, workshop participants proposed that children would be rewarded tokens based on the accuracy of each answer (directly after M9 in Fig. 2), which was suggested could increase children's extrinsic motivation to play (Proulx et al., 2017). During the Bonus Round, children would earn a "bonus multiplier" for sequential correct answers, which would be removed upon an incorrect answer, to instil a sense of consequences (Juul, 2009; Proulx et al., 2017). It was thought that this would encourage children to slow down and really think about their next move carefully, so that they did not lose their rewards. Additionally, it was proposed that, over several sessions when children displayed consistent IC-behaviours, they could earn trophies to show their progression in their IC skill.

Idea #3 explored the use of symbolic colour in the S&T icon to guide IC behaviour, instead of pulsating motion, and which would integrated directly with idea #2 on the readiness indication mechanic (Fig. 5c–d). The breakout group argued that children develop cultural understandings of colour choices from a very young age (Burkitt et al., 2003), and specifically, that the traffic light metaphor is often used in primary-level education and interventions (Bull et al., 2005; Girard, 2011; Jung et al., 2019). They proposed that the S&T icon should glow (a) RED to indicate to the child to *stop* and listen to/read the problem carefully, (b) AMBER to indicate to the child to *think* about their answer, and (c) GREEN to indicate to the child that they can give their answer. It was suggested that the readiness indication mechanic could be integrated between amber and green phases. For instance, after about five seconds of thinking time under the amber icon, an "I'm ready to answer!" button could pop-up, which the child could press (similar to raising their hand in class) once they determine they have thought about their answer long enough; this would be a new mechanic between the existing M2 and M3 in the teacher-led ATMSG. Pressing the button, would change the icon to green, and allow the child to input their answer.

Finally, idea #5 was about scaffolding the S&T activities with a walk-through on how to stop-and-think in the context of science and maths problems. It was proposed that this walk-through would be presented at the beginning of the game (before M1). It would walk children through a generic mathematics and science problem and (i) suggest that the child first stop what they are doing and read/listen to the problem very carefully; (ii) indicate that the S&T icon will flash to remind them to use the time to think carefully about the problem; and (iii) advise them they could look at each part of the problem in turn and think about what they may have learned about the concept in school or elsewhere.

3.2.5. Activity 4: Sharing and contextualising

In Activity 4, participants shared their breakout work with the whole group, further contextualising how the proposed idea would alter the existing game, and received comments and input from others. Participants determined that (a) the idea #5 on modelling IC behaviours in a tutorial would be integrated into the child-led version of the game, as a matter of best practice, and that (b) ideas #1 (rewards), #2 (readiness indication mechanic), and #3 (symbolic colour to replace pulsating motion of the S&T icon) would be tested for their efficacy in promoting IC-related behaviours an experimental user study. Ideas #4 (on adaptivity) and #6 (on game progression), which did not receive additional attention during the breakout activity of the workshop, were documented for future consideration after the user study (Section 3.3.2). Throughout the idea development and sharing activities of the workshop, the ATMSG was useful in contextualising how ideas fit and would alter the flow of the game.

3.2.6. Post-workshop activity: Illustrating the new ideas using the ATMSG

The four ideation outputs from the interdisciplinary design workshop were structured into a new child-led ATMSG visual map to include the following *additional intrinsic instructional* component 'M'echanics, which are visualised in Fig. 6: (M1) an additional walk-through tutorial on how to "stop and think" in the context of generic mathematics and science problems; (M3) use of symbolic colour in the S&T icon to indicate when to stop (red), think (amber), and respond (green); (M4) a readiness indication mechanic, where the child could indicate when they felt they were finished stopping-and-thinking; and (M12) a reward system, where the child is awarded tokens for their performance after each correct answer. The ATMSG lacks a way in which designers can communicate visual design strategies to stakeholders (e.g., symbolic colour-coding vs pulsating motion); so, in Fig. 6, we used teal colour-blocking to highlight the mechanics that were experimented upon (M3, M4, M12) and varied depending on the experimental version of the game. Substantial textual description had to be provided alongside the ATMSG for these colour blocks to carry meaning. The child-led ATMSG visual map was used to communicate changes from the design team to project leads for approval, as well as to the developer, who used it both estimate costs for and follow the planned design in subsequent development. The developer gave positive feedback regarding the capacity of the map (which used UML) to communicate what the design team wanted, leading to few miscommunications. However, they remarked that they had to rely heavily on the textual descriptions to make sense how different visual treatments mapped to different conditions, which they saw as a weakness of the ATMSG visual map.

Once the design changes were approved by the project team, four discrete versions of a redesigned child-led S&T (Fig. 7) were developed for experimental testing in Phase III: Version 1 of the child-led ATMSG was similar to the teacher-led S&T (i.e., using a pulsating S&T icon for M3), except that it integrated the walk-through tutorial (M1) at the beginning of the game. Building on Version 1, Version 2 also implemented the pulsating icon (M3), but additionally integrated the "I'm ready to answer!" button (M4) as a representation of the readiness indication mechanic. Version 3 also integrated the readiness button but, in addition, changed the visual cue in M3 from pulsating motion of the S&T icon to symbolic "traffic light" colours. Lastly, Version 4 used symbolic colour in M3, integrated the M4 readiness indication mechanic, *and* presented tokens as a reward after each correctly answered question, with bonus multiplier (M12). All versions integrated the M1 tutorial, and the efficacy of this change was not evaluated experimentally.

3.3. Phase III: Testing ideas through a randomised user study

The four experimental versions of the child-led S&T were tested in a randomised eye-tracking user study with 45 7- to 8-year-old children across two urban state primary schools in England, who had not participated in previous S&T interventions (Gauthier et al., 2022). The objective of this user study was to evaluate the efficacy of our different design ideas in supporting IC-related behaviours in the absence of extrinsic instruction. This was done to guide further design and development, rather than to test the effectiveness of the game in an ecologically valid context (e.g., during homeschooling); as such, the study was performed in a quiet room in the children's school, rather than in the home. The full methods of this study are described in detail in Gauthier et al. (2022). In summary, children (who assented to the research and received written parental consent) played one of the four child-led versions of S&T for 12 min, whilst their screens were recorded, and their gaze on the screen tracked. In an independent-samples design, we analysed children's IC-related behaviours, such as time spent stopping-and-thinking and their eye fixations on answer items (e.g., blocks in Fig. 3; animals in Fig. 7), the question textbox, and the S&T icon.

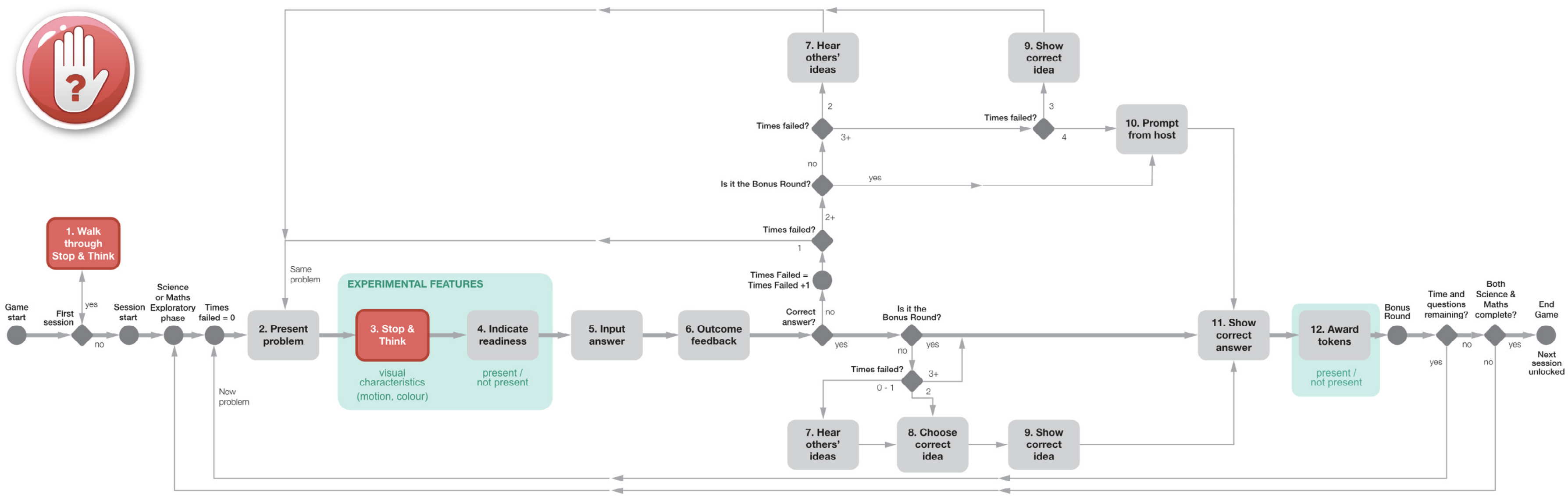
Below, we summarise the insights drawn from this study's findings about the design features that worked best to support IC-related behaviours (Section 3.3.1), and then, based on these, we propose additional modifications to be made to the child-led game using the ATMSG (Section 3.3.2).

3.3.1. Insights from user study: which ideas best promoted IC behaviour?

We drew four mains design insights that drove proposed changes. Firstly, we found that the combination of symbolic colour (M3) and the readiness indication mechanic (M4) was associated with longer stopping-and-thinking duration (itself related to better answer accuracy) and longer eye-fixations on answer elements. This suggests that Version 3 of the child-led S&T game (using symbolic colour and the readiness indication mechanic) promoted IC-training better than other versions—that is, children demonstrated better stopping-and-thinking skills when supported by these features, so these features will be included in the next iteration.

Secondly, we found, anecdotally, that giving children rewards after every correct answer (M12; Version 4) appeared motivating because children were generally more animated during gameplay (e.g., fist-pumping when earning bonus multipliers). Yet the presence of rewards seemed to distract the child from the IC-training task; the eye-tracking data revealed that children stared more at the S&T icon, presumably waiting for it to change to be able to answer quickly and then reap the reward before the end of the 12-min gameplay, rather than carefully considering the question-and-answer elements. Reflecting on this, we considered an alternative could be a trophy that is rewarded at the end of each session, rather than rewards on each turn, to slow the child down. Importantly, this end-of-session trophy could be accompanied by an Open Learner Model (OLM), to help the child connect the reward with their use of IC. A "learner model" is a structured data representation of a learner's current knowledge, skills, and other traits, that is used to *dynamically adapt* learning activities and/or interfaces to individual users, thereby creating more effective and personalised learning experiences (refer to the fourth insight below). OLMs visualise these learner models, usually through graphics and data visualisations, e.g., skills-o-meters, traffic-light icons, smiley faces, simple charts (Bull & Kay, 2016). In S&T, the OLM might display the average time the child spent "stopping and thinking", the number of clicks they made when they were supposed to be stopping and thinking, and their answer accuracy. The OLM could promote metacognitive development by encouraging children to reflect on how their in-game accuracy is related to their IC-related behaviour and how this is attributed to the trophy reward. Ultimately, we expect this could help children develop a deeper appreciation of the stop-and-think skill and how they might apply this in future learning.

Thirdly, whilst we did not test it experimentally, the results suggested that children might benefit from (a) additional integration of the S&T walk-through (M1) into the gameplay itself, to remind them of the desired behaviour, and (b) full exploratory-style scaffolding (e.g., Fig. 6 – M7-10) throughout the structured



		1. Walk through Stop & Think	2. Present problem	3. Stop & Think	4. Indicate readiness	5. Input answer	6. Outcome feedback	7. Hear others' ideas	8. Choose correct idea	9. Show correct idea	10. Prompt from host	11. Show correct answer	12. Award tokens
Gaming	Actions	x	x	x	x	Answer questions/trivia	See performance evaluation	Obtain help, listen to/read information	Select	x	Obtain help	x	See performance evaluation
	Tools	x	x	x	x	Graphic puzzle or text entry	Acknowledgement correct/incorrect	Dialogue from non-player characters	Dialogue from non-player characters	x	Advice from guide character	x	Performance metric, tokens
	Goals	x	x	x	x	Solve puzzle/trivia question	Solve puzzle/trivia question	Solve puzzle/trivia question	Solve puzzle/trivia question	x	Solve puzzle/trivia question	x	Maximise performance
Learning	Actions	Understand, generalise	x	Recognise, compare, distinguish	x	Choose, classify, complete, categorise	x	Find more information	Recognise, distinguish, choose	Recognise, relate	Recognise, relate	Memorise	x
	Tools	Tutorial: written info/ arguments, graphics	x	Delay interaction (wait)	x	Graphics, text input	x	Written info/arguments	Written info/arguments	Written info/arguments	Written info/arguments	Graphics/text	x
	Goals	Integration, adaptation	x	Reflective observation	x	Application, integration, active experimentation	x	Internalise values/other perspectives	Integration	Integration, adaptation	Understand	Remember, internalise values	x
Intrinsic instr.	Actions	Support recovery from errors, scaffold, demonstrate	Present problem	Stress importance of stopping & thinking	Scaffold	x	Assess performance	Support recovery from errors, scaffold	Reinforce idea	Demonstrate / reinforce idea, scaffold	Support recovery from errors	Demonstrate	Quantitatively assess and reward good performance
	Tools	Help text, tips and assistance	Challenge	Stop & Think icon	"I'm ready to answer" button	x	Performance measure (correct/incorrect)	Tips and assistance, limited set of choices	Limited set of choices	Display correct idea	Tips and assistance	Display correct answer	Rewards, feedback
	Goals	Provide learning guidance	Elicit performance, stimulate recall of prior learning	Provide learning guidance	Provide learning guidance; enhance confidence	x	Provide feedback	Provide learning guidance	Enhance retention and transfer	Provide learning guidance	Provide learning guidance	Provide learning guidance	Assess performance, encourage replay

Fig. 6. ATMSG visual map for the experimental child-led version of *Stop & Think*, for use in the user study.

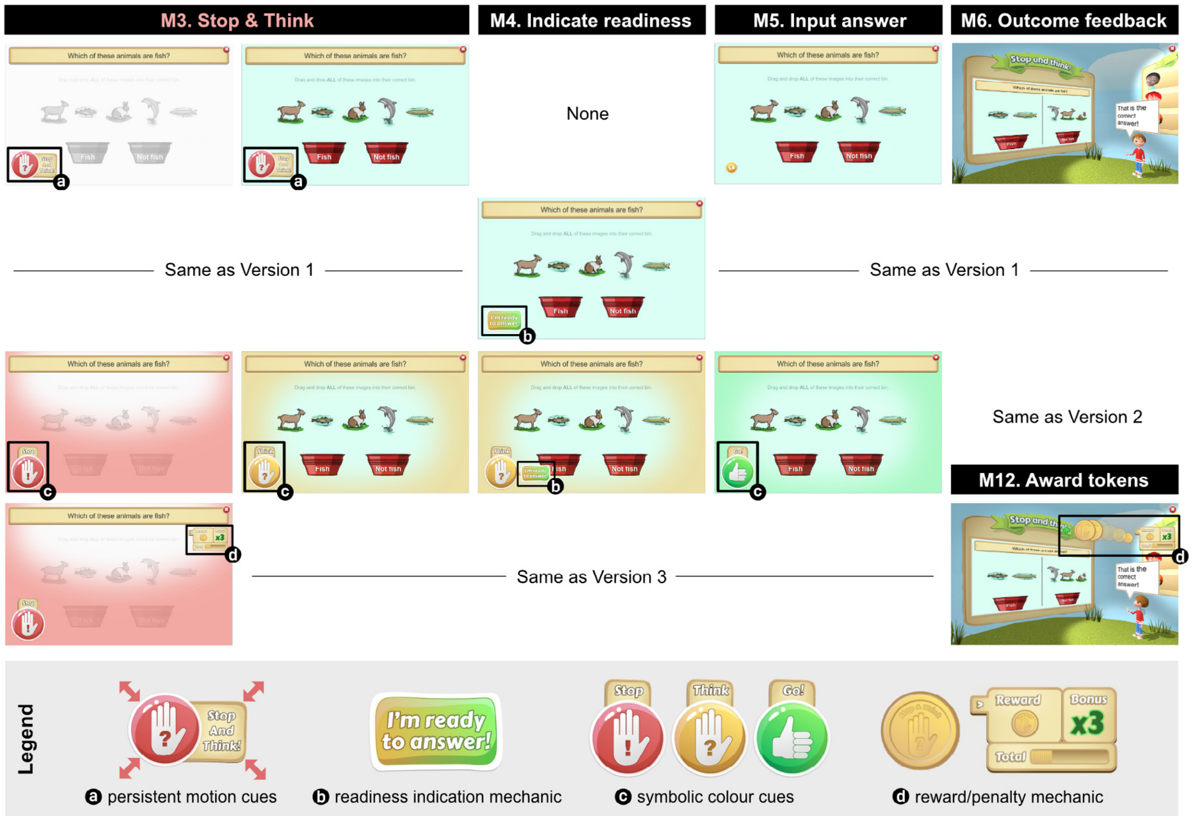


Fig. 7. Screenshots of four versions of the experimental child-led *Stop & Think* game, which integrated ideas from the Phase II workshop. A science question is depicted. M-numbered mechanics relate back to the ATMSG visual map in Fig. 6. Note for Version 4: upon outcome feedback on a correct answer (M6), awarding tokens (M12) follows very quickly.

practice problems. This is because several children in the experiment repeatedly got the same questions incorrect, whilst not displaying proper stopping-and-thinking behaviour. The repetition of these mechanics in the game flow would function to replace the role of teachers modelling the 'stop and think' behaviour in the classroom. For example, if the child consistently gets answers incorrect, then, instead of simply being presented with the correct answer (M13), they would first be "walked through" how one might "stop and think" for that particular problem. For instance, take the example of number values presented in Fig. 3. Part of the walk-through might pulsate each block-object on the screen in turn and pose a question about how these objects relate to Andy's conceptual prompt (M11), to demonstrate the IC behaviour.

Fourthly, the results of this user study also informed how we might design the adaptivity of the child-led S&T, to make the experience individualised to each child. Since we found that in-game performance was positively related to time spent stopping-and-thinking, we suggested that we might calibrate the "enforced thinking time" (i.e., when the icon is amber, before the "I'm ready" button pops up, M3-4), based on children's answer accuracy in combination with their average thinking time. In this way, children exercising IC-behaviours would transition toward engaging in these behaviours voluntarily, whilst those struggling would continue to get support. Additionally, once children consistently engage in stopping-and-thinking *voluntarily*, the visual scaffolding (i.e., the traffic-light icons and "I'm ready" button) could also be faded away, to transition children toward engaging in IC-behaviours in un-cued environments (Gauthier et al., 2022).

3.3.2. Proposed changes to S&T reflected in the ATMSG visual map

Based on the insights drawn from the user study, as outlined above, more changes were proposed to the child-led game. The ATMSG visual map from the previous iteration played a prominent role in preparing the redesign proposal, specifically by being used as a communication tool between the first author and other team members in negotiating the proposed changes. A final snapshot is shown in Fig. 8, where the visual map from Fig. 6 has been digitally marked up by the design team during this planning process.

Based on this markup, the ATMSG visual map was modified a final time (at least at the time of this publication) to first communicate the new proposed changes to project leads and then receive approval and then to the developer to follow. Fig. 9 depicts the most recent proposed redesign of the child-led S&T through a last ATMSG visual map. This version has not yet been developed, and we are currently seeking funding to explore some of these design ideas in more detail. In keeping with the remit of the project, much of the structure of the game and content is similar to the teacher-led and previous child-led versions of S&T. However, key changes were planned to support children's IC-training in the absence of extrinsic instruction, which were informed by the results of the user study, as well as by ideas generated in the workshop that were not yet evaluated. Specifically:

(1) The child receives a walk-through on how to "stop and think" in the context of a generic mathematics and science problem (Fig. 9, M1), upon first launching the game;
(2) The child chooses an upcoming unlocked session or a previously completed session (M2), to encourage goal-setting and develop a sense of progression;
(3) Symbolic colour of the S&T icon guides children through the 'stop and think' training (M4), instead of a pulsating icon;

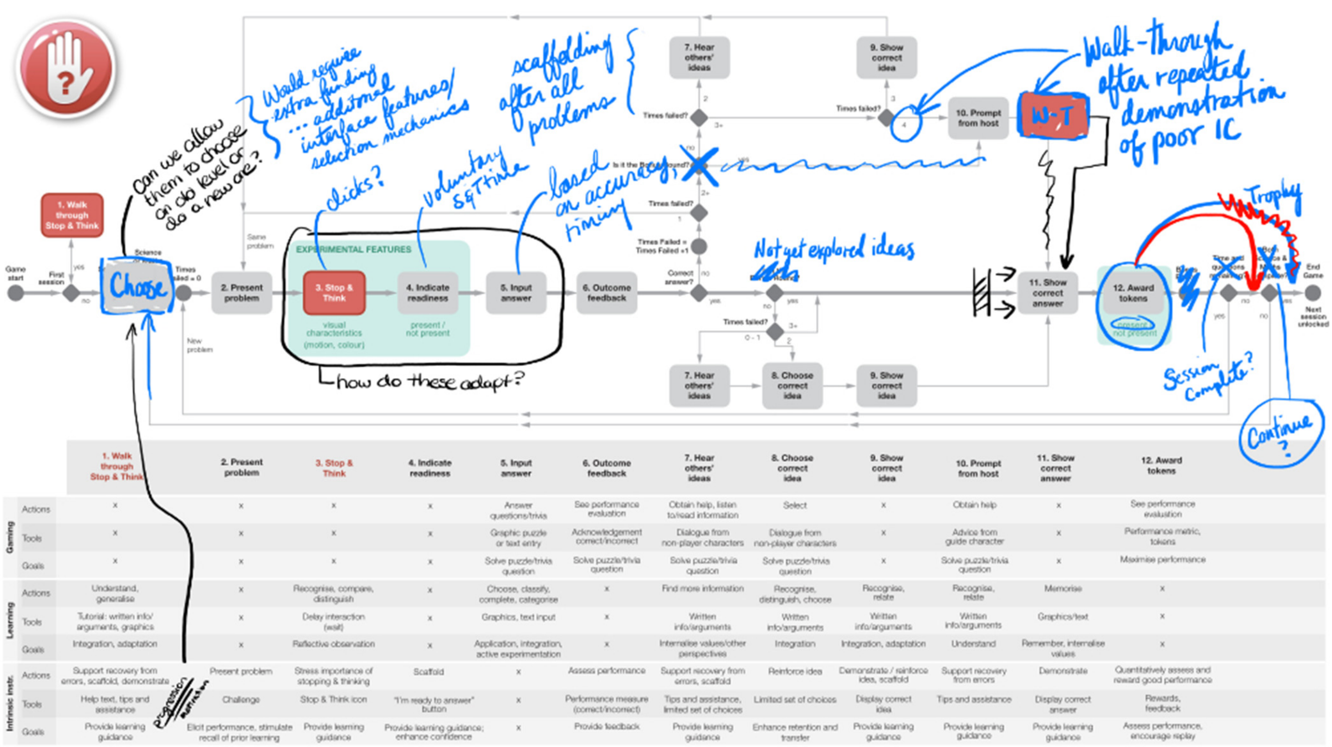


Fig. 8. Digital markup of the experimental child-led ATMSG visual map (Fig. 6), made by the designer when planning the proposed changes to the game, based on findings from the user study.

(4) The readiness mechanic (through an "I'm ready to answer!" button; M5), which boosts children's confidence in knowing that they are ready to answer, must be pressed before children are allowed to input their answer;
(5) There is no differentiation between the scaffolding provided from in the Exploratory and Structured Practice phases of the gameshow (M8-11), to ensure that children get the support they need and to reinforce the IC behaviour as the game progresses;
(6) The 12-min time restriction is removed, to compensate for the extra time associated with increased scaffolding described above. When playing at home, there is not necessarily a requirement to restrict playtime to 12 min, like at school;
(7) A walk-through tutorial is embedded within the gameplay (M12), to support children who struggle repeatedly with stopping-and-thinking or the science and mathematics content; and
(8) Upon completion of all problems in the session, the child is presented with a bronze, silver, or gold trophy based on their performance, to promote replay, alongside an OLM, to promote meta-cognition (M14).

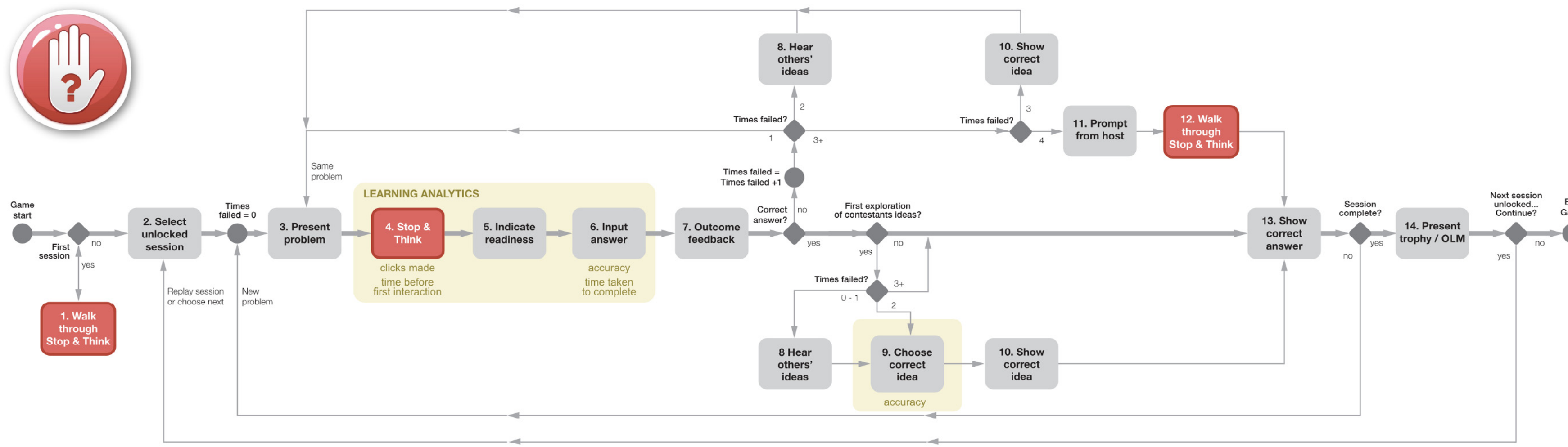
Finally, a non-structural change was the addition of adaptivity, which also needed to be documented in the new ATMSG in some way. We did this through the yellow-shaded blocks in Fig. 9, to represent where learning analytics would be collected and used to drive the adaptivity of M4 and M5, as described in Section 3.3.1.

4. Discussion

The aim of this paper was to document the process of redesigning the S&T learning game across three redesign phases, each of which leveraged the ATMSG. Several frameworks exist to help researchers document their learning game design and explain it to others (for examples, see Amory, 2007; Arnab et al., 2015; Kelle et al., 2011; Koops & Hoevenaar, 2012; Starks, 2014). We chose the ATMSG framework by Carvalho et al. (2015) because of its clear and concrete approach to mapping gaming, learning, and instructional components of mechanics to the flow of computerised interventions. It stood out because it **distinguished between intrinsic instruction (embedded in the game) and extrinsic instruction (given by a teacher, parent, or peer)** components to learning game mechanics, which we felt was particularly relevant considering we aimed to redesign the game from a teacher-led context to a child-led context. Whilst the ATMSG has been used for research dissemination and game evaluation purposes (Atmaja et al., 2020; Callaghan et al., 2016, 2018; Gauthier & Jenkinson, 2018), the iterative use of the ATMSG framework within a game design and development pipeline has not yet been documented to the best of our knowledge. Below, we discuss the advantages and challenges of implementing the ATMSG in this effort, as well as limitations in our design process and future directions.

4.1. Advantages of using the ATMSG

We found the ATMSG to be invaluable in four ways. Firstly, we found the ATMSG to be a useful tool to **describe the design** of our original, teacher-led learning game **in a concrete way**, which was a critical first step in the redesign process. Providing robust descriptions of learning games has been acknowledged as a necessity often neglected in game-based learning research (Boyle et al., 2016; Clark et al., 2016). The ATMSG provided us a clear and comprehensive way to analyse the teacher-led game and to document how the flow and composition of game mechanics was intended to promote children's use of IC and increase their proficiency of counterintuitive reasoning. Specifically, the map visually demonstrated (i) the integration of IC training (i.e., the "stop and think" mechanic) within the context of science/mathematics problem-solving, (ii) how the game scaffolded



		1 & 12. Walk through Stop & Think	2. Select unlocked session	3. Present problem	4. Stop & Think	5. Indicate readiness	6. Input answer	7. Outcome feedback	8. Hear others' ideas	9. Choose correct idea	10. Show correct idea	11. Prompt from host	12. (see 1)	13. Show correct answer	14. Present trophy / OLM
Gaming	Actions	x	Select, advance game	x	x	x	Answer questions/trivia	x	Obtain help, listen to/read information	Select	x	Obtain help		x	See performance evaluation
	Tools	x	Level map	x	x	x	Graphic puzzle or text entry	x	Dialogue from non-player characters	Dialogue from non-player characters	x	Advice from guide character		x	Performance metric, trophies
	Goals	x	Form goal	x	x	x	Solve puzzle/trivia question	x	Solve puzzle/trivia question	Solve puzzle/trivia question	x	Solve puzzle/trivia question		x	Maximise performance
Learning	Actions	Understand, generalise	x	x	Recognise, compare, distinguish	x	Choose, classify, complete, categorise	x	Find more information	Recognise, distinguish, choose	Understand, memorise	Recognise, relate		Memorise	Examine, evaluate
	Tools	Tutorial: written info/ arguments, graphics	x	x	Delay interaction (wait)	x	Graphics, text input	x	Written info/arguments	Written info/arguments	Written info/arguments	Written info/arguments		Graphics/text	Models, graphics/text
	Goals	Integration, adaptation	x	x	Reflective observation	x	Application, integration, active experimentation	x	Internalise values/other perspectives	Integration	Integration, adaptation	Understand		Remember, internalise values	Reflective observation, learning how to learn
Intrinsic instr.	Actions	Support recovery from errors, scaffold, demonstrate	Present material	Present problem	Stress importance of stopping & thinking	Scaffold	x	Assess performance	Support recovery from errors, scaffold	Reinforce idea	Demonstrate / reinforce idea, scaffold	Support recovery from errors		Demonstrate	Quantitatively assess and reward good performance
	Tools	Help text, tips and assistance	Limited set of choices	Challenge	Stop & Think icon, no interaction possible	"I'm ready to answer" button	x	Performance measure (correct/incorrect)	Tips and assistance, limited set of choices	Limited set of choices	Display correct idea	Tips and assistance		Display correct answer	Rewards, feedback
	Goals	Provide learning guidance	Gain attention/ motivation	Elicit performance, stimulate recall of prior learning	Provide learning guidance	Provide learning guidance; enhance confidence	x	Provide feedback	Provide learning guidance	Enhance retention and transfer	Provide learning guidance	Provide learning guidance		Provide learning guidance	Assess performance, provide feedback, encourage replay

Fig. 9. ATMSG visual map for the proposed adaptive, child-led version of *Stop & Think*.

supports to help the children grasp counterintuitive concepts, and (iii) the important role of the teacher in acting as an extrinsic instructional component of the game. Specifically, producing this first ATMSG visual map enabled us then to analyse the game and **identify what would be lost by removing the extrinsic instructional** component of the game—i.e., the teacher—thereby outlining areas requiring new game mechanics or modification of existing game mechanics (Phase I, Section 3.1). Removing the teacher from the equation is potentially detrimental to the effectiveness of the intervention (Benton et al., 2019). We found four gaps that would be created with the removal of the teacher related to reinforcing the stop-and-think behaviour, stimulating recall of prior knowledge during error recovery, motivating and scaffolding the gameplay, and addressing children's frustration, confusion or inappropriate behaviours.

Secondly, we printed out the teacher-led ATMSG visual map for our design workshop, **to facilitate constrained ideation** by discussing how different design ideas would impact the structure of the game and the feasibility of the approach (Phase II, Section 3.2). Previous work has indicated that adding constraints to brainstorming can be effective in helping design teams meet the objectives of complex design tasks (Bonnardel & Didier, 2020); we found that the ATMSG was a good way evoke those constraints. During the workshop's brainstorming and idea development activities, we asked participants to indicate how their ideas fit into the current structure of the game, using the ATMSG as a visual prompt, which allowed the S&T software designers/developers to explain, on the spot, how the feasibility of some ideas might be limited. For example, the idea of the "I'm ready to answer!" to be placed directly after the stop-and-think mechanic would not cause major game-flow disruption and was deemed feasible by developers; however, creating two different play modes (i.e., home vs school) would necessitate two parallel game flows, which the developers argued was out of scope for the time being.

Thirdly, we shared updated ATMSG visual maps amongst the project team **to plan and communicate proposed changes**, seek approval for these changes, and guide developers to implement these changes for the user study (Phases II and III, Sections 3.2.6 and 3.3.2). For instance, the visual format of the map was useful in quickly sketching out small modifications based on the results of the user study (Fig. 8) and discussing this within the design team before disseminating the proposal to stakeholders more broadly. UML is also a common visual language used by programmers and designers alike (Gomaa, 2006), which facilitated communication of design decisions between designers and developers.

Fourthly, the persistent use of ATMSG visual maps has allowed us to keep a **catalogue of changes** made to the game throughout the design process. Each iteration is a record of S&T's design state at a point in time, which we can refer to and use to reflect on how new proposed changes might impact learning and IC behaviours. This catalogue may also **enable reproducibility** of the game and the results of scientific findings made through their use by others (Munafò et al., 2017).

4.2. Challenges in using the ATMSG

Despite these advantages, there were some challenges in using the ATMSG during the redesign process. Firstly, the ATMSG maps illustrate connections between gaming, learning, and instructional mechanics – but **do not facilitate the illustration of visual strategies**, which may also impact learning and engagement. For instance, in our user study, both Versions 2 and 3 of the child-led S&T had identical ATMSG visual maps (Fig. 6), since their only difference was the visual treatment of the S&T icon (motion vs colour); yet Version 3 stood out as the most efficacious, which could not have been deduced by examining the ATMSG mapping alone. Similarly, in its current configuration, the framework also **lacks a way to visualise adaptive mechanisms** that rely on learning analytics; we have dealt with these two deficiencies in this project by extending the map through annotations, using differently coloured blocks to highlight mechanics received different visual treatments (as in Fig. 6), or that would adapt during play (as in Fig. 9). Others might consider using a similar approach or devise alternative methods for integrating this information into the map.

Secondly, S&T is not a very complex game, yet it **necessitated a very wide-format ATMSG map**, barely small enough to fit on a single page of A4 paper and still be legible. It may be a more effective communication tool for even simpler games (e.g., Atmaja et al., 2020; Callaghan et al., 2018). More complex games that apply the framework might have to rely on digital dissemination of their ATMSG map for it to communicate effectively.

Thirdly, **creating the initial mapping was very time consuming**; as noted by the creators of the framework themselves (Carvalho et al., 2015), the framework has a steep learning curve and requires in-depth analysis of the learning game and the available taxonomies to be able to create accurate mappings. This challenge is exacerbated by the need to use external software to create the maps; non-designers (e.g., teachers who might be involved in the design process) may not have the technical skills to create these types of visual maps, unless a specialised template were to be created to support this process—a possible future direction for research and industry. In the meantime, teams with less design-oriented members might explore the use of simpler drag-and-drop UML programmes, e.g., Lucidchart (Lucid, 2022) or Cacoo (Nulab, 2022), to see if such programmes make the creation of the visual maps less cumbersome.

Challenges aside, the ATMSG provides the most complete and detailed framework—to the best of our knowledge—for explaining how learning game mechanics incorporate gaming, learning, and instructional activities into their designs, with the specific advantage of distinguishing between intrinsic and extrinsic instructional activities. Its extension presented herein, through colour-block annotations, may be helpful for other researchers when specifying experimental treatments and/or adaptive mechanisms in learning games.

4.3. Limitations and future directions

We would like to acknowledge an obvious limitation in the design of our learning game—the lack of child involvement in the (re)design process. This was in large part due to (i) limitations in our funding remit, which was cognitive neuroscience-focused rather than design-focused, as well as (ii) access to children-participants during the pandemic. It is generally accepted that, by designing technology *for* children *with* children through participatory design approaches, designers can expect to come up with diverse child-centred ideas and empower children, thereby increasing their engagement in science and buy-in to the project (Fails et al., 2012). However, given the strict remit of our funding and constraints of ideation, involving children in the ideation process may have been challenging. For instance, Vasalou et al. (2021) faced tensions in negotiating their funded project plan with the values and ideas of participants, whilst Parsons and Cobb (2014) questioned whether outcome-focused agendas (like ours) could ever be compatible with the empowerment-focused agendas of participatory design. Despite this, we can envision the ATMSG also being useful in more exploratory, participatory design efforts that are common in child–computer interaction literature (e.g., Khaled & Vasalou, 2014; Nunes et al., 2016; Vasalou et al., 2021). For example, in addition to the functions listed above, the ATMSG could be used in the early phases of game design in participatory workshops, where taxonomies for gaming, learning, and intrinsic/extrinsic instruction components might be printed as inspiration cards (e.g., Vezzoli et al., 2020). These could be played with by both teachers and children to ideate on different ways in which the fun and serious elements of learning games might fit together. Going forward, we plan to seek more opportunities to engage children in the decision-making process once additional funding is secured. Ultimately, the final version of our child-led S&T is not yet developed and remains untested in its proposed format (Fig. 9), so there is room for further design decisions to be made through more participatory approaches before further development. For example, children could play a critical role in early phases of designing the OLM interface, which will aim to provide them with feedback on their IC-behaviour to support their metacognition (Bull & Kay, 2013).

Our final proposed child-led S&T raises many interesting design-related research questions, so variations of this design should be tested for efficacy going forward. For instance, should the OLM be presented after every problem, rather than at the completion of the session? Does the scaffolding through contestants' ideas (M6-M8 in Fig. 9, delivered at every incorrect answer) overwhelm and annoy the child, and should more conditions be met before delivering this additional scaffolding? We foresee that the ATMSG might play a critical role in further evaluation of the child-led S&T game in answering such design questions. Like work by Gauthier and Jenkinson (2018) and Gauthier et al. (2022), future research could collect click-stream interactions from multiple variations of the child-led game. Each variation would have its own unique ATMSG visual map, so that design differences can be directly compared. The click-stream data could then be overlaid on top of the ATMSG maps (see (Gauthier & Jenkinson, 2018), Fig. 6), to quantify how changes in design alter in-game IC-related behaviour—and importantly—how these alterations are related to learning outcomes.

Again, due to limitations in funding, we were unable to consider more extensive modifications to the structure of the S&T software. For instance, S&T could have been redesigned in such a way to support collaborative training of IC skills in dyads or triads over remote connection (e.g., Gallardo-Virgen & de Villar, 2011; Gijlers et al., 2013; Jackson et al., 2013; Tsuei, 2012), which would have opened new avenues of research on remote CSCL technologies for children. This raises questions about how to apply the ATMSG in this scenario; how might the framework incorporate the analysis of collaboration and cooperation in learning games? The designers of the ATMSG raise this as an open question (Carvalho et al., 2015). Constraints in the project also meant that we were unable to consider scope for different implementations of extrinsic instruction that could be feasible in homeschooling scenarios. Without such constraints, we could have designed and developed additional software, such as a teacher- or parent-facing learning analytics dashboard. This could have enabled adults to monitor their child's game progress and skills development remotely, provide feedback, and orchestrate future learning with and without S&T (Aleven et al., 2016; Knoop-van Campen et al., 2021; Pelanek, 2021; Rienties et al., 2018; Xhakaj et al., 2017). Would such dashboard mechanics fit within the extrinsic instruction activity on the child-led ATMSG map? Or would it require its own visual mapping? The ATMSG has not been tested for this purpose, which could be an interesting direction to explore in future research.

5. Conclusion

This paper describes the use of the Activity Theory-based Model of Serious Games (ATMSG) during the redesign of *Stop & Think* (S&T), a learning game to train children to apply their inhibitory control (IC) skills when answering counterintuitive science and mathematics problems. This type of computerised training has been shown to increase scores on standardised assessments, as children become better able to suppress their initial incorrect ideas and misconceptions and adopt counterintuitive (but correct) concepts by engaging the analytical parts of their brain (Diamond & Lee, 2011; Mareschal, 2016; Roy et al., 2019: Wilkinson et al., 2019). The game was previously found to be effective in a classroom context, where children played as a group, with the activity led by the teacher (Roy et al., 2019; Wilkinson et al., 2019). However, this redesign effort sought to ensure that the game would also be effective if children played the game independently, without the extrinsic support of a teacher, e.g., during homeschooling. This was identified as important in light of the Covid-19 pandemic and subsequent lockdowns.

The contribution of this paper is as an exemplar for the ATMSG's use in the iterative redesign of a game to fit this new learning context. We discuss the advantages and challenges of using the framework throughout three phases of redesign. Specifically, we found it to be a useful tool because (i) it distinguished between the *intrinsic* and *extrinsic* instructional components of games, enabling us to identify design gaps in the absence of a teacher in homeschooling contexts; (ii) it facilitated constrained brainstorming in our design workshop, to generate ideas that were feasible within our project restrictions; (iii) it enabled negotiation of design decisions between team members and communication with stakeholders; and (iv) it provided a concrete way to catalogue iterations in the design that can be used to reflect upon future design considerations and may benefit reproducible science. Yet, its use also posed some challenges, including that it does not provide means to illustrate non-structural design decisions that can greatly impact learning (e.g., visual strategies, adaptive mechanisms), that the maps require a very wide format that can be impractical for complex games, and that the initial visual map is time-consuming to create. To this end, we have proposed alternative technologies and an extension to the framework to overcome these challenges. Finally, we discussed how the ATMSG may also be valuable in both user-centred and participatory design approaches and recommend its further exploration in learning game research, (re)design, and development.

Selection and participation

In this paper, we highlight an eye-tracking experimental user study (Section 3.3) that is described in detail elsewhere (Gauthier et al., 2022) but which deeply informed our design process. In this study, recruited occurred via emails sent out to around 500 primary schools in the region. For schools who agreed to participate, a short 20-min presentation was made to children in Year 3 classrooms, introducing them to the research. All these children were sent home an information sheet and informed consent form for parents to review and opt-in. Assent was also gained from the children who received parental consent before conducting the experiment. Secondly, participants in the design workshop (Section 3.1), who were all adults selected through existing academic networks, were emailed information about the workshop and told that, if they agreed to participate, any information provided by them might be used in future game design, development, and research publication.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Denis Mareschal reports financial support was provided for this project by the Wellcome Trust and by the Education Endowment Foundation.

Acknowledgement

We would like to thank the UnLocke Team, including Michael Thomas, Andrew Tolmie, Doug Lapsley, Wayne Holmes, Hannah Wilkinson, Dilini Sumanapala, Keith Anderson, and Su Morris, who have contributed substantially to the design, development, and evaluation of *Stop & Think* along the way.

References

Adobe (2022). Adobe illustrator. https://www.adobe.com/uk/products/illustrator.html.

Ajzen, I. (1985). From intentions to actions: A theory of planned behaviour. In K. J, & B. J (Eds.), *Action-control* (pp. 11–39). Springer, http://dx.doi.org/10.1007/978-3-642-69746-3_2.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. http://dx.doi.org/10.1016/0749-5978(91)90020-T.

Aleven, V., Xhakaj, F., Holstein, K., & McLaren, B. M. (2016). Developing a teacher dashboard for use with intelligent tutoring systems. In *Proceedings of the 4th international workshop on teaching analytics* (pp. 44–50). http://dx.doi.org/10.1145/12345.67890.

Allen, M. (2014). *Misconceptions in primary science* (2nd ed.). New York, NY: McGraw Hill Education, Issue.

Amory, A. (2007). Game object model version II: A theoretical framework for educational game development. *Educational Technology Research and Development*, *55*(1), 51–77. http://dx.doi.org/10.1007/s11423-006-9001-x.

Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F., De Freitas, S., Louchart, S., Suttie, N., Berta, R., & De Gloria, A. (2015). Mapping learning and game mechanics for serious games analysis. *British Journal of Educational Technology*, *46*(2), 391–411. http://dx.doi.org/10.1111/bjet.12113.

Atmaja, P. W., Muttaqin, F., & Sugiarto, S. (2020). Facilitating educational contents of different subjects with context-agnostic educational game: A pilot case study. *Scientific Journal of Information System Technology*, *6*(1), 53–65. http://dx.doi.org/10.26594/register.v6i1.1726.

Baptista, G., & Oliveira, T. (2019). Gamification and serious games: A literature meta-analysis and integrative model. *Computers in Human Behavior*, *92*(2018), 306–315. http://dx.doi.org/10.1016/j.chb.2018.11.030.

Benton, L., Vasalou, A., Barendregt, W., Bunting, L., & Révész, A. (2019). What's missing : The role of instructional design in children's games-based learning. In *CHI conference on human factors in computing systems proceedings (CHI 2019), 11-undefined*.

Bofferding (2019). Understanding negative numbers. In *Constructing numbers* (pp. 251–277). Springer.

Bonnardel, N., & Didier, J. (2020). Brainstorming variants to favor creative design. *Applied Ergonomics*, *83*(2019), Article 102987. http://dx.doi.org/10.1016/j.apergo.2019.102987.

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C., & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers and Education*, *94*, 178–192. http://dx.doi.org/10.1016/j.compedu.2015.11.003.

Bull, S., & Kay, J. (2013). *Open learner models as drivers for metacognitive processes* (pp. 349–365). http://dx.doi.org/10.1007/978-1-4419-5546-3_23.

Bull, S., & Kay, J. (2016). SMILI: A framework for interfaces to learning data in open learner models, learning analytics and related fields. *International Journal of Artificial Intelligence in Education*, *26*(1), 293–331. http://dx.doi.org/10.1007/s40593-015-0090-8.

Bull, S., Mangat, M., Mabbott, A., Abu Issa, A., & Marsh, J. (2005). Reactions to inspectable learner models: seven year olds to university students. Proceedings of workshop on learner modelling for reflection. In *International conference on artificial intelligence in education* (pp. 1–10).

Burkitt, E., Barrett, M., & Davis, A. (2003). Children's colour choices for completing drawings of affectively characterised topics. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *44*(3), 445–455. http://dx.doi.org/10.1111/1469-7610.00134.

Byun, J., & Joung, E. (2018). Digital game-based learning for K-12 mathematics education: A meta-analysis. *School Science and Mathematics*, *118*(3–4), 113–126. http://dx.doi.org/10.1111/ssm.12271.

Callaghan, M. J., Mcshane, N., Eguíluz, A. G., Teillès, T., & Raspail, P. (2016). Practical application of the learning mechanics – game mechanics (LM-GM) framework for serious games analysis in engineering education. In *13th International conference on remote engineering and virtual instrumentation, February* (pp. 382–386). http://dx.doi.org/10.1109/REV.2016.7444510.

Callaghan, M. J., McShane, N., Gómez Eguíluz, A., & Savin-Baden, M. (2018). Extending the activity theory based model for serious games design in engineering to integrate analytics. *International Journal of Engineering Pedagogy*, *8*(1), 109–126. http://dx.doi.org/10.3991/ijep.v8i1.8087, Michael.

Carvalho, M. B., Bellotti, F., Berta, R., De Gloria, A., Sedano, C. I., Hauge, J. B., Hu, J., & Rauterberg, M. (2015). An activity theory-based model for serious games analysis and conceptual design. *Computers and Education*, *87*, 166–181. http://dx.doi.org/10.1016/j.compedu.2015.03.023.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, *86*(1), 79–122. http://dx.doi.org/10.3102/0034654315582065.

Coulanges, L., Abreu-Mendoza, R. A., Varma, S., Uncapher, M. R., Gazzaley, A., Anguera, J., & Rosenberg-Lee, M. (2021). Linking inhibitory control to math achievement via comparison of conflicting decimal numbers. *Cognition*, *214*(May), Article 104767. http://dx.doi.org/10.1016/j.cognition.2021.104767.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old [article]. *Science*, *333*(6045), 959–964. http://dx.doi.org/10.1126/science.1204529.

Elbers, E., & Streefland, L. (2000). Collaborative learning and the construction of common knowledge. *European Journal of Psychology of Education*, *15*(4), 479–490. http://dx.doi.org/10.1007/BF03172989.

Evans, J. S. B. (2003). In two minds: dual-process accounts of reasoning [article]. *Trends in Cognitive Sciences*, *7*(10), 454–459. http://dx.doi.org/10.1016/j.tics.2003.08.012.

Fails, J. A., Guha, M. L., & Druin, A. (2012). Methods and techniques for involving children in the design of new technology for children. *Foundations and Trends in Human-Computer Interaction*, *6*(2), 85–166. http://dx.doi.org/10.1561/1100000018.

Fitzgerald, E., Jones, A., Kucirkova, N., & Scanlon, E. (2018). A literature synthesis of personalised technology-enhanced learning: What works and why. *Research in Learning Technology*, *26*, http://dx.doi.org/10.25304/rlt.v26.2095.

Gallardo-Virgen, J. A., & de Villar, R. A. (2011). Sharing, talking, and learning in the elementary school science classroom: Benefits of innovative design and collaborative learning in computer-integrated settings. *Computers in the Schools*, *28*(4), 278–290. http://dx.doi.org/10.1080/07380569.2011.621803.

Garneli, V., Patiniotis, K., & Chorianopoulos, K. (2021). Designing multiplayer serious games with science content. *Multimodal Technologies and Interaction*, *5*(3), http://dx.doi.org/10.3390/mti5030008.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming*, *33*(4), 441–467. http://dx.doi.org/10.1177/1046878102238607.

Gauthier, A., & Jenkinson, J. (2018). Designing productively negative experiences with serious game mechanics: Qualitative analysis of game-play and game design in a randomized trial. *Computers & Education*, *127*(2018), 66–89. http://dx.doi.org/10.1016/j.compedu.2018.08.017.

Gauthier, A., Porayska-Pomsta, K., Dumontheil, I., Mayer, S., & Mareschal, D. (2022). Manipulating interface design features affects children's stop-and-think behaviours in a counterintuitive-problem game. *Transactions on Computer-Human Interaction*, *29*(2), 1–22.

Gijlers, H., Weinberger, A., van Dijk, A. M., Bollen, L., & van Joolingen, W. (2013). Collaborative drawing on a shared digital canvas in elementary science education: The effects of script and task awareness support. *International Journal of Computer-Supported Collaborative Learning*, *8*(4), 427–453. http://dx.doi.org/10.1007/s11412-013-9180-5.

Girard, S. A.-S. (2011). *Traffic lights and smiley faces: do children learn mathematics better with affective open- learner modelling tutors?* University of Bath.

Gomaa, H. (2006). Designing concurrent, distributed, and real-time applications with UML. In *Proceedings - international conference on software engineering, 2006* (pp. 1059–1060). http://dx.doi.org/10.1145/1134285.1134504.

Jackson, A. T., Brummel, B. J., Pollet, C. L., & Greer, D. D. (2013). An evaluation of interactive tabletops in elementary mathematics education. *Educational Technology Research and Development*, *61*(2), 311–332. http://dx.doi.org/10.1007/s11423-013-9287-4.

Johnson, C. I., Bailey, S. K. T., & Van Buskirk, W. L. (2017). Designing effective feedback messages in serious games and simulations: A research review. In P. Wouters, & H. van Oostendorp (Eds.), *Instructional techniques to facilitate learning and motivation of serious games*. Springer International Publishing Switzerland, http://dx.doi.org/10.1007/978-3-319-39298-1.

Jonassen, D. H., & Rohrer-Murphy, L. (1999). Activity theory as a framework for designing constructivist learning environments. *Educational Technology Research and Development*, *47*(I), 61–79, https://link.springer.com/content/pdf/10.1007%2FBF02299477.pdf.

Jung, T., Huang, J., Eagan, L., & Oldenburg, D. (2019). Influence of school-based nutrition education program on healthy eating literacy and healthy food choice among primary school children. *International Journal of Health Promotion and Education*, *57*(2), 67–81. http://dx.doi.org/10.1080/14635240.2018.1552177.

Juul, J. (2009). Fear of failing? the many meanings of difficulty in video games. In M. J. P. Wolf, & B. Perron (Eds.), *The video game theory reader. Vol. 2* (pp. 237–252). Routledge: http://dx.doi.org/10.1017/CBO9781107415324.004.

Kelle, S., Klemke, R., & Specht, M. (2011). Design patterns for learning games. *International Journal of Technology Enhanced Learning*, *3*(6), 555–569. http://dx.doi.org/10.1504/IJTEL.2011.045452.

Khaled, R., & Vasalou, A. (2014). Bridging serious games and participatory design. *International Journal of Child-Computer Interaction*, *2*(2), 93–100. http://dx.doi.org/10.1016/j.ijcci.2014.03.001.

Knoop-van Campen, C. A. N., Wise, A., & Molenaar, I. (2021). The equalizing effect of teacher dashboards on feedback in K-12 classrooms. *Interactive Learning Environments*, 1–17. http://dx.doi.org/10.1080/10494820.2021.1931346.

Koops, M., & Hoevenaar, M. (2012). Conceptual change during a serious game: Using a lemniscate model to compare strategies in a physics game. *Simulation & Gaming*, *44*(4), 544–561. http://dx.doi.org/10.1177/1046878112459261.

Lucid (2022). Lucidchart. https://www.lucidchart.com.

Mareschal, D. (2016). The neuroscience of conceptual learning in science and mathematics. *Current Opinion in Behavioral Sciences*, *10*, 114–118. http://dx.doi.org/10.1016/j.cobeha.2016.06.001.

Mason, L., & Zaccoletti, S. (2021). Inhibition and conceptual learning in science: a review of studies. *Educational Psychology Review*, *33*(1), 181–212. http://dx.doi.org/10.1007/s10648-020-09529-x.

Meyer, B., Haywood, N., Sachdev, D., & Faraday, S. (2008). *What is independent learning and what are the benefits for students ? how is independent learning viewed by teachers? Schools and families research report, 051*, (pp. 1–6). London: Department for Children.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., Simonsohn, U., Wagenmakers, E. J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1–9. http://dx.doi.org/10.1038/s41562-016-0021.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*(2), 220–246. http://dx.doi.org/10.1037/0033-2909.126.2.220.

Nulab (2022). Cacoo. https://cacoo.com/.

Nunes, E. P. S., Luz, A. R., Lemos, E. M., & Nunes, C. (2016). Approaches of participatory design in the design process of a serious game to assist in the learning of hospitalized children. In M. Kurosu (Ed.), *Human-computer interaction. novel user experiences* (pp. 406–416). Springer International Publishing.

Parsons, S., & Cobb, S. (2014). Reflections on the role of the users: Challenges in a multi-disciplinary context of learner-centred design for children on the autism spectrum. *International Journal of Research and Method in Education*, *37*(4), 421–441. http://dx.doi.org/10.1080/1743727X.2014.890584.

Pelanek, R. (2021). Analyzing and visualizing learning data : A system designer's perspective. *Journal of Learning Analytics*, *8*(2), 93–104. http://dx.doi.org/10.18608/jla.2021.7345.

Proulx, J.-N., Romero, M., & Arnab, S. (2017). Learning mechanics and game mechanics under the perspective of self-determination theory to foster motivation in digital game based learning. *Simulation & Gaming*, *48*(1), 81–97. http://dx.doi.org/10.1177/1046878116674399.

Rienties, B., Herodotou, C., Olney, T., Schencks, M., & Boroowa, A. (2018). Making sense of learning analytics dashboards: A technology acceptance perspective of 95 teachers. *International Review of Research in Open and Distributed Learning*, *19*(5), 19. http://dx.doi.org/10.19173/irrodl.v19i5.3493.

Roy, P., Rutt, S., Easton, C., Sims, D., Bradshaw, S., & McNamara, S. (2019). Stop and think: learning counterintuitive concepts. In *NFER Evaluation Reports*.

Serif Europe LTD (2020). Affinity designer. https://affinity.serif.com/en-gb/.

Soloway, E., Guzdial, M., & Hay, K. E. (1994). Learner-centered design: The challenge for HCI in the 21st Century. *Interactions*, *1*(2), 36–48. http://dx.doi.org/10.1145/174809.174813.

Starks, K. (2014). Cognitive behavioral game design: A unified model for designing serious games. *Frontiers in Psychology*, *5*(FEB), 1–10. http://dx.doi.org/10.3389/fpsyg.2014.00028.

Thorell, L., Skoglund, C., Peña, A. de la, & Baeyens, D. (2020). *Psychosocial effects of homeschooling during the COVID-19 pandemic: differences between seven european countries and between children with and without mental children adolescent. 0–2*. https://psyarxiv.com/68pfx/.

Tsuei, M. (2012). Using synchronous peer tutoring system to promote elementary students' learning in mathematics. *Computers and Education*, *58*(4), 1171–1182. http://dx.doi.org/10.1016/j.compedu.2011.11.025.

Van Der Meijden, H., & Veenman, S. (2005). Face-to-face versus computer-mediated communication in a primary school setting. *Computers in Human Behavior*, *21*(5), 831–859. http://dx.doi.org/10.1016/j.chb.2003.10.005.

Vasalou, A., Ibrahim, S., Clarke, M., & Griffiths, Y. (2021). On power and participation: Reflections from design with developmentally diverse children. *International Journal of Child-Computer Interaction*, *27*, Article 100241. http://dx.doi.org/10.1016/j.ijcci.2020.100241.

Vezzoli, Y., Mavrikis, M., & Vasalou, A. (2020). Inspiration cards workshops with primary teachers in the early co-design stages of learning analytics. In *Proceedings of the 10th international conference on learning analytics and knowledge* (pp. 73–82). http://dx.doi.org/10.1145/3375462.3375537.

Wilkinson, H. R., Smid, C., Morris, S., Farran, E. K., Dumontheil, I., Mayer, S., & The UnLocke Team (2019). Domain-specific inhibitory control training to improve children's learning of counterintuitive concepts in mathematics and science. *Journal of Cognitive Enhancement*, *4*, 296–314.

Wood, D., & O'Malley, C. (1996). Collaborative learning between peers: An overview. *Educational Psychology in Practice*, *11*(4), 4–9. http://dx.doi.org/10.1080/0266736960110402.

Xhakaj, F., Aleven, V., & McLaren, B. M. (2017). *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*, *Effects of a teacher dashboard for an intelligent tutoring system on teacher knowledge, lesson planning, lessons and student learning* (pp. 315–329). http://dx.doi.org/10.1007/978-3-319-66610-5_23.