

Playtest Planning and Reporting Toolkit for Game Design in Academic Contexts

Mikhail Fiadotau^{1,2}, Peadar Callaghan¹, Leonardo Sorrentino¹, Laura de Carli^{2,3}, Martin Sillaots¹, Anni Hõ¹ and Juhung Shin⁴

¹Tallinn University, Estonia

²IT University of Copenhagen, Denmark

³Malmö University, Sweden

⁴Osaka Electro-Communication University, Japan

fiadotau@tlu.ee

Abstract: The article presents the Playtest Planning and Reporting Toolkit: a set of materials created to help game designers plan, structure, and report on their playtests in a way that is consistent, structured, and recoverable by others. The toolkit is particularly aimed at game creators who operate in academic contexts such as research and development projects and higher education. The toolkit consists of three parts: an information sheet outlining the different types and methods of playtesting, a template for planning playtest sessions, and a separate template to record playtest observations and report on the results. The article establishes the need for a playtesting toolkit in academic contexts; provides a functional definition of playtesting; and introduces the three parts of the toolkit, contextualizing their underlying design considerations.

Keywords: Playtesting, Game design didactics, Games education, Game development

1. Introduction

Playtesting is an integral part of the game development process. From the first prototype to the final build, playtesting is key to game designers' ability to validate their ideas and assumptions, as well as to verify that their design goals are accomplished for both the game as a whole and its individual atoms.¹ The importance of regular and rigorous playtesting is not only widely recognized in the game industry (Mirza-Babaei, Moosajee and Drenikow, 2016; Kramarzewski and De Nucci, 2023), but emphasized by educators in game design programs around the globe.

Fullerton (2018, p. 277) argues that playtesting is “the single most important activity a designer engages in,” yet it remains one of the least understood aspects of game design. Choi et al. (2016) further note that textbooks and educational materials on game design tend to provide limited practical guidance on the subject, forcing designers to improvise their playtesting procedures until they have developed a process that makes sense to them.

In academic contexts—from game design programs in higher education to research and development projects on game-based learning—playtesting poses an additional challenge: the expectation that the playtesting process is structured and reported in an academically rigorous way. As Khaled et al. (2018) argue, a core aim of game design research is to “materialise tacit design knowledge,” and key to that is *recoverability*: documenting and describing not just one's game, but the design process behind it, in such a way that readers can understand and scrutinize it. In the context of playtesting, this involves explicating and justifying the decisions that go into the playtesting procedure, consistently adhering to that procedure, and methodically documenting the evidence it produces—the same general expectations that apply to any research method.

Yet, the processes for running playtests and especially reporting their results are far from standardized. While there is no shortage of game design research that discusses playtesting, existing work lacks an established, transparent format to present the playtesting process and its results in accordance with academic standards. As a result, game designers in academic settings often find themselves reinventing the wheel and reporting on their playtest procedure and results in an ad hoc way.

To address this challenge, this paper presents the Playtest Planning and Reporting Toolkit, which we developed specifically for academic settings (including both game design education and research that involves game

¹We borrow the term “game atoms” from Elias et al. (2020, p.13), who define it as the “smallest complete unit of play,” i.e., the smallest subset of a game that can be meaningfully played. For example, having one platform and one playable character with a jump ability would probably not be enough to constitute an atom of a platformer game, but a screen containing the same character and several platforms to jump between would.

creation). The toolkit consists of three parts: (1) an outline of key approaches and considerations in playtesting; (2) a playtest plan template that allows designers to explicate their procedure, from setting goals, to recruiting playtesters, to gathering relevant data; and (3) a playtest report template that can be used to collect data from individual playtest sessions and summarize it. The toolkit has been used in two game design courses, refined during a collaborative workshop with a group of educators, and validated in three focus groups with game design students and recent alumni. While the toolkit is a work in progress, our intention is to provide a practical and referenceable set of materials that make it easier for game design students and researchers to plan their playtests, execute them, and report them in their papers and theses.

2. Overview of the Toolkit

The Playtest Planning and Reporting Toolkit was created with a dual purpose: first, to help game designers in academic settings plan and structure their playtests, and second, to provide a format to report playtest results more consistently. Drawing on the action research tradition (Tripp, 2005), we took an iterative approach, integrating the toolkit in teaching practice and refining it based on actual use and feedback from students and fellow practitioners.

The initial version was created in fall 2025 for a game design course at the IT University of Copenhagen, aimed mostly at MSc in Games students, where it was introduced during a workshop on conducting playtesting. In February 2026, the toolkit was refined during a two-hour workshop with six game design educators. Following this, an updated version of the toolkit was used in spring 2026 in a game design course at Tallinn University offered as part of the Digital Learning Games MSc program. In April 2026, three focus groups, involving a total of 16 game design students and recent alumni, were conducted to evaluate the updated version of the toolkit and refine it further. While the focus of this paper is on introducing the toolkit rather than presenting its empirical evaluation, some insights gained during the focus groups have informed the discussion.

Crucially, the Playtest Planning and Reporting Toolkit is not a set of guidelines on how to conduct playtests. For that, chapters on playtesting in textbooks by Fullerton (2018), Schell (2019), and Kramarzewski and De Nucci (2023) provide a useful starting point. Rather, the intention behind the toolkit is to make the process of planning and reporting playtest results more streamlined and transparent. While designed for use in academic game creation contexts, the toolkit may also be helpful in other settings (e.g., independent game development); however, reflecting on additional use cases is beyond the scope of this paper.

The Playtest Planning and Reporting Toolkit consists of three parts. The first is an information sheet outlining different types of playtesting. This is intended to help game designers decide and articulate what type of playtest to conduct, what data to gather, etc. The second part is the Playtest Plan: a template to help plan and prepare for an upcoming playtest. The third part is the Playtest Report: a template that follows a similar structure to the Playtest Plan, but is intended to both document a playtest as it happens and report the results afterwards. The following sections introduce each of the three parts in order.

The current version of the toolkit, including downloadable materials that can be printed out or filled out digitally, is available at <https://glg.tlu.ee/playtest-kit>.

3. Overview of Playtesting

A starting point for introducing the toolkit is to establish what playtesting is, what methods it commonly relies on, and what forms it takes. These aspects form the basis of the first part of the toolkit: an outline of approaches and considerations in playtesting.

3.1 Defining Playtesting

Game design literature contains many, sometimes contradictory, definitions of playtesting. Some of these are very descriptive and informal; others are narrow in scope and overlook specific facets or forms of playtesting (e.g., Denisova et al., 2024, define playtesting as a “formal research method,” whereas in reality many playtests are informal and ad hoc). During a workshop with six game design educators, we discussed what criteria a definition of playtesting should meet and agreed on the following:

1. It should conceptually differentiate playtesting from other forms of testing, such as various types of software testing (functional testing, performance testing, etc.) and usability testing.²
2. It should encompass diverse forms and methods of playtesting, rather than reflect a specific approach (in other words, the definition should be about *what* playtesting is, not *how* to do it).
3. It should be inclusive of both digital and non-digital games, reflecting the diversity of forms game design can take (e.g., tabletop games, location-based games, escape rooms, etc.).
4. It should not assume that the goal of playtesting is to maximize player satisfaction (e.g., some games are designed to challenge the player rather than entertain them, cf. Schell 2019), nor that the goal is always to ensure the game perfectly matches the designer's vision, as design intent is often a starting point rather than an end goal (cf. Myers 2020, p.17).

As we could not identify a definition in the literature that met all of the above criteria, we formulated our own: playtesting is the process by which game design decisions are evaluated to ensure alignment between designer expectations and player experience. This definition warrants unpacking. Game design is a second-order problem: we cannot design play itself, but only design for play by creating “the circumstances under which play might arise” (Zimmerman 2022, p.27). Designing such circumstances is a purposeful process that encompasses a broad spectrum of design decisions, ranging from individual game mechanics and resources, to progression, to how the player interacts with the game. As the designer makes those decisions, he or she forms a mental model of how the game and its constituent parts may behave when played and what experience this may create for the player: this is what we refer to as designer expectations. However, as games are complex emergent systems, this mental model often differs from what actually happens when people play. This may be because the game system itself does not function as intended (e.g., some of the game mechanics interfere with each other) or because the players engage with the game differently than expected (e.g., misinterpret the goal, overuse or ignore specific mechanics, or give up in frustration). Furthermore, players often “prioritize ... constructive and expressive play that undermines” the designer's intent (Myers 2020, p.17): in other words, players may recognize how the game is intended to be engaged with, but refuse to go along with it, preferring to play on their terms (cf. Meades 2015). Playtesting involves having players engage with the game or its atoms and using player observation and feedback to identify such gaps between the designer's mental model of the game and the reality of how people play it.

When these gaps are identified, the designer faces three options: make changes to the game to bring it more in line with their design goals, adjust their expectations to better reflect how players engage with the game, or both. For example, if during the playtest of a card game it becomes apparent that players never use a specific type of card, the designer may choose to adjust that card's values or presentation in an effort to change player behavior, or they may decide to accept that the card is unnecessary and remove it altogether. In some cases, designers may have to resort to the fourth option: accepting that their overall concept is not viable. If playtests show that the gap between designer expectation and player experience is too large to bridge, the designer may need to “kill [the project] and move to the next concept” (Clark, 2026) to avoid investing more resources into an idea that does not appear to be workable.

This last scenario points to the importance of starting playtesting early in the game development process. It is a mistake to delay playtesting until there is a polished version of the game, as by that point too much time may have been spent on implementing and refining features that should have been scrapped to begin with. Playtests should be continued on a regular basis throughout the game development process, accompanying every iteration of the game (Fullerton, 2018; Zimmerman, 2022). This is important to validate any new additions to the game, as well as to ensure that the changes introduced since the last iteration have addressed previously identified issues without creating new ones.

3.2 Playtesting Methods and Data

Playtesting can involve gathering many kinds of data from the players. The vast majority of playtests involve some form of **observation** of the game being played. In-person observation is the most common approach, but may create a sense of external pressure for the playtester, affecting their performance (Mirza-Babaei et al., 2016). Remote observation, on the other hand, may be less intrusive but may also be less informative since it

²There is some overlap between playtesting and functional testing (both can uncover bugs and technical issues), as well as playtesting and usability testing (both can identify player experience issues). However, in playtesting, finding bugs and usability issues (e.g., “the purpose of this UI element is unclear”) is secondary to evaluating how meaningful, cohesive, and engaging the underlying game system is (cf. Bond 2014, Schell 2019).

makes it more difficult to capture the rich contextual information provided by being physically in the same room as the playtester (e.g., their body movements, facial expressions, etc.). A common technique used in conjunction with observation is **think-aloud**, i.e., inviting playtesters to verbalize their thoughts as they play. While this can provide immediate, “in the moment” insights that players may struggle to recall later, think-aloud works better for turn-based or slower-paced games than those involving fast-paced action, and even then can be challenging for inexperienced playtesters (Pellicone et al., 2022). Playtesting sessions are also often **recorded** (e.g., via screen capture for digital games, video camera for tabletop games, microphone for think-aloud, etc.) for later analysis. An additional form of indirect observation that can be used in playtesting is tracking **in-game analytics** (Denisova et al., 2024): logging quantitative data about specific player behaviors (e.g., the win vs. loss ratio for each character type in a multiplayer game, the most frequent points at which players use the hint system in a puzzle game, etc.).

Following the playthrough, **post-play interviews** and **questionnaires** are often used to gather additional information from the playtesters (Kramarzewski and De Nucci, 2023). Post-play questionnaires often contain a mixture of close-ended (e.g., Likert-style) and open-ended questions. These are more time-efficient than interviews, since facilitators can get the next playtester to play while the previous one is completing the questionnaire, and can reduce courtesy bias (i.e., the tendency to provide more positive feedback for fear of displeasing the facilitator). At the same time, they tend to produce more limited, less in-depth qualitative responses, as people tend to find it easier to provide detailed responses in a conversation than write them down. In addition to providing more in-depth responses, interviews also afford the facilitator an opportunity to clarify and ask follow-up questions.

Biometric data, e.g. from eye tracking, brainwave recording, and heartrate monitoring, is also sometimes used in playtesting (Choi et al. 2016); and approaches from other fields, from A/B testing to computational modeling, may also be integrated into the process; but a discussion of these methods is outside the scope of this paper.

3.3 Types of Playtesting

In addition to relying on different methods and data, playtests can differ in scope, participant type, and structure. The choice between these depends on the goals for the specific playtest.

3.3.1 By scope

Game design literature emphasizes that playtesting is a purposeful process that must be carried out with a specific, clearly defined goal in mind (Schell 2019; Kramarzewski and De Nucci 2023; Denisova et al. 2024). We propose a distinction that the literature often fails to make, but which in our view is useful: between holistic and focused playtests.

Holistic playtesting aims to evaluate the overall game experience. It may be carried out with an exploratory mindset early in the development process, when the designer has an initial prototype and needs to decide whether the concept is viable or which of its aspects to focus on for future development (Choi et al. 2016). Holistic playtesting may also be carried out during the final stages of production, when the game is mostly polished and almost ready for release, in order to identify any remaining (presumably minor) issues (Kramarzewski and De Nucci 2023). Importantly, holistic playtests must still be purposeful: the designer should have a set of assumptions about how gameplay will proceed that the playtest helps validate or prove wrong.

Shawn Patton (2017) proposed the following six reflection (post-play) questions that can be posed to participants in a holistic playtest, using the mnemonic FFWWDD:

- What was *frustrating* about playing the game?
- What was your *favorite* part of the game?
- Was there anything you *wanted* to do but could not?
- If you had a magic *wand* and could change anything about the game, what would it be?
- Can you summarize what you were *doing* in the game?
- How would you *describe* the game to your friends?

The questions above, slightly paraphrased from the original, help the designer identify the core strengths and frustration points of their game, gather ideas for improvement, and validate their assumptions about how players understand the game and see its appeal. (For a more extensive question list that can be used for more in-depth interviews, but requires a greater time commitment from the playtesters, see Fullerton 2018, p. 301.)

In contrast to holistic playtesting, **focused playtesting** has a more narrowly defined scope and aims to evaluate a specific aspect of the game. This can relate to game balancing, the difficulty curve, the control scheme, level design, the player's comprehension of the game message, or any other aspect of game design. For example, when working on a multiplayer role-playing game, many of the playtests might focus on balancing the different characters' abilities, to make sure that each character feels fun to play and does not put the player at a disadvantage compared to other characters (cf. Clark 2026). If a focused playtest involves a post-play questionnaire or interviews, the questions should align with the specific focus. (That being said, a focused playtest will often uncover additional issues that are outside of the main focus. These may be treated as incidental discoveries to consider for the next iteration or serve as a basis for a follow-up playtest with a differently formulated goal.)

3.3.2 By structure

Although Denisova et al. (2024, p.3) describe playtesting as a "formal process of setting research objectives," the degree to which the playtesting process is structured and formalized may vary. We propose to distinguish between informal, semi-formal, and formal playtests.

Informal playtests are opportunistic and do not follow a pre-planned protocol or involve formal data collection beyond observation and possibly unstructured verbal feedback by the testers. The designer is still likely to take notes, but on an impromptu basis. Informal playtests are typically exploratory and holistic in scope and take place in the early stages of the development process (e.g., when validating an initial prototype).

Semi-formal playtests are pre-planned and structured, with observation often followed by post-play interviews or questionnaires. At the same time, they follow an opportunistic approach to player recruitment (e.g., whoever volunteers to play at a public game event) and take an informal approach to data analysis, where the designer mostly relies on their notes from the playtest session.

Formal playtests are rigorously planned, structured, and documented, fully adhering to the standards of empirical research. Playtester recruitment follows clearly defined criteria, and data analysis is rigorous (e.g., interviews are recorded, transcribed, and coded using thematic analysis or another qualitative analysis method). We contend that, even in the context of academic research, it may be practical to use formal playtests strategically (e.g., for final evaluation) and rely on semi-formal playtests for early prototypes and intermediary playtesting.

3.3.3 By participant type

It is often suggested that playtesting must be carried out with testers who are part of the game's target audience (Denisova et al. 2024), as that is the only reliable way to evaluate how the game may be received. While broadly true, this view overlooks the fact that playtesting may be conducted with other demographics for both pragmatic and conceptual reasons. For example, early exploratory playtests often involve the most readily available testers: members of the game development team (which also helps the team align their vision and expectations) or friends and family (Fullerton, 2018). In the context of game-based learning and serious games more broadly, playtesting sessions may involve stakeholders who are outside of the intended player demographic (Bunt, Greeff and Taylor, 2024). These may be, for example, subject matter experts who can provide guidance on the accuracy of the game content (e.g., historians consulting the developers of a history-themed game) or teachers who will be deploying the game in their classrooms.

An argument can also be made for deliberately playtesting *outside* of the target audience, particularly in the early stages of the development process. By doing so, designers can test their assumptions about what their target audience should be, as the game may appeal to other demographics (e.g., a game intended for children may resonate with older players); and even if not, the process may still produce useful feedback.

Building on Fullerton (2018), Table 1 illustrates the different audiences playtests can be conducted with and the main considerations that apply to each audience.

Table 1: Types of playtesting by audience

Type	Usual context	Purpose	Pros	Cons
Internal (within the team)	Informal playtest	Validate prototype/build before testing with other audiences	Good for team alignment Quick to implement	Not reliable: biased, too much familiarity with the game
Friends and family	Informal playtest	Validate basic game concept understanding	Easily available	Heterogeneous group with inconsistent expertise in games Personal bias
Fellow developers	Informal or semi-formal playtest	Invite expert feedback on game system and implementation	Good for new ideas and inspirations Specific, well-articulated suggestions	Professional perceptions are not representative of target group Atomic (tendency to focus on detail over the big picture)
Target audience	Any type of playtest	Validate game system and implementation with intended players	The only way to know whether the game fulfils its design goals	Not reliable for early-stage prototypes Playtesters cannot always articulate their feedback clearly
Stakeholders	Games that rely on expert knowledge or involve a secondary user group (e.g., teachers for a learning game)	Ensure the game reflects stakeholders' expertise and expectations	Can improve content accuracy Can increase stakeholder buy-in (i.e., convince stakeholders of the game's value)	Stakeholders are not always familiar with games Unrealistic demands and expectations that are at odds with game design considerations

4. Playtest Plan

While early exploratory playtesting may be conducted in an impromptu manner, any playtesting sessions beyond that require thoughtful planning if they are to provide useful and reliable data for game designers. To facilitate playtest planning and streamline the reporting process, we propose the following Playtest Plan template (Figure 1) that can be filled out in preparation for a playtest and then used as a guide. (The text in the fields on the right is only provided for illustrative purposes; in the actual template, the fields are blank.) Throughout the development process, a separate plan should be formulated for each round of playtests, based on the state of the project and the designers' needs.

The contents of the Playtest Plan are based on the key considerations outlined by Kramarzewski and De Nucci (2023): focus and questions; format and methods; participant recruitment.

On a conceptual level, the template is intended to help designers specify the main goal of the playtest and translate it into questions to investigate and observation items to keep track of (these may include both recurrent patterns of behavior and quantifiable items, depending on the designer's needs). By limiting the designer to one main goal per round of playtesting, we aim to help them maintain a clear focus, as without it, a playtest may produce an overwhelming amount of observational data that is difficult to make sense of. Even for a holistic, exploratory playtest, a single main goal can be formulated, such as validating that the game concept is viable. We also included fields relating to playtester recruitment, to ensure that the designer clearly considers who they need to test with in order to fulfill their goal and where to find them.

The template also encourages the designer to consider the logistics of the playtest session: how much time is needed given the number of testers and the amount of content to test, what equipment and materials to prepare, and any additional considerations that may be relevant. The "Additional considerations" field was added based on a suggestion made during the focus group interviews that there may be miscellaneous, context-specific things to record that do not naturally fit any of the other fields.

Playtest plan: Game Title

Version	E.g., first prototype; version 0.1
Date	
Venue	

Goal <i>State one main goal</i>	E.g.: validate that the core loop is fun Or: calibrate the difficulty curve, etc.
Main questions <i>Things you want to find out through the playtest (aligned with the goal)</i>	E.g.: 1) Do players find the controls intuitive? 2) Do players enjoy the level of challenge throughout? 3) Does the hint system help reduce player confusion and frustration?
Observation points <i>Things to pay attention to in order to answer the main questions</i>	• E.g., frequency of failure in different levels • E.g., common frustration points in the level design • E.g., frequency and context of the use of hints
Format <i>In-person vs. remote, individual vs. group</i>	E.g., individual, in-person
Methods and data <i>What kind of data you will collect and how</i>	E.g., observation + post-test interview

Recruiting strategy <i>How you will find playtesters</i>	E.g., friends of friends; via the XYZ Discord; at a game event
Recruiting criteria <i>What you are looking for in playtesters</i>	E.g., familiarity with tower defense games, experience with game controllers
No. of playtesters <i>How many testers you aim to recruit</i>	E.g., at least 5

Scope <i>What part of the build will be playtested</i>	E.g., levels 1–5; the paladin campaign; entire game
Est. time per session <i>How long a single session will take (e.g., per person)</i>	E.g., 15 minute playtest + 10 minute interview
Materials needed <i>What to have ready for the playtest</i>	E.g., PC, two controllers, pen and paper, snacks

Additional notes <i>If/as needed</i>	E.g., who will be running the playtest on which day?
--	--

Figure 1: Playtest Plan

5. Playtest Report

The final part of the toolkit is the Playtest Report template. A version of the template is presented in Figure 2; however, the actual template spans multiple pages, with observation points and revision ideas occupying at least one blank page each.

We envisage the template as having two uses. First, it can be used as an observation protocol during the playtest, helping the facilitator document the process. To that end, it restates the main goal, questions, and observation points from the Playtest Plan, so as to help the facilitator maintain their focus on the relevant

aspects. The “Observation notes” field can be left unstructured for a holistic playtest, but for a focused playtest, it may be customized based on the main goal. For example, if the goal is to evaluate the difficulty curve in a linear level-based game, the notes may be taken on a per-level basis, as shown in Table 2. If a round of playtesting spans multiple days or is run in parallel by several team members, the first block of fields at the top of the page can help keep track of which individual playtest session the specific data come from.

Playtest report sheet: <i>Game Title</i>	
Person(s) responsible	
Version	
Date	
Venue	
No. of playtesters	
Goal	(Copy from the Playtest Plan)
Main questions	(Copy from the Playtest Plan)
Observation points	(Copy from the Playtest Plan)
Observation notes	(This can be further broken down into, e.g., levels or tasks. But keep some space for general observations too.)
Revision ideas	
High priority	
Consider for the future	

Figure 2: Playtest Report

When conducting playtesting, ideas about how to address the issues that are coming to light, or about possible new features, will often occur to the designer immediately. Recognizing this (and building on Fullerton 2018), we added the “Revision ideas” section, allowing the designer to keep track of such ideas. As not all ideas are equally important and relevant, the section encourages the designer to divide them into those that are high priority and those that can be considered in the future. (This prioritization can later be reflected upon and discussed with the team.)

Table 2: Example observation notes structure for a difficulty curve-focused playtest of a linear game

Level	Intended difficulty	Actual difficulty	Observations
1	Easy	Easy	No major issues. Many testers skip the introduction text (P1, P2, P4).
2	Easy	Hard	Unclear goal (P1, P3, P4). Difficulty finding the exit (P2, P4).
3	Medium	Trivial	Unexpected shortcut: jump off the first ledge (P1, P2, P5).

The second way the template can be used is to synthesize and present the findings from a round of playtests in an academic text or report. Since the template combines both observation notes and revision ideas, notes made during the playtests can be reviewed and cleaned up, supplemented with data from post-play interviews

or questionnaires, and then presented using the same format. This approach reduces the effort on the game designer's part and is particularly useful in projects that follow a tight schedule and rely on quick iteration.

An alternative format, suggested during both the focus groups and peer review for the article, pairs the observations and the revision ideas, so that it is clear which observed issue (or several issues) each idea addresses. This approach, shown in Table 3, is useful in more advanced stages of the development process, when the main design is locked-in and the focus is on ironing out issues rather than introducing new ideas.

Table 3: Example of observation–revision idea pairing in the Playtest Report

No.	Issue	Priority	Revision idea
1	First weapon upgrade feels underwhelming	Low	Increase pushback, more dramatic visual and sound effect
2	Level goals are unclear	High	Replace “How to play” screen with a playable tutorial
3	Controls are confusing	High	

While the Playtest Report template is well-suited to presenting the results of semi-formal playtests, for formal playtests, it may need to be complemented with more rigorous data analysis techniques (e.g., thematic analysis of interview transcripts), which would need to be documented separately.

6. Discussion and Conclusion

The three parts of the Playtest Planning and Reporting Toolkit have been adopted by students in graduate-level game design courses, and, in our observation, this has helped make their approach to playtesting both more purposeful and more recoverable, i.e., easier to follow for the instructors. The utility of the toolkit was further confirmed during focus group interviews, whose participants reported that they found the materials to be easy to understand and use, as well as recognized the value of employing them. Focus group participants also highlighted that the toolkit is not prescriptive and can be used with different types of playtests without forcing designers to follow a rigid process that may not suit their specific needs.

A point that became clear during the focus groups is that the toolkit may require adjustment depending on the structure of the particular game design program or course, or the needs of the specific project. We would like the materials to be seen as flexible and customizable, as well as modular (for example, the Playtest Plan may be used on its own, with a different format adopted to report the playtest results). Additionally, the toolkit may, and in many cases should, be complemented with other materials such as interview protocols for post-play interviews, questionnaires, etc. A playtest introduction script can be added at the end of the Playtest Plan, as considering how to explain the playtest session to participants is also part of the planning process.

We view the toolkit as a work in progress, which we intend to iterate upon and add to in the future based on the feedback we receive and our own continued reflection. The version presented here, however, is relatively stable and has been successfully used by both students and the authors in practice. It is our hope that the toolkit can help game design researchers, educators, and students make playtest planning and reporting more thoughtful and transparent than the impromptu manner in which it is often done currently; this in turn would contribute to more rigorous game design research and higher-quality games.

Ethics Declaration: no ethnical clearance was required for this study. The workshop and focus groups mentioned in the article were conducted with the informed consent of the participants. The authors have no conflicts of interest to declare.

AI Declaration: no generative AI tools were used when writing this article.

References

- Bunt, L., Greeff, J. and Taylor, E. (2024) “Enhancing Serious Game Design: Expert-Reviewed, Stakeholder-Centered Framework,” *JMIR Serious Games*, Vol 12, No.1, article e48099. <https://doi.org/10.2196/48099>
- Choi, J.O., Forlizzi, J., Christel, M., Moeller, R., Bates, M. and Hammer, J. (2016) “Playtesting with a Purpose,” *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY’16)*, ACM, New York, pp. 254–265. <https://doi.org/10.1145/2967934.2968103>
- Clark, O. (2026) *Playing with Balance: Game Economy Design*. CRC Press, Boca Raton.
- Denisova, A., Bromley, S., Mirza-Babaei, P. and Mekler, E.D. (2024) “Towards Democratisation of Games User Research: Exploring Playtesting Challenges of Indie Video Game Developers,” *Proceedings of the ACM on Human-Computer Interaction*, Vol 8, No. CHI PLAY, article 343. <https://doi.org/10.1145/3677108>

- Elias, G.S., Garfield, R. and Gutsche, K.R. (2020) *Characteristics of Games*. MIT Press, Cambridge, MA.
- Fullerton, T. (2018) *Game Design Workshop: A Playcentric Approach to Creating Innovative Games* (fourth edition), CRC Press, Boca Raton.
- Khaled, R., Lessard, J. and Barr, P. (2018) "Documenting trajectories in design space: a methodology for applied game design research," *Proceedings of the 2018 Foundations of Digital Games Conference (FDG'18)*, ACM, New York. <https://doi.org/10.1145/3235765.3235767>
- Kramarzewski, A. and De Nucci, E. (2023) *Practical Game Design: A Modern and Comprehensive Guide to Video Game Design* (second edition), Packt, Birmingham.
- Meades, A. (2015) *Understanding Counterplay in Video Games*, Routledge, New York.
- Mirza-Babaei, P., Moosajee, N. and Drenikow, B. (2016) "Playtesting for indie studios," *Proceedings of the 20th International Academic Mindtrek Conference*, ACM, New York, pp. 366–374. <https://doi.org/10.1145/2994310.2994364>
- Myers, D. (2020) *Games Are Not: The Difficult and Definitive Guide to What Video Games Are*. Manchester University Press, Manchester.
- Patton, S. (2017) "The Definitive Guide to Playtest Questions" [online]. Available at: <https://sglabs.schellgames.com/post/160052690343/the-definitive-guide-to-playtest-questions>
- Pellicone, A., Weintrop, D., Ketelhut, D.J., Shokeen, E., Cukier, M., Plane, J.D. and Rahimian, F. (2022) "Playing Aloud: Leveraging Game Commentary Culture for Playtesting," *International Journal of Gaming and Computer-Mediated Simulations*, Vol 14, No.1, pp. 1–16. <https://doi.org/10.4018/IJGCM.296705>
- Schell, J. (2019) *The Art of Game Design: A Book of Lenses* (third edition), CRC Press, Boca Raton.
- Tripp, D. (2005) "Action research: a methodological introduction," *Educação e Pesquisa*, Vol 31, No.3, pp. 443–466. <https://doi.org/10.1590/S1517-97022005000300009>
- Zimmerman, E. (2022) *The Rules We Break: Lessons in Play, Thinking, and Design*. Princeton Architectural Press, New York.