

This is a repository copy of *Towards Democratisation of Games User Research: Exploring Playtesting Challenges of Indie Video Game Developers*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/216480/>

Version: Published Version

Article:

Denisova, Alena orcid.org/0000-0002-1497-5808, Bromley, Steve, Mirza-Babaei, Pejman et al. (1 more author) (2024) *Towards Democratisation of Games User Research: Exploring Playtesting Challenges of Indie Video Game Developers*. *Proceedings of the ACM on Human-Computer Interaction*. 343. ISSN: 2573-0142

<https://doi.org/10.1145/3677108>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Towards Democratisation of Games User Research: Exploring Playtesting Challenges of Indie Video Game Developers

ALENA DENISOVA, University of York, UK

STEVE BROMLEY, gamesuserresearch.com, UK

PEJMAN MIRZA-BABAEI, University of Ontario Institute of Technology, Canada

ELISA D. MEKLER, IT University of Copenhagen, Denmark

Playtesting is a games user research (GUR) method used to evaluate design decisions based on feedback gathered from players with the goal to improve player experience. HCI games research has been actively working on and promoting best practices in GUR. However, these practices often require resources, knowledge and expertise, which are not readily available for indie video games developers. Thus, to better understand how GUR can support these developers, we conducted an interview study with 13 indie games professionals to learn about their practices and the challenges they face when doing playtesting. We report on the key findings from this study, including challenges with finding appropriate participants and handling the data from playtests. We provide a discussion of how existing GUR practices can be adapted and what HCI games research can do to help mitigate these challenges to make playtesting more accessible and impactful to indie video games developers.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; • **Applied computing** → **Computer games**.

Additional Key Words and Phrases: Playtesting, player experience, PX, user experience, UX, games user research, GUR, indie, challenges, practices

ACM Reference Format:

Alena Denisova, Steve Bromley, Pejman Mirza-Babaei, and Elisa D. Mekler. 2024. Towards Democratisation of Games User Research: Exploring Playtesting Challenges of Indie Video Game Developers. *Proc. ACM Hum.-Comput. Interact.* 8, CHI PLAY, Article 343 (October 2024), 25 pages. <https://doi.org/10.1145/3677108>

1 Introduction

Recent advances in digital distribution services like Steam and affordable game engines like Unreal and Unity have helped spark a recent renaissance among independent (indie) developers – individuals or small development teams/studios working on games without the financial and technical support of a large game publisher. This has resulted in an upswing in the diversity, quality, and execution of indie games. There are over 45k games tagged as ‘indie’ on Steam alone as we write this paper (~58% of all games on this platform).

The prevalence of indie games on the market demonstrates the importance of delivering good quality, positive player experiences (PX) to survive in the competitive environment. Playtesting is commonly used to evaluate whether the game indeed offers such an experience. This games user research (GUR) approach allows designers to gather feedback from the players to assess whether

Authors’ Contact Information: [Alena Denisova](mailto:alena.denisova@york.ac.uk), alena.denisova@york.ac.uk, University of York, York, UK; [Steve Bromley](mailto:steve@stevebromley.com), steve@stevebromley.com, gamesuserresearch.com, London, UK; [Pejman Mirza-Babaei](mailto:pejman.mirza-babaei@uoit.ca), pejman.mirza-babaei@uoit.ca, University of Ontario Institute of Technology, Oshawa, Canada; [Elisa D. Mekler](mailto:elme@itu.dk), elme@itu.dk, IT University of Copenhagen, Copenhagen, Denmark.



This work is licensed under a [Creative Commons Attribution International 4.0 License](https://creativecommons.org/licenses/by/4.0/).

© 2024 Copyright held by the owner/author(s).

ACM 2573-0142/2024/10-ART343

<https://doi.org/10.1145/3677108>

their game evokes the experience for which it was designed [47]. That said, playtests may not always be feasible or affordable for independent game developers, as they require specialised equipment and expertise.

Given this recent growth and prevalence of indie developers, academic researchers [42–44] have called for supporting indie developers in playtesting by creating new tools and adapting existing GUR processes to meet the tight budgets, strict timelines, and resource limitations of these developers. While these topics have been discussed in workshops and SIGs, indicating significant interest, they primarily reflect anecdotal evidence from academic perspectives and experiences rather than empirical research based on direct feedback from indie developers. Furthermore, much of the existing GUR literature does not clearly distinguish between the playtesting practices of indie developers and larger studios, despite the fact that only the bigger studios typically have dedicated UX teams, making their requirements and capabilities significantly different. Despite these valuable discussions, little research has since gone into understanding the specific challenges indie developers face in playtesting due to their constrained resources. This paper aims to address this gap by directly engaging with indie developers to explore their unique playtesting practices and current struggles.

Hence, to learn more about how HCI games research can better support indie developers, we interviewed 13 participants (see Table 1) in indie PC, console and mobile game development about their current playtesting practices and the challenges they face in their existing processes. The main contribution of our work is the reporting of the challenges that indie game developers face. We also highlight the playtesting practices that are not currently well-understood and well-supported by existing tools, indicating important areas for the future design of a range of possible technical and empirical solutions.

This is an important contribution to the HCI games research as it highlights the need for reducing the inequality between the currently underrepresented groups (small indie studios, freelance developers and non-profit organisations) and those who are able to make financial, logistical and time investments for playtesting efforts (to organise and run playtesting sessions with real players and to form an action plan from the collected data). We discuss the need for democratising GUR for indie developers by offering these teams the impact of research and return on investment (ROI) through GUR training and empowerment of non-research team members to conduct user research.

2 Related Literature

GUR is a prominent interdisciplinary research stream in HCI games, which is concerned with ensuring the optimal quality of usability and user experience (UX) in video games. GUR's primary focus is on evaluating games during development using different user research methods with playtesting being the most widely used approach [20].

2.1 Indie Game Development

Indie game developers constitute a distinctive community within the realm of game development, embodying a multifaceted identity shaped by alternative production and distribution structures outside mainstream gaming corporations [35, 36, 49]. The term 'indie' extends beyond a mere categorisation of hobbyists or amateurs – it refers to a more complex concept encompassing social, artistic, cultural, and ethical dimensions [48]. Within the indie game development landscape, individuals and small studios dissociate themselves from affiliations with large game companies or publishers, opting instead for alternative pathways such as self-funding, self-publishing, collaboration in small teams or studios, and engagement in free labour practices [24].

Indie game developers are recognised for their capacity to push the boundaries of conventional gaming experiences [36] – they often pursue ideas that may be riskier, operate on varying time

scales, and have greater autonomy and self-direction. The creative freedom they wield allows for the production of games that are inherently ‘out there’ – offering players distinct and diverse experiences that deviate from mainstream gaming norms (e.g. the avantgarde games discussed in [14]). This freedom to explore nuanced and novel designs contributes to the richness of the gaming landscape.

From an HCI perspective, studying indie game development brings nuanced insights to the discourse on technology. Existing research on indie developers has predominantly focused on labour conditions and the relationship between cultural production and technology [24, 25] and ethnographic research capturing the cultural nuances and dynamics within the community [32, 53], often highlighting the importance of collaboration and community support for indie developers. However, a notable gap exists in the literature concerning the playtesting practices of indie developers and the challenges associated with it. By empirically examining the processes through which these developers conceive, create, and evaluate their games, researchers gain understanding of how alternative development and evaluation practices contribute to the generation of unique gaming experiences [15]. This exploration not only helps us better understand indie game development, especially around its crucial aspect that significantly influences the final product’s quality and reception, but also serves as a foundation for informing the design of future technologies, with the potential to unlock opportunities for even more novel and diverse gaming encounters.

2.2 Definition of Playtesting

Playtesting itself is often a combination of user research methods (mixed-methods) that are triangulated to answer particular research questions. For example, a playtesting study may use observation and interview methods, where another playtesting study might employ surveys and physiological measures. Understanding the advantages and limitations of each user research method and identifying the best mixture of these methods are often the key tasks of user researchers when deciding on a playtesting study. That said, there is no single definition of playtesting and it has been interpreted differently depending on the development context [41]. In the context of this paper, we define *playtesting* as a formal process of setting research objectives and using insights from players’ opinions and behaviour to identify adjustments necessary to bring the game closer to the experience envisioned by the designer, as well as the players’ expectations (cf. [27, 41, 47]).

2.3 Playtesting Process

To do this, games user researchers (or any other party responsible for the playtest) start by defining precise research questions and objectives for the playtest. Then, they choose appropriate methods to address these objectives, which typically involves gathering behavioural (‘can the player do this’) and attitudinal (‘does the player like this’) player data by means of observations of the player interacting with the game and its components while taking detailed notes, often followed by interviews or questionnaires to gather more subjective data about players and their experience with a game [41]. At the end of a playtesting round, researchers produce a report based on the analysis of the generated data about possible issues related to usability or UX. They then share this report with the developers on the team or the leads who make the decision about what appropriate changes need to be made to the game based on these findings [40]. This summary provides only the indicative steps of the process; for a more detailed breakdown, please refer to [9, 20].

2.4 Playtesting as Indie and AAA Game Developers

Game designers [27, 47], professional PX researchers [9, 20] and academic researchers [3, 23, 30, 41, 42, 44] have described ‘best’ practices for conducting playtests. These studies were often done with larger studios with higher UX research maturity levels, where the uptake of UX research is fairly

good, and typically focused on understanding the methods that developers could use to gather feedback from players, as well as the ways in which this feedback is used to improve the game. While playtesting processes vary from company to company depending on their resources and goals, there are some key factors in offering effective solutions for bringing the game closer to the designer's intent and the player's expectations.

Many major game development studios and publishers have in-house GUR facilities and staff, which facilitate the integration of playtesting into the development cycle. Alternatively, third-party GUR service providers offer contractual services to developers of all sizes. However, small and mid-size studios may not always be in the position of hiring an external agency to do playtesting for them as they try to maintain a strict development budget and avoid unnecessary expenditures, which playtesting might be seen as. This is not only because of the differences in resources and scale between indie and larger game studios, but due to the lack of publicly available research discussing the potential impact of playtesting on a game's quality and ROI [41]. The consensus is that a lack of GUR impacts sales, increases fail rates, and leads to a low persistence on the market – Mirza-Babaei et al. [41] provides several real-world examples illustrating where inadequate playtesting had led to certain problems being overlooked, ultimately contributing to a more negative reception than might have been the case.

2.5 Challenges of Playtesting

Despite the differences in the resources and scale between indie developers and their more mature, juggernaut counterparts, the needs for playtesting and UX optimisation is the same regardless of the studio size. Some indie developers may take on the load of doing playtesting internally. The main challenge here is that playtesting methods are often not as well-understood by developers in non-research roles as effective use of these methods requires experience acquired through practice and/or theoretical research [42]. The lack of relevant research expertise on the team is not the only limitation – limited resources also could mean that the specialised equipment and infrastructure necessary to follow the best practices might also not be in place. As discussed above, these best practices typically involve tips and resources on administering playtesting sessions, selecting from research methods and approaches in communicating the research findings to the development team.

Playtesting is a challenging process that requires careful and purposeful decision-making [13] and effective communication of the playtesting results to the various stakeholders involved in game development (e.g., programmers, designers, producers, or marketing) [44]. Not following appropriate practices can lead to less useful results, ineffective use of time and resources, ultimately impacting the quality of the end product, increasing failure rates, leading to a low persistence on the market [30], as well as contributing to the lack of justification for playtesting ROI in the industry.

Hence, there is an increasing need to understand how to support developers who lack resources, knowledge, and experience to include playtesting in their game design and development processes. GUR offering recommendations specific to indie developers already exists, e.g. [11, 29, 39]. However, there is a marked absence in the literature of direct studies of games developers and their current practices [22, 34] with a notable exception is Mirza-Babaei, et al.'s work on playtesting for indie developers [40, 41]. Their work studied indie playtesting based on the case studies that the research team were involved with and then a postmortem discussion on playtesting based on game critics reviews on the case studies. However, they did not study indie developers' understanding and perspective on playtesting. This is where our paper makes its main contribution. We believe to better support indie developers in their efforts to create better games, we need to learn about what

methods and processes are currently used by them and whether there are any profound challenges that need to be addressed to make playtesting practices more accessible and useful for all.

3 Method

To learn more about the playtesting practices of indie developers as well as the challenges they face related to playtesting of their games, we conducted individual semi-structured interviews with 13 indie game developers. We used the following research questions to guide the conversation with them:

- (1) What research methods and processes do indie game developers use for playtesting?
- (2) What do indie games developers find challenging about playtesting currently?

3.1 Participants

We recruited 13 participants through three primary channels: (i) social media platforms, such as X (formerly Twitter), (ii) personal inquiries, and (iii) game developer communities, including the IGDA Games Research and User Experience SIG Discord channel. The selection of channels (i) and (iii) aimed to cast a wide net, ensuring a diverse participant pool regarding geographic location and professional roles. This strategy was complemented by (ii) purposeful sampling to achieve a balanced representation within our network. The number of participants was deemed sufficient to address our research questions, given the richness of the interview data and the specificity of our participant sample [8, 37].

Our recruitment call was aimed to recruit participants who were presently or previously involved in the playtesting of video games, which were either developed by them as a freelancer (commercially-driven, i.e. have worked on at least one published title) or produced by an independent games studio they are part of. This range of roles and studio sizes was purposeful to understand the dynamics of playtesting in different contexts – because there is no single representation of the ‘playtesting’ role in indie development, we embraced the diversity of the roles that were responsible for performing playtesting. We did not explicitly use the number of titles these developers or studios had previously published or the number of years they had worked in industry as our recruitment criterion since these temporal measures were not believed to be important without outside influence and training.

Our participants held a range of roles at the studios they worked at and several of them wore many hats: alongside their main role, for example, as a designer, they would also be responsible for playtesting their games. None of our participants had formal training in user research; their backgrounds were predominantly in creative or technical fields. The breakdown of the self-described roles of our interviewees is provided in Table 1.

Our participants described their experiences of working mostly in North America, Australia and Europe. Some of these experiences were based on working in more than one company – some interviewees worked in two places at once or recalled their experiences related to UX research when working in other companies before starting working on their own games. Team sizes of the studios ranged from as few as 2 people up to 35 people, with a median number of team members being around 10. They have worked on titles released for PC, mobile, and console platforms. The games they developed were diverse in content and genre, encompassing adventure and narrative games, strategy and simulation, as well as survival and shooter games.

3.2 Procedure

Prior to the interview, each participant was provided with an information sheet and a consent form. The semi-structured interviews began with a brief about the research project, followed by

ID	Roles and Responsibilities	Team size	Location
P1	Independent Developer. As the founding director, they constitute one of the two key team members responsible for the comprehensive development process, encompassing game design, programming, marketing, and various other aspects of game creation. Their portfolio includes the successful release of three games, which underwent in-house playtesting.	Small	UK
P2	QA Manager and Producer, previously Publishing Producer at another studio. They are responsible for overseeing project scheduling, budget management, and comprehensive testing across various projects. Their duties include organising and conducting user playtests. They have successfully contributed to the release of several titles.	Mid-size	Australia
P3	Freelancer. P3 has worked as a consultant and game designer on several published titles for various companies on a flexible basis over two years at the time of the interview. Their responsibilities include game design, production, content integration, balancing, as well as setting up and conducting playtests for the games they have worked on.	Solo	France
P4	Lead game designer, who has worked on several published titles. They were previously a senior game designer at another studio, where playtesting was outsourced to an external user research consultancy. They drew inspiration from this process and leveraged the experience to implement in-house playtesting methodologies in their new role.	Mid-size	UK
P5	Producer. They have previous experience in playtesting for large companies, where they observed and managed playtests which were overseen by a lead researcher. Now, as a producer for an indie team, they handle all aspects of playtesting themselves, from organising and defining studies to observing and analysing results.	Small	Canada
P6	Production Director. They currently run a freelance company and have been in the games industry for about a decade, working on published titles. In their present role, they are responsible for design, development, client interfacing, and playtesting.	Small	USA
P7	Creative Director. They have been working in the games industry for a decade, working on several published titles, starting as a graphic designer and transitioning to a specific focus on user experience. They have been in their current role for 2 years, overseeing three projects.	Mid-size	USA
P8	Community manager. They have been working in the games industry for 2 years and, in this role, they manage the online community of a multiplayer sandbox exploration game developed by the studio they are working at.	Small	USA
P9	Playtesting lead at a company specialising in offering a range of user research services to video games companies. They have been with the company for seven years, starting as a tester, and their main responsibilities include network stress testing with occasional general playtesting. They have contributed to a range of published titles, many of which were created by indie studios.	Mid-size	Global
P10	Producer and Product Manager. In their current role, they oversee over a dozen of legacy titles supported by their studio (not in active production). Playtesting at this studio is conducted by an internal user research team, with whom P10 has established a rapport.	Mid-size	Netherlands
P11	Designer. They have been working in the industry for 5 years, having published a number of titles, and in this role they are working on a game for which they have a range of responsibilities from high concept creation to the iteration and refinement of these concepts and getting onto playtesting.	Small	UK
P12	UX Lead and Freelancer. Working for a UX consultancy with a focus on video games. Prior to that, they were a UX researcher and UI artist at a small indie studio, where they were responsible for the playtesting of their main project.	Small/Solo	UK
P13	Co-Founder of an indie studio. They also serve as the Director of Insights at a video games company, having previously worked as a designer for the same studio. At this large studio, they contributed to several published titles and are responsible for guiding teams across various production stages.	Small	USA

Table 1. Summary of participants’ professional roles and responsibilities, their experience, locations and team sizes based on the description given by the interviewees themselves (where ‘mid-size’ referred typically to a dozen or more employees).

the questions about the participant’s role at the company they work at and their current or past playtesting practices, including questions about what the participants were trying to learn from these playtests and where they have learnt how to playtest this way. This led up to the discussion of the challenges they face related to playtesting. All interview questions are in Appendix A.

All interviews were conducted by the first and second authors using video conferencing software and were manually transcribed by the same person. Interviews lasted between 45 minutes and 1

hour. Participants were offered 40USD for their time. Overall, we collected approximately 58,000 words of content.

3.3 Positionality Statement

Our team consists of academic researchers and industry practitioners with expertise in HCI, games and GUR. Despite our collective depth in GUR, it is essential to acknowledge our distinct positionality. We are not indie developers ourselves, which positions us as observers seeking to understand and refine known and commonly used practices in the context of limited resources. Our primary interest lies in unravelling how indie developers conduct playtesting, delving into their perspectives, values, and evaluation methods. Our inquiry is rooted in bridging the gap between established GUR principles and the lived experiences of those operating on the indie landscape.

Geographically, our team is dispersed across the UK, Denmark, and Canada, reflecting our primary focus on European and North American companies. This regional orientation shapes our interpretation and potentially limits the generalisability of our findings to other global contexts. We recognise the necessity for future work to extend our exploration into the practices of game developers in Asia, Eastern Europe, South America, and Africa. By openly acknowledging these facets of our positionality, we aim to foster transparency and encourage further collaborative research that enriches the understanding of GUR practices across diverse gaming landscapes.

3.4 Thematic Analysis

We chose a codebook approach to Thematic Analysis (TA) [6, 7] to analyse the collected data, as it aligns well with the exploratory nature of the research questions and the complexity of the data provided by indie game developers, helping us achieve a comprehensive and authentic understanding of their playtesting practices and challenges.

The first author conducted inductive TA [4] on the interview dataset, discussing the coding and the development of the themes with the rest of the team. Due to the interpretative nature of the analysis, inter-rater coding was not carried out since it is not considered appropriate for this form of analysis [5].

First, the first author carefully read through all interview transcripts multiple times to familiarise themselves with the content. Second, initial observations and ideas were noted down – they formed the basis for the 52 initial codes. The smallest coding units ranged from a few words to a full sentence, and multiple codes could be assigned to a coding unit. Both the semantic content (i.e., what participants said) and on a latent level (i.e., participants' assumption underlining the semantic content) were used in the creation of the codes. Third, the initial codes were collated and the relationships between them were examined to identify potential themes. Next, these candidate themes were reviewed to refine, redefine and reorganise the themes and sub-themes in successive discussions among the authors. This phase was repeated until we identified the final themes that represented the data. The final stage was defining and naming the themes, which can be found in Table 2. We introduce these themes accompanied by illustrative quotes in the next section. All participants, except for P6 and P10, agreed for their direct quotes to be used in the paper.

In the next section, we begin by outlining our interviewees' views on playtesting to establish the grounds for the subsequent discussion of the challenges. We present these challenges along with the practices articulated by our interviewees as thematic insights to contextualise the difficulties. A thematic map of the themes and specific challenges is provided in Figure 1.

We then discuss these challenges in more depth in section 5, where we highlight the impact of these challenges on the value of playtesting, explore potential strategies to mitigate their effects, and suggest avenues for future work for the HCI community to address these open problems.

Theme Name	Definition	Example Quote
Regular and Early Playtesting as Reassurance	Participants reflect on the role of playtesting at various stages of game development, emphasising the perceived benefits and challenges of conducting early and frequent playtests, as well as the underlying reasons for engaging in or foregoing these activities.	<i>"Playtests just aren't long enough and we weren't able to invest the amount of money to figure out."</i> (P7)
Recruiting Enough Representative and Trusted Players to Gather Objective and Honest Feedback	Participants describe challenges associated with recruiting suitable playtesters, including how they find and balance the recruitment of both expert and novice players to evaluate the usability and player experience of the game at different development stages and how the (lack of) balance of different types of players impact the reliability, validity, and usefulness of the findings.	<i>"I am very interested in ways to connect with new players or don't play our game on getting feedback on features and updates as a new player but of course I don't have that rapport with them [...] so making those connections with other people who we could build a trustful relationship with is difficult."</i> (P8)
Choosing Appropriate Research Methods to Gather Valid and Reliable Data	Participants discuss the challenges associated with selecting suitable research methods for collecting valid and reliable data from playtests and the reasons for choosing certain methods over others. This includes their familiarity with research methods and their understanding of how to use them appropriately, alongside the technical and logistical capabilities and requirements.	<i>"[The challenge is] not having people playtesting something and giving feedback that is honest [and] knowing whether someone is giving feedback just because you've asked them to give feedback – are they just giving feedback because you asked for it or they're trying to find something that's wrong when actually it's not?"</i> (P1)
Interpreting and Analysing Collected Data to Extract Valuable Insights	Participants discuss the challenges associated with reviewing and analysing data to understand player behaviour and experiences, including the process of interpreting large and noisy datasets, which requires careful consideration to distinguish useful insights. They discuss the impact of these challenges on how effectively feedback is implemented into development practices.	<i>"Then on the technology side, even something as simple as recording the screen and the player and the input all at the same time is not something that everyone can do."</i> (P4)
Dealing with Feedback and Prioritising Actions	Participants discuss the challenges related to handling and prioritising actions based on collected data. They highlight the complexities of integrating feedback into development practices and the effects of these decisions on the overall effectiveness of playtesting.	<i>"[...] people have to rush off to their desk and be like 'we need to change this right now' – they're reacting essentially to what they're saying, which isn't representative of a full user."</i> (P2)

Table 2. Themes identified using Thematic Analysis (TA), their descriptions and example quotes.

4 Results

Our interviewees used the term ‘playtesting’ to refer to processes such as quality assurance (QA playtest), ‘internal’ group testing with the colleagues (internal playtest) before showing the game to the ‘real’ players, showcasing the game at events to gather general feedback (demo/event playtest), as well as to engage with the community and for marketing purposes (marketing playtest). While these are valid uses of playtesting, in our paper, we looked at playtesting as strictly a formal process of setting research objectives and using insights from players’ opinions and behaviour to identify adjustments necessary to bring the game closer to the experience envisioned by the designer, as well as the players’ expectations (UX/usability playtesting).

4.1 Regular and Early Playtesting as Reassurance: Creating the Right Experiences for the Right Audience

Our participants recognised the value of playtesting. They believed it offered them reassurance that they are making the right game: *“the only way you can make a game that is playable and understandable and fun for the player, in a way that you’re sure of”* (P3) for the right audience.

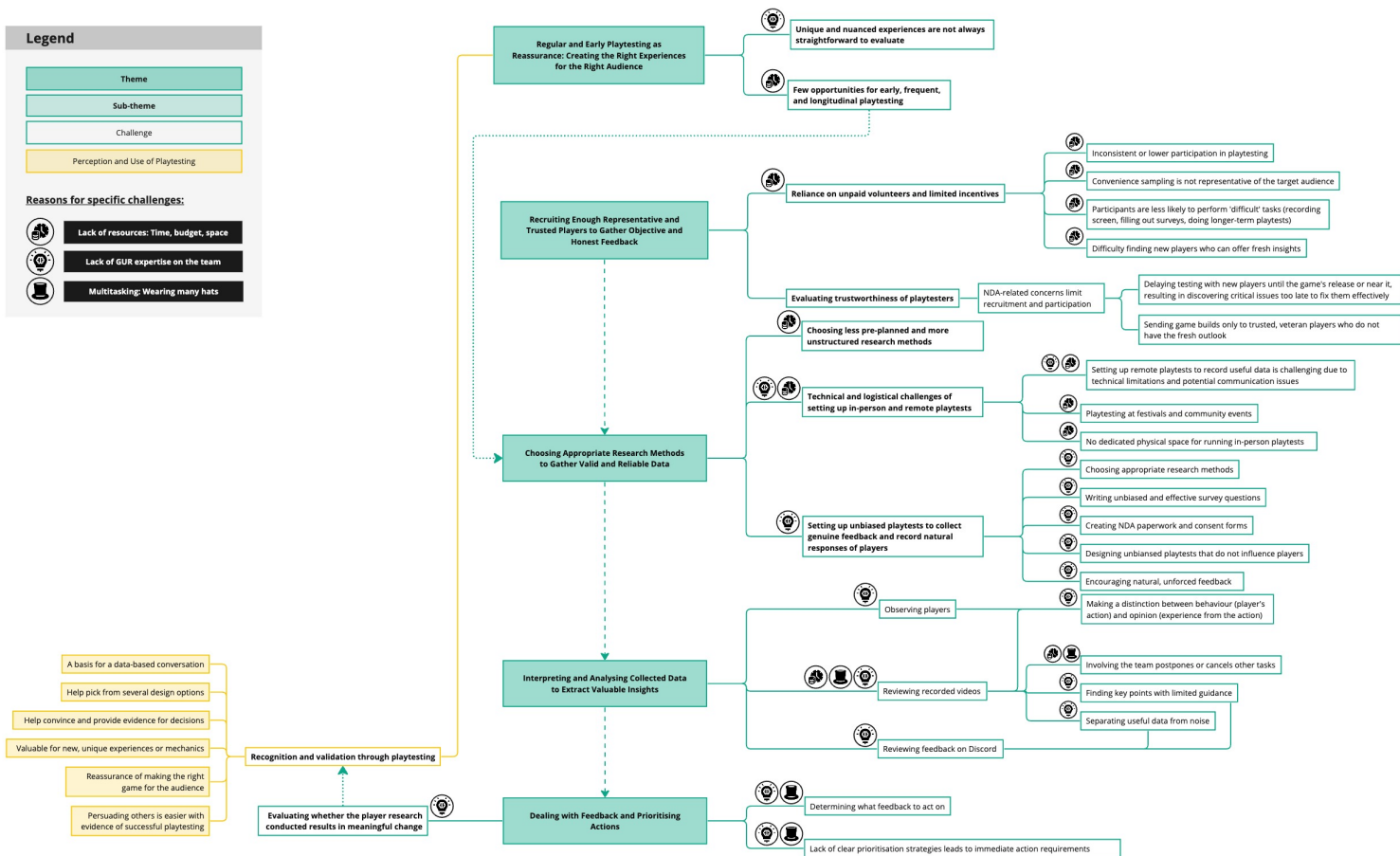


Fig. 1. A diagram representing the five identified themes (in green) and the corresponding sub-themes and the inferred relationships between these. Icons represent the reason for specific challenges identified in our analysis, e.g. a lack of resources, a lack of GUR expertise on the team, and indie developers wearing multiple hats.

For our participants, usability playtesting also served as a form of validation as it can be used as a basis for a data-based conversation. They described it as *“an opportunity to prove or disprove assumptions and hypotheses you have on the game”* (P4) and to help pick from several design options: *“sometimes it will give you new ideas, sometimes it will make you throw away something”* (P3). It could also help convince people when making decisions and provide evidence for decision making, which was particularly valuable when creating new and unique experiences or mechanics without a point of reference for what should be done, which is characteristic of many indie games: *“because the type of game we’re making isn’t kinda out there, playtesting is especially important because we need to know what we’re actually making is fun and use that feedback to figure out what about our prototype and current build that is not fun for different testers and players.”* (P5).

Persuading others of the value of playtesting was perceived as less challenging knowing that there is evidence of others doing it as well: *“it’s validating to at least pitch it to the higher-ups on the team, it’s easy to say to them ‘hey, this has been done before, here’s a point of reference, this is what we can do and it’s gonna be valuable’ and getting that buy-in was definitely an easy process”* (P2), highlighting the importance of evangelism for this field.

Playtesting was also seen an opportunity to engage with the community, as well as being used for marketing purposes: to encourage players to wishlist the game on Steam or similar platforms, get people to sign up to their mailing list and to encourage streaming of the game: *“the content creators [e.g. Youtube and Twitch streamers] do have something exciting to share when they get to stream early or do that video then they can share with their community that is maybe an exclusive from them”* (P8).

Our interviewees noted the importance of early playtesting, e.g. *“when you’re this early in development in the prototype phase this is when playtesting has the biggest impact”* (P5). But despite this, there had not been many opportunities to do early or frequent, more formal, playtests. This decision was not an easy one, as often they did not have the autonomy and UX maturity to make this decision: *“your boss is not gonna pay you to just watch people play for some reasons”* (P3). This often meant that they had to make an explicit decision not to engage in playtesting with ‘real’ players until later stages in the development when making substantial changes to the game would be challenging, if not impossible: *“by that point, we were in beta, so there was nothing we could do for the most part ... All we could do is bug test”* (P5).

Without a dedicated UX researcher on the team, the responsibility of setting up a playtest and analysing the data was placed on the team members. This often meant that other jobs they would have otherwise been doing, would be neglected: *“you’re gonna have to be there during the day of the playtest, that means you’re not gonna be doing design during this time”* (P3). Other tasks got prioritised, often at the cost of doing fewer playtests: *“Without trying to bring more people into the team and adding more cost to the project, we’re [...] trying to ship it under budget”* (P2).

4.2 Recruiting Enough Representative and Trusted Players to Gather Objective and Honest Feedback

Our interviewees described two broad categories of playtesters they gathered feedback from at different stages in the development: ‘internal’ and ‘external’. The ‘internal’ playtests were done at earlier phases and did not typically include ‘real’ players. As discussed earlier, this highlights that “playtesting” is a term that is used in different contexts in game development. While these ‘internal’ playtests do not fit in GUR’s definition of playtesting, developers may see them as a way to gather feedback before the ‘polished’ version can be shown to the ‘real’ players in ‘external’ playtests.

Our interviewees noted that gathering *“objective, relevant feedback, to find the right people”* (P2) for their playtests was challenging, because of the *“[limited] technology and resources”* (P4). Specifically, they had trouble finding new players representative of their target audience, had

challenges associated with the lack of incentives to give to playtesters as well as NDA-related issues, which are discussed in more detail below.

Most of our interviewees relied on volunteers to playtest their games. Limited budget meant that there was no or little compensation to give to the participants, which often resulted in inconsistent or lower participation in playtesting, adding to the time resource. This often also meant that playtesters were less likely to be asked to do ‘difficult’ things, like recording their screen or filling out a survey, or do longer term playtests: *“there are some [questions] that we still can’t answer because the playtests just aren’t long enough and we weren’t able to invest the amount of money to figure out, because services that do features if you want to test those, that is like days and days and days and hours of playing and we weren’t ready to put in that type of investment”* (P7), which limited method choice.

In absence of a dedicated budget for recruiting playtesters, interviewees had to often rely on convenience sampling for finding participants in the early stages of the development. Friends and family members, as well as colleagues within the studio or other developers would take part in informal, ‘internal’ playtesting. While ‘internal’ playtesters provided with useful initial feedback, our interviewees understood that these opinions were not necessarily *“representative of full players”* (P2) and were not objective because of the pre-existing relationship: *“it’s not gonna be the most objective thing, because first of all she’s my girlfriend so she’s gonna say ‘you’re doing a great job, congratulations!’ ”* (P3) or because of their previous exposure to the game: *“once they have played the game once or if they are someone who has developed a feature, obviously they know so you won’t be getting any valuable data.”* (P3). Hence, to gather more objective feedback, our interviewees turned to ‘external’ playtesters to *“sanity check that with outside public”* (P2).

Signing an NDA form before the playtests is often part of the playtesting routine as leaks are a major concern for games studios, particularly for remote playtesting sessions. This meant that their recruitment decisions were often based on trust and it encouraged teams to use panels of trusted players even when new players would have been more appropriate. Some interviewees mentioned screening their playtesters based on resumes and questions (P6). Asking players to sign NDA often meant that the interviewees had to *“account for the drop off caused by [NDA requirements...]”* (P5).

Nonetheless, our interviewees often aimed to include a diverse range of experiences and opinions. And so they *“would screen the participants beforehand [...] and make sure that you have a population diversity that you want with various interests”* (P9). And to evaluate players’ first-time user experience (FTUE), some interviewees tried *“to find real naive people that you don’t really know”* (P3).

However, reaching out to new players was particularly difficult: *“I am very interested in ways to connect with new players or don’t play our game on getting feedback on features and updates as a new player but of course I don’t have that rapport with them, I don’t have their email address, so making those connections with other people who we could build a trustful relationship with is difficult”* (P8).

This balance between trust and getting fresh range of opinions was hard to strike: *“you need a lot of trust and you need to know that they are going to take the thing seriously, you have to know them. But everyone you know has played the game for ages, so how do I find a group who have never played it but that I do trust will take it seriously and will respect our NDA”* (P8). For some, this meant delaying testing with new players until the game’s release or near it.

4.3 Choosing Appropriate Research Methods to Gather Valid and Reliable Data

Our interviewees aimed to collect both behavioural (does the player understand what they need to do) and attitudinal (what is their gaming experience) data in their playtests using a mixture of approaches, including some formal research methods like surveys, data analytics and observations. Discord was mentioned as the platform used for connecting with the community, as well as collecting *“general feedback”* (P8) and having an informal discussion of gaming experiences.

The challenges associated with the research method choice ranged from technical difficulties of setting up playtests online and in-person to having the experience and knowledge (and often the resources) to choose appropriate methods for the identified research objectives.

Tight budget and looming deadlines meant that the interviewees often looked for research methods that were less pre-planned and more unstructured: *“it’s the more basic stuff which is kind of where we skew towards. And always with our time between now and actually getting early access, we definitely focus more on easier playtesting sessions and just making sure we’re getting some data, rather than getting the data when you release the game build”* (P2).

The developers we spoke with did not have a dedicated physical space for running playtests. Thus, most tests with ‘real’ players were conducted remotely or at events or local gaming communities, which affected the choice of research methods and likely compromised playtest validity. For instance, ‘internal’ playtesting with colleagues was done at their desks: *“people, in my experience, behave differently and playing differently if they are at their normal working desk than they do if they’re in like a specific space [...] With that kind of space issue, you don’t necessarily have an option to observe anonymously as well, there is no one-way mirror that you hide behind – you’re always lurking behind someone”* (P4). Playtesting at events, while enticing for its potential outreach, has notable limitations. The lack of ecological validity and potential participant bias can negatively affect data reliability. The awareness of being observed may induce confirmation bias, impacting the authenticity of insights gathered.

In recent years, remote playtesting have become even more popular, especially considering that it offers more flexibility in setting up and running tests than its in-person alternative. Interviewees noted that remotely they could gather large numbers of diverse participants relatively quickly and cost-effectively. Nonetheless, one interviewee noted that setting up a remote session that produces useful data could be technically challenging: *“On the technology side, even something as simple as recording the screen and the player and the input all at the same time is not something that everyone can do.”* (P4). The challenge was in getting the build to the player and being confident it will work on their kit, as well as having the confidence in them being able to record their screen and send it back to the developer, which requires a level of technical understanding.

Interviewees with previous experience of in-person playtests at festivals and in local games community meetups expressed their preference for in-person interactions over remote sessions because *“you can’t see their eyes, you can’t see where their hands are going and that sort of thing you can get from in-person tests”* (P7) and you *“can’t have a conversation necessarily with someone while they’re playing, only when we’re doing stuff at events”* (P1). This extra information was deemed important as it offered more context for the potential problems: *“[without] seeing what buttons they are pressing, you don’t get the insight whether someone is pressing the wrong button combination or we don’t know if they are trying to run up this ramp or they’ve just stood still because we don’t have that additional information”* (P4).

It was challenging *“knowing what type of questions to ask based on what my team was needing”* (P12) to be able to gather insights that are useful for the team, especially regarding writing survey questions, NDA paperwork and consent forms. This was typically done ad-hoc: *“we did have a session where we’re just kind of coming up with questions for the form”* (P2), often by ‘googling’ terms like ‘user feedback questionnaires’ (P2) and *“how do you ask this kind of question? Or do I ask an open question in the form without an answer? For what kind of question can I do that?”* (P3). Some interviewees participated in other companies’ playtests to learn about what goes into a playtesting session and to gauge the type of questions they ask in their surveys: *“was looking at other playtesting questionnaires from other indie studios. [...] I’ll participate in a playtest, I’ll see the questions that they wrote and then I’ll put those into my notes”* (P5). However, merely taking part in other studios’ playtests did not offer the insights into the decision making and the thought process

that goes into setting up one of these sessions: *“the large companies have [GUR] professionals [who] know what they’re doing, [...] I did not see how they did that, I only observed the end result”* (P5).

Another challenge around collecting valid data from players was knowing how to design playtests in a way that does not influence the players: *“Something I do think is best practice for playtesting is not to tell your playtesters what you are looking for, it is important to give them as clear a slate as possible because it means that the reactions you get from them will be as genuine as possible”* (P11) and *“as pure of an experience as possible [...], to be natural”* (P7). Similarly, one participant was interested in knowing more about how to write a brief that is indeed brief and ‘unbiased’: *“We just try to be subtle about it and not overly difficult about what we ask, but sort of guide them towards what we want them to do, but also not. So I think that’s probably the hardest part for us”* (P7).

It was also deemed important to be able to ask for feedback that comes out naturally and their players are not being forced to just say something. This sometimes meant that no specific instructions would be given to the players: *“we’ll just say ‘here’s our new stuff, see what you think?’”* (P7), *“[The developers] didn’t really know the right questions to ask. So they had a build, they stuck it in front of a bunch of people in a room, they got general ‘Yeah, looks pretty good’, which was useless feedback for them. They didn’t know what to do with that feedback [...] It was hugely expensive, hugely underwhelming in terms of value that they got from it.”* (P13). Such an unstructured approach runs the risk of introducing noise into the gathered data – the lack of a specific focus may lead players to discuss topics unrelated to the intended research objectives.

4.4 Interpreting and Analysing Collected Data to Extract Valuable Insights

Observation studies were commonly used to learn more about player behaviour and sometimes as means of evaluating player experience. These were conducted either in-person or online (for example, via zoom), and often recording these sessions for future reference or asking for videos directly from the players: *“[our players] will send videos of what they’ve done [via] Discord”* (P1). Knowing what to look out for in an observation was, however, considered challenging – making a distinction between behaviour (player’s action) and their opinion (experience from the action) during observations was something that was *“hard to persuade people to do”* (P7).

Observation videos of people playing the game recorded either by the players or the moderators were typically watched by everyone in the studio: *“we’ll give people about a week to look at all of the videos and then we’ll get together on zoom to debrief. So everyone gets together with the notes that they took and we’ll go around the table listing off our observations”* (P7). *“I’d say 80% of what you’re gonna get from your data is not relevant [...] For most of them, it won’t be useful, we’re just gathering it just in case”* (P3). Despite that, no specific guidance was given to those who did the analysis of the videos: *“just watch it and kind of analyse it, I guess, but very loosely [...] we’d watch a bunch and see if there were common things, so we’d be like ‘ah everyone’s making that mistake’ so then that tells us we need to change this”* (P1), making it more challenging to find points of interest that needed to be actioned.

Dealing with informal feedback on Discord resulted in a similar challenge of looking for patterns in the data, which also needs to be aggregated manually: *“it is difficult to scroll through and read 2-3 days’ worth of Discord messages. It’s not hard, but it is very time-consuming. And it would be cool to have a scale or a tool or something like that”* (P8). Separating useful data from noise was a challenging task.

4.5 Dealing with Feedback and Prioritising Actions

Once the data from surveys and observations have been analysed, the observations would typically be compiled into a report for internal use. At this stage, it was challenging to know how to handle the collected data – what to ignore and what to action. In some cases, the report was sent to the

'higher-ups', e.g. the product managers or the studio director, who "*mak[e] informed decisions*" (P2) on "*what we should prioritise or if there is a problem we should deal with soon*" (P7). After that, "*the designer team will go through and read what everyone has said and adjust accordingly*" (P8).

However, in smaller indie studios, this decision about "*how we fine-tune things and adjust the design in response to the things that they have said*" (P8) would often be left for the team to handle: "*we then decide on what we're going to actually act on, or what's valued feedback, what's something that we've identified but it's definitely not a priority for right now. How do we want to take it?*" (P2) and "*if there's an issue that comes up, we would then ask everyone like 'Oh what does everyone think about this?' and then they would often start discussing with each other about like 'Yeah, no, I don't think this thing is overpowered' or like... and then there'd be a discussion.*" (P1).

Without a clear prioritisation strategy, most recent feedback based on one's observations, survey results or comments on Discord would result in immediate action: "*there is inherent fear going like 'oh my God, somebody said this, we can simplify this and make it clear'*" (P2). This was prominent in 'internal' playtesting sessions within the studio: "*sometimes people have to rush off to their desk and be like 'we need to change this right now'... they're reacting essentially to what they're saying, which isn't representative of a full user*" (P2). This could be particularly burdensome when every member on the team has many other development tasks that require immediate attention.

Acting upon feedback to "*provide more meaningful impact*" (P13) and being able to incorporate research insights into the design was also seen as a challenge for someone without formal GUR training and experience: "*respecting the design enough to think, okay, I understand that you need this information to make this decision and I know it doesn't seem that this feedback is relevant, but let me explain to you why it is*" (P12).

Two interviewees raised concerns about the lack of clarity in knowing or evaluating whether the player research they have conducted had much impact or resulted in a meaningful change: "*We don't really know if we're doing it right because it's situational, it might not be right, we're just winging it, we don't really know what we're doing, but we are trying our best*" (P7).

5 Addressing challenges and advancing playtesting practices for indie developers

Our interviews mark a pioneering effort towards documenting the challenges and needs of indie game developers in preparing and conducting playtests, directly capturing their unique perspectives and experiences. Notably, these open problems have not been comprehensively documented before, especially from the viewpoint of a representative sample of this community.

Our findings suggest that there are certain implicit challenges (i.e., identified in our analysis), as well as needs and barriers explicitly noted by our interviewees that require further investigation and attention from the HCI community.

Many of these challenges were rooted in the precarious labour conditions – a subject of focus for notable HCI scholars, such as Freeman et al. [24], Keogh [32], Whitson et al. [53]. Specifically, the lack of infrastructure, time, and funding significantly hampers researchers' ability to conduct studies with the same scale and rigour as their larger counterparts. Beyond resource constraints, these challenges were compounded by a lack of research expertise within the teams.

In our discussion, we first explore the impact of these limitations on the quality and feasibility of playtesting. Subsequently, we propose further steps to help make playtesting more accessible and impactful to these developers, which can in turn extend the research agenda and find new avenues for future academic work in the field of games HCI. The proposed tasks encompass both technical and empirical approaches to streamline specific aspects of playtesting and deepen the understanding of playtesting practices among indie developers. Additionally, we advocate for systemic changes, emphasising their importance within the HCI games community to foster collaboration in addressing the identified challenges.

5.1 Empirical and technological suggestions for addressing playtesting challenges

5.1.1 The shift towards online playtesting. Our findings clearly highlight that indie developers understand the conceptual importance of getting their games into players' hands to gather feedback and make improvements, even if they lack a deep understanding of the scientific methodologies behind effective playtesting.

Traditionally, indie developers would rely on expos like PAX, GDC, where they could showcase their games, observe player interactions, and gather feedback directly from attendees. This approach, though not scientifically rigorous, provided valuable insights that could guide the development process.

However, the landscape of playtesting has shifted (started by the disruptions caused by COVID-19 pandemic). Developers have adapted by turning to online platforms to conduct private or closed beta tests. These tests involve distributing game demos through channels like Discord, Twitch, or Reddit, and inviting players to provide feedback in more controlled environments. This 'playtesting' style has become a practical solution for indie developers who often have limited resources, as also noted by our interviewees.

The shift to online beta testing presents both challenges and opportunities. While some developers may lack systematic approaches to collect and analyse feedback, relying instead on informal methods like reading forum posts or Discord messages, others have begun to implement more advanced techniques such as analytics tracking within their games. This evolution in playtesting practices highlights a significant area where academic knowledge (such as our paper) and professional expertise can contribute. By developing tools, frameworks, and best practices for online beta testing, the academic community can help indie developers refine their feedback collection processes, making them more structured and effective.

The growing prevalence of closed beta tests among indie developers underscores a trend towards more rigorous playtesting methods, even in resource-constrained environments. As this practice becomes more common, there is a clear need for collaboration between developers and academics to enhance the validity and utility of the feedback gathered, ultimately leading to better game development outcomes.

5.1.2 Dual Purpose of Playtesting: Enhancing Game Quality and Validating Market Potential. Playtesting has also evolved to serve not only as a tool for game development but also as a critical method for business validation, especially for indie developers. This dual purpose of playtesting encompasses both improving the game through feedback and gauging market interest and engagement. By conducting private beta tests and collecting analytics, developers can measure how many people are interested in their game and how engaged they are during the testing phase.

For instance, indie developers might post about their game on platforms like Reddit to attract beta testers. By tracking how many people sign up and participate, developers can determine the initial level of interest. During the beta testing period, they can analyse player behaviour – such as playtime, number of play sessions, and in-game actions – to understand player engagement. This data not only helps in refining the game's mechanics and user interface but also provides insights into the game's potential market success.

By integrating playtesting with business strategies, indie developers can validate their games' commercial potential. This approach combines creativity with business acumen, allowing developers to refine their games based on player feedback while simultaneously building evidence of market demand. This dual focus on game improvement and market validation can attract more attention from publishers and investors, enhancing the chances of commercial success.

5.1.3 Understanding the limitations and potential biases in recruitment practices. Targeting and recruiting appropriate playtesters is crucial for obtaining relevant insights. Considering that the most common approach for finding playtesters of our participants was through recruitment calls to their community, it is important to address the impact of sampling bias. Hence, more research is needed to understand the positive or negative impact of informal practices, such as interviews with personal contacts and unofficial discussions [46]. Useful insights still can be gathered from informal playtesting, but in those cases behavioural data from this audience should be prioritised over opinions.

A diverse participant pool consisting of both new and experienced players is vital for conducting effective playtests. Our interviewees noted, however, that balancing player expertise and trustworthiness poses a challenge, especially for early prototypes. While commercial solutions such as *PlaytestCloud* and *User Interviews* enable developers to find and screen players, not all developers can afford to use these services.

Addressing challenges related to participant recruitment involves exploring alternative recruitment strategies and non-monetary incentives, such as community recognition and exclusive access to build a dedicated community around the game. Alternatively, simulating diverse player profiles during playtesting is an avenue worth exploring, offering insights into how a game caters to different preferences and demographics, contingent on a prior understanding of the target audience. Further research is, however, needed to help with building such profiles.

5.1.4 Exploring new tools and research methods for evaluating new and unique player experiences. Indie games present a distinctive challenge due to the often niche, nuanced, and unique player experiences they offer. Understanding the goals of playtesting requires a deeper comprehension of the specific experiences that indie developers aim to create. In addressing this challenge, HCI researchers can contribute by conceptualising such experiences and designing methods and tools to facilitate their evaluation.

Real-time feedback can offer valuable insights into players' moment-to-moment experiences. Integrating random self-report requests within gameplay by indie developers can capture effective states and player experience data. However, placing these requests without disrupting game immersion requires a nuanced understanding of player interaction. One approach could be to track players' "downtime" using biometric data (e.g. [45]) and presenting questionnaires during those moments to evaluate their present experience (e.g. see [26]). However, for this approach to be effective, further research is necessary to explore how to integrate PX questionnaires into games more organically.

Related to that is a growing recognition of the need for a diverse range of user research methods that extend beyond conventional observation and interviews. Exploring alternative approaches to assess early prototypes, including aspects like narrative elements, does not only enhance the evaluation process, but also positions itself as a persuasive tool for garnering support from budget holders and decision-makers.

5.1.5 Learning to ask effective and unbiased questions to gather honest and reliable data. Effective playtesting requires the use of appropriate methods to gather reliable and useful data to make reasonable assumptions about the necessary design improvements. Clear research objectives defined at the start of the process are crucial for selecting the right data collection and analysis methods. However, without the relevant training, exacerbated by a scarcity of available resources for guidance, this could be a challenging task. Our interviewees struggled with formulating effective, unbiased questions for interviews and designing surveys, hindering the gathering of useful and reliable data. They also expressed the desire for organic, honest player feedback, avoiding forced responses while still eliciting constructive insights. Striking a balance between providing guidance for meaningful

exploration and avoiding excessive, overly detailed instruction was, however, difficult. The absence of a centralised learning platform left participants relying on fellow developers or limited online resources, making it difficult to understand the decision-making process behind method selection or question phrasing.

The overarching challenge of balancing mixed-method studies with practical decisions faced by game developers could be effectively addressed through a co-created framework, a recommender system, or a digital or physical tool (e.g. ideation cards). Such tools could guide developers systematically through the decision-making process, assisting in defining objectives and identifying effective methods aligned with their goals. Some existing commercial solutions for the outlined challenges already are on the market, for instance, *The Playtest Kit* [10] is a checklist for setting up and running a playtest that was designed based on the findings presented in this paper.

Supplementary, more specialised solutions could help with more nuanced tasks, like finding an appropriate questionnaire or a survey for evaluating specific gaming experiences, and help analyse the collected survey data to draw meaningful conclusions. Work is underway to develop an interactive online tool to help browse validated and commonly used questionnaires in GUR, e.g. [1, 18, 31], using a user-centred approach. As discussed in Denisova et al. [15], nuanced experiences emerging from indie games deserve special attention and can be measured using established, validated scales. Hence, the goal of such a tool is to allow its users to learn more about the different questionnaires measuring a range of experiences to make an informed choice about the experience(s) one might want to measure and how they can do it. Related, a tool that can assist in defining clear research objectives, perhaps drawing from the expertise of industry professionals and GUR academics, could be valuable.

5.1.6 Handling large amounts of unstructured data to extract relevant insights. Observations, an affordable and remote-friendly research method, were widely used by our participants, who manually reviewed gameplay videos, noting observations along the way – a process that proved time-intensive and challenging for novice observers like our interviewees. Similarly, informal approaches to gathering feedback, whether it is examining Discord comments or going through gameplay footage, without defined playtest objectives can result in excessive quantities of superfluous data, which can be time-consuming to analyse, particularly if there are no specialised tools to help with the analysis. It would also make it difficult to extract relevant information leading to actionable points. Combining methods like data analytics and observations can be effective for pinpointing relevant information without the need to watch the entire video. Tools based on academic research [38, 52] can help address this challenge by combining in-game metrics and affective data with the video footage effectively and visualising the results of this combined data. More automated tools, however, are needed to assist junior user researchers with this process by providing helpful clues to the observer based on the specified objectives.

Some participants noted that a combination of eye gaze, controller interactions, and emotional responses of players in addition to the screen recording and other in-game metrics allowed them to paint the most accurate picture of what was happening in the game. However, there were some technical challenges in setting up remote sessions that record all this data, including sending the build to the player and making it work on their kit, as well as relying on the willingness to perform this with little to no financial incentives and the level of technical understanding of players who are able to screen record and send the footage back to the developer(s). Recent academic research suggested tools for cost-effective and time-saving data collection and analysis suitable for remote playtesting. Facial expressions [51], physiological data (e.g. heart rate [45, 50]) as well as gaze, emotions and movement [21, 33] can be similarly captured alongside the video footage. Further

solutions, e.g. around sentiment analysis and identifying recurring themes, could be co-designed to improve the ease of doing remote playtesting by overcoming the aforementioned challenges.

5.1.7 Improving accessibility of indie games. Despite being a key part of user research, accessibility was not mentioned by any of our participants. Recent work by Beeston et al. [2] has highlighted the need for more research into accessible player experiences (APX), which prompted the creation of the APX design cards [12] that aim to help game designers offer access for players with disabilities to be able to enjoy a range of challenges offered by their games. Similarly, we call for further exploration of what accessibility issues are most prominent in the creation of indie games (considering the wide range of mechanics and in-game factors) and how accessibility can be improved to allow players enjoy indie games regardless of their (dis)abilities.

5.1.8 Dealing with Feedback and Prioritising Actions. Addressing feedback and prioritising actions emerged as challenging tasks for our interviewees. Effective communication of usability and UX issues within the team is crucial to prevent important details from being lost ‘in translation’. Once identified, these issues must be organised by priority, with critical concerns taking precedence over less urgent matters, such as cosmetic issues. Co-designed solutions from additional research could enhance the efficiency of this process and provide guidance on prioritising feedback and the scale and timeliness of decision-making. Community involvement becomes pivotal in this context – encouraging designers to share decision processes and reflect on outcomes. This collaborative approach, drawing on diverse perspectives and strategies, generates collective knowledge valuable for assessing the impact of specific design decisions on game reception and ROI. However, further research can help us develop a more comprehensive understanding of how to prioritise actions based on factors like issue frequency, its type, usability, player experience and the impact on game reception. Crowd-sourced information, combined with existing research findings (e.g., [16, 17, 19]), may offer insights that inform design choices, potentially reducing the necessity for extensive early playtesting, especially considering challenges related to NDA restrictions and limited resources.

5.2 Democratising user research to empower indie developers to conduct research themselves

The need for GUR is growing as it plays a crucial role in building games that players enjoy. Past research has highlighted that many studios, particularly small to mid-sized companies, either do not have a research team or the research team cannot handle the demand [44]. The need for ‘maturing’ and ‘scaling’ research within the studio and the field has been emphasised and deserving more attention. Previous research suggested that companies need to allocate more research resources and ‘scale’ their research personnel. However, ‘scaling’ by hiring more researchers is not always feasible due to various reasons, such as the cost and the need to prioritise hiring developers to build games. One way to address this unequal access to user research is through democratisation to support the team to scale the impact of research by empowering and educating non-research team members to conduct research themselves and providing them with the necessary tools to get access to players’ insights.

Our results highlight that indie developers need user research and the research is often conducted by non-researchers. Therefore, empowering and training non-researcher team members to conduct GUR could have prevented some of the issues we discussed in this paper and would lead to better quality insights. Here, we would like to highlight a few action items to address this need.

First is the opportunity to build on a culture of learning in game studios towards supporting team members to learn about user research. Game studios already operate at the frontier of new technologies and development practices. These often include utilising new hardware and developing for different gaming platforms, introducing new business models and so on. Hence,

game development naturally attracts and retains workforce that has the desire for continuous learning. Moreover, innovation in games is often seen as an indicator of a studio's success. We see an opportunity to build on this desire to encourage and support non-researchers to learn more about conducting GUR. Within game studios, this could include a planned sustainable framework for learning, for example, covering the cost of courses and certifications, rewarding people who take the initiative to learn, providing access to appropriate learning resources, and specific ways to apply the new skills and receive feedback from experienced mentors while respecting team members' autonomy and agency. This would also be a potential answer to the need voiced by our interviewees to evangelise user research: a top-down incentive would better communicate how UX research may add value, from improving the quality of their titles to helping team members better align on their KPIs. Established researchers from both the industry and academia could support this by providing guidelines and in mentoring capacities. The mentoring program offered by the IGDA's GRUX-SIG is a notable example.

There are also various risks that we feel are important to highlight here. One major risk for studios without established research practices and maturity is the lack of processes and rigour necessary to ensure that the research results are valid and actionable (as highlighted in Chapter 5 of [20]). Creating and sharing research guidelines is another area that academic researchers can contribute. Another concern particularly in companies with existing research team(s) might be that researchers' impact and power would be reduced if non-researchers are trained to conduct research studies. We argue that educating and empowering others to be able to do research could be also seen as a researcher's main duty. Disseminating GUR expertise would have a more substantial impact on making better products, as more people are involved in conducting research and are aware of the impact of GUR on ROI and player engagement. In this case, it is also important to decide which research projects should be conducted by the research team (e.g., complex foundational projects) and which projects can be done by non-researchers (e.g., projects suitable for formative, co-design, and competitor evaluation). Finally, it is important to highlight the need for supporting resources, as discussed earlier; these include access to appropriate research tools, step-by-step guides, templates and knowledge banks.

5.3 Further Reflections and Future Work

The overarching long-term goal is, therefore, to advance GUR through learning about practices and challenges of indie developers and determining best practices for this community. Recognising constraints as potential drivers of innovation, indie developers may have devised innovative approaches that hold promise for broader adoption in industry, presenting interesting avenues for academics to explore in more depth. A large-scale survey of these practices could potentially capture some of the lesser known or used approaches. Future work could also consider how these indie practices compare to the established practices from larger studios.

Co-created innovative technological and methodological solutions could help address some specific challenges and needs their practicality and effectiveness in specific contexts. This would also establish a feedback loop that would allow for the refinement of these solutions and effective dissemination to the game industry, research communities, and the educational sector.

Further ethnographic research is necessary to explore current practices of indie developers 'in the wild' and 'in situ' to help researchers address the 'real' rather than the 'projected' misconceptions and challenges [28] and create novel, applicable, effective, and widely available solutions to address these. This offers an opportunity for longer term involvement for both parties, addressing the need to create better relationships and dialogue between academics and indie game developers doing playtesting.

5.4 Limitations

This study has a few limitations. First, we acknowledge that our self-selecting sample consists of indie developers based in Europe, Australia and North America, with a predominant majority being native English speakers. Our interviewees have worked on published titles before and so the sample did not include those who treat game development as primarily for artistic expression, experimentation, or even student indies. This sample might not fully capture the practices and challenges encountered by the entire global population of indie game developers. Second, there could be a bias towards indie game developers who actively engaged on social media during our recruitment period. Future studies could employ large-scale surveys to capture a wider variety of experiences from different cultures, games and personal backgrounds to further validate the interview findings.

It is also important to note that the participant selection along with any analysis we have conducted is subject to the inherent biases of the researchers involved. As noted in the positionality statement, three members of our research team are GUR academics and two authors work as GUR researchers in industry. Thus, our analysis of the playtesting practices and challenges was approached from academic research perspective and also based on the commonly used, well-established practices in the games industry. We acknowledge that some prevalent practices are rooted in broader UX and HCI research and GUR-specific knowledge is largely derived from the experiences of larger companies and GUR consultancies, thus are not necessarily tailored for or developed with the indie community in mind. We, therefore, hope that future work will consider the needs and constraints of indie developers, aiming to make playtesting practices more accessible and inclusive for those who are presently overlooked and underrepresented.

6 Conclusions

Playtesting is an important part of the game creation process, which aims to assist developers in achieving intended experiences in their games and identifying and resolving potential problems during development. This method is commonly used in the video games industry, but indie developers do not always have the necessary expertise on the team or the specialised equipment to follow the practices as promoted by GUR. Hence, to understand what indie developers find challenging about playtesting, we conducted and analysed 13 interviews using TA. The prevalent challenges were organised into five overarching themes: (1) Regular and early playtesting as reassurance: creating the right experiences for the right audience; (2) Recruiting enough representative and trusted players to gather objective and honest feedback; (3) Choosing appropriate research methods to gather valid and reliable data; (4) Interpreting and analysing collected data to extract valuable insights; and (5) Dealing with feedback and prioritising actions. Based on these findings, we highlight several key tasks for the HCI community that extend existing research through further investigation of the playtesting practices of indie game developers and their contextual influences and call for the development of methods and lightweight tools tailored to the unique challenges faced by these developers.

Acknowledgments

We would like to extend our sincere gratitude to all the participants for their time and for sharing their valuable thoughts and experiences with us. We would also like to thank the reviewers and acknowledge the extensive review process this paper underwent – this experience has highlighted the diverse and often conflicting opinions within our community regarding the significance, sensitivities, and knowledge surrounding playtesting and indie developers. This divergence of views

highlights the complexity and importance of this topic, which we believe requires further collective attention from the community.

A Semi-Structured Interview Guide (Indicative)

The interview guide was designed for approximately 60 minutes and divided into five segments: Background/Warming Up, Understanding and Use of Playtesting, Learning about Playtesting, Playtesting Challenges, and Wrap-Up. Each segment was defined with primary questions (denoted by Latin numerals) and sub-questions (denoted by bullet points) for follow-up. The interviewer also posed additional ad-hoc questions if the participant introduced topics of interest that warranted further exploration.

The interview began with an introduction where the interviewer, welcomed the participant and explained the purpose of the study on playtesting practices and challenges. The participant was asked to review the consent form to ensure they understood and agreed to the terms. Participants were given the option to consent to the use of direct quotes from their interview and to disclose the company they work for if they felt comfortable doing so. The interviewer emphasised the voluntary nature of the interview and informed the participant that the conversation would be recorded solely for transcription purposes. The participant was assured that any topics they wished to discuss off the record could be accommodated. After addressing any initial questions, the recording of the interview commenced.

Background/Warming up (5 min)

The first part of the interview focused on the participant's background. The interviewer asked about the participant's role at their company and the specific responsibilities their position entailed to provide context for understanding their experience and perspective on playtesting:

- (1) What is your role at «company»?
 - What are the usual tasks you do in your day-to-day job?

Understanding and Use of Playtesting (20-25 min)

The discussion then shifted to the topic of playtesting. The participant was invited to define what playtesting means to them. They were asked to describe their past experiences with playtesting in previous projects, including the methods they used and how these experiences unfolded. The interviewer probed further into the objectives of playtesting, seeking to understand what the participant aimed to learn from these sessions and why this information was considered important.

- (1) In your own words, describe what 'playtesting' is?
- (2) Have you done much playtesting in your past projects?
- (3) Can you tell me about how you playtested these games?
 - How did that go?
- (4) What are you trying to learn from playtests?
 - Why is that important to learn?

Learning about Playtesting (10-15 min)

The conversation then explored the rationale behind their current playtesting methods and whether they had experimented with other approaches. The interviewer inquired about the outcomes of any alternative methods they had tried. The participant was also asked about their sources of knowledge on playtesting. This included how they learned to conduct playtests, the sources or tools they considered, and how they discovered these resources. The ease of finding these sources and their overall evaluation of them were also discussed.

- (1) Why are you playtesting the way that you are? Have you tried other ways?
 - How did that go?
- (2) Where did you learn how to run playtests like that?
 - What sources or tools have you considered?
 - How did you come across these sources or tools?
 - How easy was it to find these sources or tools?
 - How effective were these sources or tools in aiding your playtesting process?

Playtesting Challenges (10-15 min)

The participant described the difficulties they currently face and the extent to which these challenges are frustrating. The interviewer sought to understand whether these issues were seen as major problems or minor inconveniences. The participant also discussed any perceived limitations and the strategies they employed to overcome these barriers. To conclude, the interviewer asked what changes the participant would make to facilitate easier playtesting if given the opportunity.

- (1) What's difficult about playtesting currently?
 - How frustrating is that difficulty?
 - Is it a big problem, or not a big deal?
- (2) What kind of limitations do you think there may be to the approaches you have used?
 - Do you do anything to overcome these limitations?
- (3) If you could change one thing to make playtesting easier, what would it be?

Wrap-Up (5 min)

The interview concluded with expressions of gratitude for the participant's time and contributions. The session ended after addressing any remaining questions from the participant.

- (1) Is there anything else that you would like to add or discuss?

Details for a \$40 reimbursement were also requested at this stage, or alternatively, a donation to a charity of the participant's choice was offered. The participant was given the option to review the interview transcript.

References

- [1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135 (2020), 102370. <https://doi.org/10.1016/j.ijhcs.2019.102370>
- [2] Jen Beeston, Christopher Power, Paul Cairns, and Mark Barlet. 2018. Accessible Player Experiences (APX): The Players. In *Computers Helping People with Special Needs*, Klaus Miesenberger and Georgios Kouroupetroglou (Eds.). Springer International Publishing, Cham, 245–253. https://doi.org/10.1007/978-3-319-94277-3_40
- [3] Regina Bernhaupt. 2010. *Evaluating User Experience in Games: Concepts and Methods*. Springer London, USA. <https://doi.org/10.1007/978-1-84882-963-3>
- [4] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- [5] Virginia Braun and Victoria Clarke. 2016. (Mis)conceptualising themes, thematic analysis, and other problems with Fugard and Potts' (2015) sample-size tool for thematic analysis. *International Journal of Social Research Methodology* 19, 6 (2016), 739–743. <https://doi.org/10.1080/13645579.2016.1195588>
- [6] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. <https://doi.org/10.1080/2159676X.2019.1628806>
- [7] Virginia Braun and Victoria Clarke. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and psychotherapy research* 21, 1 (2021), 37–47. <https://doi.org/10.1002/capr.12360>

- [8] Virginia Braun, Victoria Clarke, and Nikki Hayfield. 2019. ‘A starting point for your journey, not a map’: Nikki Hayfield in conversation with Virginia Braun and Victoria Clarke about thematic analysis. *Qualitative Research in Psychology* (2019), 1–22. <https://doi.org/10.1080/14780887.2019.1670765>
- [9] Steve Bromley. 2021. *How to be a games user researcher*. Amazon, UK.
- [10] Steve Bromley. 2022. The Playtest Kit. <https://playtestkit.com/>
- [11] Lysiane Charest. 2018. Punching Above Your Weight: How small studios can leverage data for an unfair advantage. In *Games User Research*. Oxford University Press, 357–380. <https://doi.org/10.1093/oso/9780198794844.003.0020>
- [12] The AbleGamers Charity. 2018. Accessible Player Experiences (APX). <https://accessible.games/accessible-player-experiences/>
- [13] Judith Oden Choi, Jodi Forlizzi, Michael Christel, Rachel Moeller, MacKenzie Bates, and Jessica Hammer. 2016. Playtesting with a Purpose. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA) (*CHI PLAY '16*). Association for Computing Machinery, New York, NY, USA, 254–265. <https://doi.org/10.1145/2967934.2968103>
- [14] Tom Cole, Paul Cairns, and Marco Gillies. 2015. Emotional and Functional Challenge in Core and Avant-garde Games. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (*CHI PLAY '15*). Association for Computing Machinery, New York, NY, USA, 121–126. <https://doi.org/10.1145/2793107.2793147>
- [15] Alena Denisova, Julia Ayumi Bopp, Thuy Duong Nguyen, and Elisa D Mekler. 2021. “Whatever the Emotional Experience, It’s Up to Them”: Insights from Designers of Emotionally Impactful Games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (*CHI '21*). Association for Computing Machinery, New York, NY, USA, Article 120, 9 pages. <https://doi.org/10.1145/3411764.3445286>
- [16] Alena Denisova and Paul Cairns. 2015. Adaptation in Digital Games: The Effect of Challenge Adjustment on Player Performance and Experience. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (London, United Kingdom) (*CHI PLAY '15*). Association for Computing Machinery, New York, NY, USA, 97–101. <https://doi.org/10.1145/2793107.2793141>
- [17] Alena Denisova and Paul Cairns. 2015. First Person vs. Third Person Perspective in Digital Games: Do Player Preferences Affect Immersion?. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (*CHI '15*). Association for Computing Machinery, New York, NY, USA, 145–148. <https://doi.org/10.1145/2702123.2702256>
- [18] Alena Denisova, Paul Cairns, Christian Guckelsberger, and David Zendle. 2020. Measuring perceived challenge in digital games: Development & validation of the challenge originating from recent gameplay interaction scale (CORGIS). *International Journal of Human-Computer Studies* 137 (2020), 102383. <https://doi.org/10.1016/j.ijhcs.2019.102383>
- [19] Alena Denisova and Elliott Cook. 2019. Power-Ups in Digital Games: The Rewarding Effect of Phantom Game Elements on Player Experience. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) (*CHI PLAY '19*). Association for Computing Machinery, New York, NY, USA, 161–168. <https://doi.org/10.1145/3311350.3347173>
- [20] Anders Drachen, Pejman Mirza-Babaei, and Lennart Nacke. 2018. *Games User Research*. Oxford University Press, Inc., USA.
- [21] Brandon Drenikow and Pejman Mirza-Babaei. 2017. Vixen: interactive visualization of gameplay experiences. In *Proceedings of the 12th International Conference on the Foundations of Digital Games* (Hyannis, Massachusetts) (*FDG '17*). Association for Computing Machinery, New York, NY, USA, Article 3, 10 pages. <https://doi.org/10.1145/3102071.3102089>
- [22] Henrik Engström. 2020. *Game Development Research* (1 ed.). University of Skövde, Sweden. 230 pages.
- [23] Remigius Fierley and Stephan Engl. 2010. User Experience Methods and Games: Lessons Learned. In *Proceedings of the 24th BCS Interaction Specialist Group Conference* (Dundee, United Kingdom) (*BCS '10*). BCS Learning & Development Ltd., Swindon, GBR, 204–210. <https://doi.org/10.14236/ewic/HCI2010.26>
- [24] Guo Freeman, Jeffrey Bardzell, Shaowen Bardzell, and Nathan McNeese. 2020. Mitigating Exploitation: Indie Game Developers’ Reconfigurations of Labor in Technology. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–23.
- [25] Guo Freeman, Nathan McNeese, Jeffrey Bardzell, and Shaowen Bardzell. 2020. “Pro-Amateur”-Driven Technological Innovation: Participation and Challenges in Indie Game Development. 4, GROUP, Article 04 (01 2020), 22 pages. <https://doi.org/10.1145/3375184>
- [26] Julian Frommel, Katja Rogers, Julia Brich, Daniel Besserer, Leonard Bradatsch, Isabel Ortinau, Ramona Schabenberger, Valentin Riemer, Claudia Schrader, and Michael Weber. 2015. Integrated Questionnaires: Maintaining Presence in Game Environments for Self-Reported Data Acquisition. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (London, United Kingdom) (*CHI PLAY '15*). Association for Computing Machinery, New York, NY, USA, 359–368. <https://doi.org/10.1145/2793107.2793130>
- [27] Tracy Fullerton, Christopher Swain, and Steven Hoffman. 2004. *Game Design Workshop: Designing, Prototyping, and Playtesting Games*. CMP Books, USA.

- [28] Colin M. Gray, Erik Stolterman, and Martin A. Siegel. 2014. Reprioritizing the Relationship between HCI Research and Practice: Bubble-up and Trickle-down Effects. In *Proceedings of the 2014 Conference on Designing Interactive Systems* (Vancouver, BC, Canada) (*DIS '14*). Association for Computing Machinery, New York, NY, USA, 725–734. <https://doi.org/10.1145/2598510.2598595>
- [29] Julien Huguenin. 2018. Running user tests with limited resources and experience. In *Games User Research*. Oxford University Press, 417 – 430. <https://doi.org/10.1093/oso/9780198794844.003.0024>
- [30] Katherine Isbister and Noah Schaffer. 2008. *Game usability: Advancing the player experience*. CRC press, USA.
- [31] Charlene Jennett, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. 2008. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies* 66, 9 (2008), 641–661. <https://doi.org/10.1016/j.ijhcs.2008.04.004>
- [32] Brendan Keogh. 2023. *The videogame industry does not exist: Why we should think beyond commercial game production*. The MIT Press.
- [33] Daniel Kepplinger, Günter Wallner, Simone Kriglstein, and Michael Lankes. 2020. *See, Feel, Move: Player Behaviour Analysis through Combined Visualization of Gaze, Emotions, and Movement*. Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376401>
- [34] Boriana Koleva, Peter Tolmie, Patrick Brundell, Steve Benford, and Stefan Rennick Egglestone. 2015. From Front-End to Back-End and Everything In-Between: Work Practice in Game Development. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (London, United Kingdom) (*CHI PLAY '15*). Association for Computing Machinery, New York, NY, USA, 141–150. <https://doi.org/10.1145/2793107.2793131>
- [35] Nadav Lipkin. 2013. Examining Indie's Independence: The meaning of "Indie" Games, the politics of production, and mainstream cooptation. *Loading...* 7, 11 (2013).
- [36] Sebastian Long, Alena Denisova, and Pejman Mirza-Babaei. 2024. From Pixels to Play: Opportunities and Challenges of a Diverse and Democratized Games Industry. *Interactions* 31, 2 (February 2024), 54–58. <https://doi.org/10.1145/3647646>
- [37] Kirsti Malterud, Volkert Dirk Siersma, and Ann Dorrit Guassora. 2016. Sample Size in Qualitative Interview Studies: Guided by Information Power. *Qualitative Health Research* 26, 13 (2016), 1753–1760. <https://doi.org/10.1177/1049732315617444>
- [38] David Melhart, Antonios Liapis, and Georgios N. Yannakakis. 2019. PAGAN: Video Affect Annotation Made Easy. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. 130–136. <https://doi.org/10.1109/ACII.2019.8925434>
- [39] Pejman Mirza-Babaei and Thomas Galati. 2018. *Affordable and data-driven user research for indie studios*. Oxford University Press, 381–392.
- [40] Pejman Mirza-Babaei, Naeem Moosajee, and Brandon Drenikow. 2016. Playtesting for Indie Studios. In *Proceedings of the 20th International Academic Mindtrek Conference* (Tampere, Finland) (*AcademicMindtrek '16*). Association for Computing Machinery, New York, NY, USA, 366–374. <https://doi.org/10.1145/2994310.2994364>
- [41] Pejman Mirza-Babaei, Samantha Stahlke, Günter Wallner, and Atiya Nova. 2020. *A Postmortem on Playtesting: Exploring the Impact of Playtesting on the Critical Reception of Video Games*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3313831.3376831>
- [42] Naeem Moosajee and Pejman Mirza-Babaei. 2016. Games User Research (GUR) for Indie Studios. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI EA '16*). Association for Computing Machinery, New York, NY, USA, 3159–3165. <https://doi.org/10.1145/2851581.2892408>
- [43] Lennart E. Nacke, Pejman Mirza-Babaei, Katta Spiel, Zachary O. Touns, and Katherine Isbister. 2018. Games and Play SIG: Engaging Small Developer Communities. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI EA '18*). Association for Computing Machinery, New York, NY, USA, 1–4. <https://doi.org/10.1145/3170427.3185360>
- [44] Lennart E. Nacke, Christiane Moser, Anders Drachen, Pejman Mirza-Babaei, Andrea Abney, and Zhu (Cole) Zhenyu. 2016. Lightweight Games User Research for Indies and Non-Profit Organizations. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI EA '16*). Association for Computing Machinery, New York, NY, USA, 3597–3603. <https://doi.org/10.1145/2851581.2856504>
- [45] Xiaolan Peng, Xurong Xie, Jin Huang, Chutian Jiang, Haonian Wang, Alena Denisova, Hui Chen, Feng Tian, and Hongan Wang. 2023. ChallengeDetect: Investigating the Potential of Detecting In-Game Challenge Experience from Physiological Measures (*CHI '23*). Association for Computing Machinery, New York, NY, USA, Article 587, 29 pages. <https://doi.org/10.1145/3544548.3581232>
- [46] Jullia Saad, Suéllen Martinelli, Leticia S. Machado, Cleidson R.B. de Souza, Alexandre Alvaro, and Luciana Zaina. 2021. UX work in software startups: A thematic analysis of the literature. *Information and Software Technology* 140 (2021), 106688. <https://doi.org/10.1016/j.infsof.2021.106688>
- [47] Jesse Schell. 2008. *The Art of Game Design: A Book of Lenses*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

- [48] Bart Simon. 2013. Indie Eh? Some Kind of Game Studies. *Loading...* 7, 11 (2013).
- [49] Alexander Styhre and Alexander Styhre. 2020. Who Is an Indie Developer? Sorting Out the Categories. *Indie Video Game Development Work: Innovation in the Creative Economy* (2020), 107–128.
- [50] Philipp Sykownik, Katharina Emmerich, Jochen Peketz, and Maic Masuch. 2019. Blending Science and Practice: A Collaborative Approach for Evaluating the Value of Heart Rate Measurement. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (Barcelona, Spain). Association for Computing Machinery, New York, NY, USA, 211–222. <https://doi.org/10.1145/3341215.3354644>
- [51] Chek Tien Tan, Daniel Rosser, Sander Bakkes, and Yusuf Pisan. 2012. A Feasibility Study in Using Facial Expressions Analysis to Evaluate Player Experiences. In *Proceedings of The 8th Australasian Conference on Interactive Entertainment: Playing the System* (Auckland, New Zealand) (IE '12). Association for Computing Machinery, New York, NY, USA, Article 5, 10 pages. <https://doi.org/10.1145/2336727.2336732>
- [52] Günter Wallner, Nour Halabi, and Pejman Mirza-Babaei. 2019. *Aggregated Visualization of Playtesting Data*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3290605.3300593>
- [53] Jennifer R Whitson, Bart Simon, and Felan Parker. 2021. The Missing Producer: Rethinking indie cultural production in terms of entrepreneurship, relational labour, and sustainability. *European Journal of Cultural Studies* 24, 2 (2021), 606–627. <https://doi.org/10.1177/1367549418810082>

Received February 2024; revised June 2024; accepted July 2024