IADIS International Journal on WWW/Internet
Vol. 23, No. 2, pp. 33-49
ISSN: 1645-7641

AI-GENERATED VIDEO AS A PROTOTYPING LANGUAGE FOR GAME DESIGN COMMUNICATION: INSIGHTS FROM THERAPEUTIC GAME DEVELOPMENT USING OPENAI'S SORA

JaeJun Lee[1], Soyoun Eom[1], Joo-Eun Kang[1], Anderson Sungmin Yoon[2], Hayoung Oh[3] and Min Ah Kim[2*]

[1]Department of Game Design, School of Art, SungKyunKwan University. Seoul, Korea
[2]Department of Social Welfare, Sungkyunkwan University, Seoul, Republic of Korea
[3]Department of Applied Artificial Intelligence, Sungkyunkwan University, Seoul, Republic of Korea

ABSTRACT

Effective interdisciplinary collaboration is increasingly important in game development for therapeutic applications, yet traditional text-based design documents often fail to convey nuanced play experiences to nongaming experts. This study introduces a three-stage AI-generated video-prototyping pipeline, combining OpenAI's ChatGPT and SORA, as a communication medium in the development of Mind Trekking, a therapeutic game grounded in mindfulness-based cognitive therapy. The pipeline enabled game designers, psychologists, and medical professionals to align gameplay mechanics iteratively with therapeutic objectives, visualize alternative design directions, and expand ideation beyond initial concepts. A fidelity assessment using written semistructured questions with domain experts demonstrated that video prototypes enhanced shared understanding and reduced misinterpretation compared to conventional design artifacts. Although limitations included prompt sensitivity and restricted fine-grained control over generated outputs, the findings suggest that AI-generated video prototyping can function as a prototyping language for interdisciplinary collaboration, particularly in specialized domains such as therapeutic game design.

KEYWORDS

Video prototyping, Generative AI, game design communication, therapeutic game, Mindfulness-Based Cognitive Therapy, SORA

[*]Corresponding author: Professor Min Ah Kim - *minahkim@skku.edu*

1. INTRODUCTION

Game development involves collaboration among various stakeholders, including designers, programmers, artists, domain experts, and project managers, from the initial design phase to the final product. During this process, a shared understanding of the game among project participants is crucial.

Our interdisciplinary research team engaged in a project titled “Development of Game-Based Digital Therapeutics Technology for Adolescent Mental Health Management” to create a game called Mind Trekking. With the support of the Korea Creative Content Agency, we collaborated with more stakeholders compared to typical game development as an interdisciplinary research team. The project aimed to develop a therapeutic game grounded in evidence-based practice with mental health professionals, in collaboration with interdisciplinary professionals from various school departments such as social work and applied artificial intelligence (AI). Additionally, Samsung Medical Center was in charge of conducting clinical trials in collaboration with medical device manufacturers. All game design and development processes were overseen and supervised by game design professionals from a department of game design.

The game was designed based on mindfulness-based cognitive therapy (MBCT), a therapeutic program that integrates mindfulness practices with cognitive behavioral therapy to prevent depressive relapse. Mindfulness, rooted in Buddhist meditation, emphasizes present-moment awareness without judgment and fosters “decentering”—the ability to observe thoughts and emotions without identifying with them (Kabat-Zinn, 1994; Baer et al., 2006; Segal et al., 2013). A precise understanding of the game by mental health professionals was significant for successful game production. However, communicating the planned play experience using game design documents including text, storyboards, and concept images proved to be extremely challenging for the interdisciplinary research team, given domain professionals might have little to no experience with or understanding of games. Moreover, due to the uniqueness of the research team affiliated with a university, the intensive development environment typical of professional game development companies was limited.

Consequently, the environment required minimizing communication iterations by using high-resolution and high-fidelity prototypes. The prototyping method with the highest fidelity is working prototyping, which provides the highest level of understanding through direct interaction, surpassing explanations or gameplay observations (Buchenau et al., 2000). However, this approach is expensive, slow, and difficult to modify. This study focused on a prototyping method we adopted in this context. Our research team noted the potential of generative AI in generating videos; thus, we tested video-based prototypes to convey desired game play experiences. The team’s goals included building a video-based prototype that combined generated videos that richly demonstrated the game experience for communication with nonexperts in games along with a visual novel engine to enable interaction, achieving a balance between fidelity and efficiency. The final playable prototype version is downloadable at the URL found in the Appendix A.

2. LITERATURE REVIEW

Despite the crucial role of prototyping in computer game design (Fullerton et al, 2006; Fullerton, Tracy, 2008), there is no specific methodology for prototyping (Manker, 2011). Research on the role of prototyping varies, including its use for communication (Otto & Wood, 2001; Camburn et al., 2017; Israel et al., 2016; Lauff et al., 2018; Ulrich & Eppinger, 2016), enabling exploration of game ideas (Lee et al., 2023), confirming feasibility (Otto & Wood, 2001; Menold et al, 2017), and establishing milestones (Otto & Wood, 2001; Ulrich & Eppinger, 2016). Game designers generally share the opinion that prototypes serve as sketches, communication of functionality, and idea exploration (Manker, 2011). Prototypes act as a language in the game development process and become the center of communicating the game experience (Manker, 2011).

Research on prototyping largely has considered the fidelity of prototypes and the efficiency of prototype production. Storyboarded game scenes are clear and efficient for explaining sequences, but they have low fidelity in describing interaction experiences due to their static nature. Paper prototyping is effective and fast (Snyder, 2003), but the facilitator can interfere with the player's experience (Bolchini et al., 2009). Interactive document methods that do not require a human facilitator have been studied (Signer & Norrie, 2007). To increase fidelity and recreate the gaming experience in combination with devices, papers are photographed and placed on devices (Bolchini et al., 2009), with animations used to create prototypes to explain situations arising from interactions (Kultima et al., 2012). For efficiency, advancements in AI and game production technology have led to the use of generative AI to create storyboards (Lee et al., 2023) and automatically generate prototypes for specific genres (Reyno & Carsí Cubel, 2009).

Recent advancements in video and world-generation technologies are further transforming the landscape of prototyping. Tools such as Google's Veo 3 enable the creation of high-resolution, general-domain videos from textual descriptions (Veo3, DeepMind), whereas DeepMind's Genie 3 can generate controllable, prompt-driven world environments that allow navigation and interaction over extended periods (Genie3, DeepMind). These developments suggest a shift from static or prerendered representations toward dynamic, explorable prototypes, offering unprecedented fidelity in conveying gameplay experiences. As these technologies become more sophisticated, they are not only streamlining asset creation but also opening new possibilities for simulating interactive spaces without extensive manual production.

Generative AI technology enables Text to Image and Image to Video generation, with OpenAI's SORA capable of Text to Video (SORA, OpenAI). To disseminate our research better, we selected SORA, the most widely accessible video generator, so that anyone could apply the flow we developed. The reason for choosing SORA was its popularity, ease of access, and the ability to replicate our process without specialized or proprietary tools. These technologies are used to innovate game design, efficiently produce game visual images (Lee & Lee, 2023), design characters (Ling et al., 2024), and manage game services (Riedl & Zook, 2013), thereby changing the game production process. Therefore, our research team looked into enhancing the efficiency of showcasing game experiences, employing generative AI to generate video-based prototypes.

3. VIDEO-PROTOTYPING PIPELINE

The Video Prototyping Pipeline using Generative AI consists of three main stages (Figure 1), each of which plays a distinct role in transforming abstract design concepts into tangible interactive experiences.

First, the process begins with a theory-driven game design and documentation stage. At this step, the core framework of the game is established, grounded in relevant theories. For instance, in the case of therapeutic games, theoretical constructs such as MBCT inform the narrative structure, gameplay mechanics, and target user experience. Designers and researchers collaborate to articulate these ideas in written design documents, flowcharts, and concept sketches. This documentation ensures that the objectives of the game—whether clinical or entertainment oriented—are explicitly stated and can be later translated into prototype form.

Second, the pipeline moves to the prompt engineering and content generation stage, during which textual prompts are carefully crafted to instruct generative AI tools. These prompts need to balance precision and flexibility, specifying elements such as characters, settings, mood, and visual style while leaving room for creative variation. Generative AI models, such as OpenAI's SORA, are then employed to produce still images or short video clips that align with the design documents. Because prompt outputs can vary significantly, iterative refinement is often required: Multiple prompt versions are tested, adjusted, and run until the generated video adequately reflects the intended game and visual design. This stage thus is a critical bridge between abstract design intentions and concrete video and image representations.

Third, the integration and simulation stage takes the generated assets and incorporates them into a visual novel engine. Here, the AI-generated videos are not viewed passively but rather embedded in branching storylines that allow players to make choices at critical points. This interactivity enables stakeholders such as psychologists and medical professionals to experience the prototype as if they were actual players. By simulating decision-making moments, designers can evaluate whether the intended emotions, therapeutic effects, or narrative progressions are communicated effectively. Furthermore, this stage allows for rapid iteration: Changes in narrative flow or decision points can be tested without full-scale programming or asset production.

In combination, these three stages transform theoretical concepts into experiential prototypes, providing a novel "language" through which diverse stakeholders can evaluate, critique, and refine game designs before committing to full-scale development.

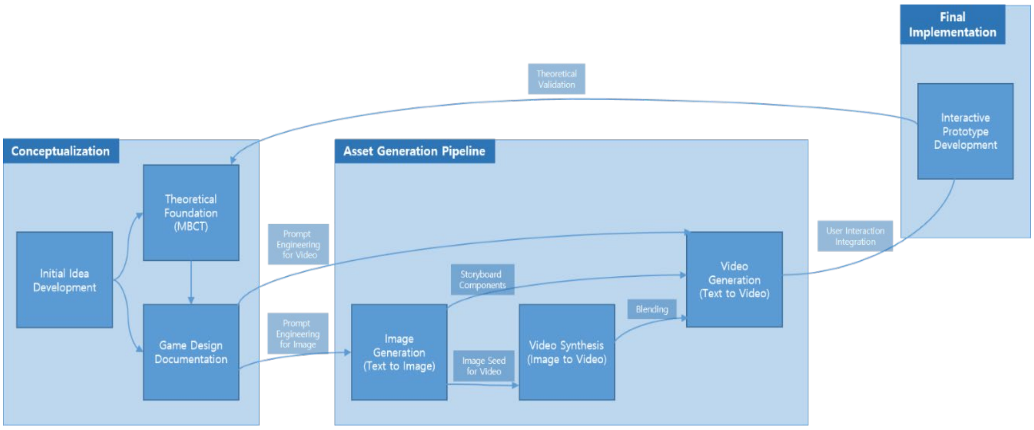

Figure 1. Pipeline Process

3.1 Conceptualization

In the first stage of the process, MBCT theory and its mechanisms of action, as derived from the *Mindfulness-Based Cognitive Therapy for Life* (Bernard et al., 2020), were organized into eight episodes based on offline MBCT sessions in collaboration with mental health professionals guided by evidence-based practice and theory (Feldman and Kuyken, 2019). The game design was developed to ensure that the offline MBCT mechanisms could also be effective in a digital environment (Table 1).

Table 1. MBCT sessions and corresponding video game design episodes

Episode	MBCT session	Video Game Design
1	Awareness and automatic pilot	An experience of focusing one's senses on birdsong amid various natural sounds.
2	Living in our heads	An experience of perceiving and understanding the sensations of body parts displayed on the screen through one's body.
3	Gathering the scattered mind	An experience of walking along a mountain trail, appreciating the surroundings, and inputting emotional responses.
4	Recognizing aversion	While struggling on a difficult hike, participants are shown a bus, which they discover is actually on a return loop and does not ease their journey.
5	Allowing or letting be	During the walk, participants encounter a very noisy group of hikers but have the opportunity to show kindness to the group's dog.
6	Thought are not facts	A scene is staged in which a squirrel, encountered during the walk, becomes angry when its acorn is taken away, only to realize that the acorn was rotten and feel relieved.
7	How can I best take care of myself?	An exercise in distinguishing between positive and negative activities experienced during the game.
8	Maintaining and extending new learning	The provision of items and opportunities to decorate a "lake" that represents one's mind or emotional state.

3.2 Asset Generation

To generate videos for the eight episodes, OpenAI's ChatGPT Pro, SORA and Midjourney were used. Through prompt engineering, videos were generated directly using the text-to-video approach or by first going through the text-to-image process and then using the generated images as references for prompt engineering to create videos.

A total of 182 videos were iteratively generated using SORA, and after going through the remixing and blending process with SORA, 13 final videos were selected to compose the eight episodes (Figure 2). The process took approximately 35 hours. This work was carried out by one game designer who was not an expert in prompt engineering; another game designer cross-checked whether the generated videos matched the intended design. Prompts were written with the aid of OpenAI ChatGPT Pro, and results were generated in SORA or used to generate images in Midjourney, which were then used as image prompts in SORA. Initial prompt inputs made in SORA that led to the final results are given in Appendix B. Designers selected the final results based on internal game design criteria, choosing those most effective in delivering the desired game experience.

Due to the nature of generative AI, very specific aspects of the videos could not be controlled. In particular, it was impossible to control the specific facial expressions of characters; the movement paths of characters, animals, or objects; and the game's user interface for Episodes 3, 5, and 6. In Episode 7, for which the generation of very specific photo panel designs was required, it was difficult to create videos directly using prompts in SORA. In such cases, images were generated through Midjourney and then used as seeds to create videos in SORA. For Episodes 3, 4, and 5, changes in actions and situations were needed while maintaining the composition and featuring the same characters. Therefore, to create Episodes 4 and 5, the video from Episode 3 was set as a seed, and the videos were generated by adjusting the remix intensity using SORA. When scenario changes over time were needed in a single video, such as in Episodes 2 and 6, SORA's storyboard feature was used. When blending elements of a video, such as the scene in Episode 2 during which a cube transforms into an apple, the blend function was used. It was not possible to generate specific sections of Episode 1 that focused on training auditory sensations using SORA.

Week 1	Week 2	Week 3	Week 4
			
Seed Image, Prompt Remix 38 Iterations	Prompt, Storyboard Blend 26 Iterations	Prompt, Seed Video Remix 10 Iterations	Prompt, Seed Video Remix 6 Iterations

Week 5	Week 6	Week 7	Week 8
			
Prompt, Seed Video Remix 11 Iterations	Prompt, Storyboard Blend 26 Iterations	Seed Image, Prompt Remix 39 Iterations	Prompt Remix 32 Iterations

Figure 2. Weeks 1–8 Videos

3.3 Final Implementation

To reproduce the experience of making narrative choices in a computer game faithfully in the prototype, player interaction was deliberately incorporated at critical points in the story flow. The prototype was developed using Unity and the Utage Visual Novel Tool, a framework designed for building branching narratives, managing dialogue sequences, and integrating visual and audio assets in a visual novel format. Utage's event system and scenario-scripting features allowed for the creation of structured decision points that could seamlessly alter the progression of the story without requiring complex coding.

When a single episode consisted of multiple prerecorded video segments, a choice selection screen, implemented through Utage's built-in branching command, was inserted between the videos. This ensured that players could make meaningful decisions that would determine which video played next, effectively simulating the decision-making process common in narrative-driven games. Additionally, during transitions from one episode to another, user input was gathered to trigger conditional branching, a function that Utage natively supports through its scenario management system. Depending on the choices made, different video sequences were loaded as outcomes, giving players a sense of agency and replayability. The combination of Unity's robust asset handling and Utage's specialized narrative tools enabled the creation of a smooth, interactive storytelling experience that closely mirrored the dynamics of commercial visual novels and branching video games (Figure 3).

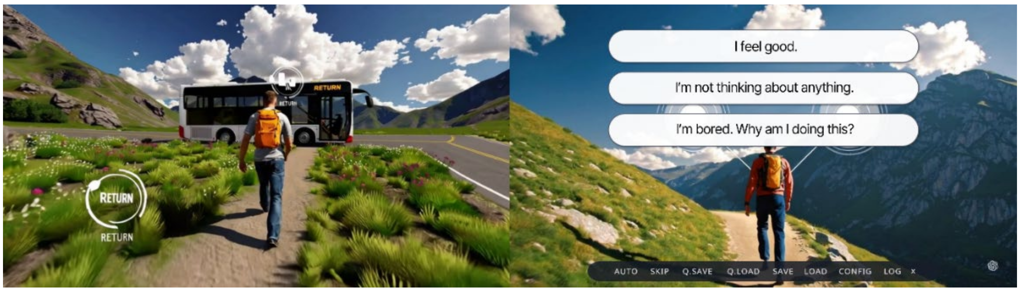

Figure 2. Choice Function

4. ASSESSMENT OF FIDELITY FROM A THEORETICAL PERSPECTIVE

Two domain experts in our interdisciplinary research team provided valuable insights into the alignment between the game prototype and MBCT principles. To evaluate the effectiveness of the video prototyping method, written semistructured assessments of its theoretical fidelity were completed by two professionals with research knowledge and practice skills in social work and mental health who had received a theoretical foundation throughout the process between 2024 and 2025. Their feedback was categorized into three major themes: alignment with theoretical expectations, differences in visual representation, and expansion of game design ideas. Notably, although the initial analysis focused predominantly on positive remarks, this refined evaluation incorporated more critical perspectives to capture the full depth and nuance of the expert insights. Key questions from the assessment and feedback text are appended (Appendix C and Appendix D, respectively).

The mental health professionals generally acknowledged that the video prototype effectively communicated the theoretical principles of MBCT. Mental Health Professional 2 noted, “The video aligns almost perfectly with what I imagined.” Similarly, both mental health professionals agreed that video-based prototyping increased the fidelity of the prototyping process.

Strengths:

- The video provided a more tangible and concrete representation of the theoretical framework.
- It allowed for the identification of gaps between the intended theory and its practical application.

Challenges and Limitations:

- Although the video clarified many aspects of the theory, some nuances of MBCT, such as the subtle cognitive shifts necessary for mindfulness practice, were not explicitly represented.
- Certain elements, such as the implementation of decision-making processes, could be more deeply integrated to reflect MBCT’s focus on mindful reactions versus automatic responses.

One professional (Mental Health Professional 1) acknowledged this limitation, stating that although the prototype effectively showcased theory-to-practice translation, it could be refined by considering “whether the game truly encourages a mindful response or if it merely presents a prescripted experience.”

Although the video prototype successfully communicated core ideas, the other professional (Mental Health Professional 2) pointed out differences between their expectations and the actual visual representation. “It was checked whether we had the same thoughts. It’s very similar to what I had in mind. However, I could see that the graphics, landscapes, people’s appearances, and shapes were different from what I imagined.”

Strengths:

- The visual representation helped experts verify their alignment with the development team’s interpretation.
- It provided a basis for refining art direction to fit the conceptual framework better.

Challenges and Limitations:

- Differences in expectations regarding the aesthetic style created some cognitive dissonance.
- Mental Health Professional 2 initially envisioned a more animated and stylized depiction, whereas the prototype leaned toward a documentary-like approach.
- The graphical design might influence the perceived target audience; as one expert mentioned, the visuals seemed more suited for younger audiences rather than university students or adults.

This feedback suggests that visual alignment plays a critical role in how theoretical concepts are perceived and internalized. Further iterations could explore stylistic variations to ensure that the representation remains engaging while maintaining theoretical fidelity.

Mental health professionals also highlighted the potential of video prototyping in facilitating creative expansion and idea generation. Mental Health Professional 1 remarked, “Watching the video, I think it generates more expanded ideas about Episode 2.” Similarly, Mental Health Professional 2 noted that the interactive discussion was enriched because they could visually engage with the prototype, leading to more refined feedback.

Strengths:

- Video prototyping encouraged deeper engagement and iterative refinement.
- It enabled domain experts to suggest improvements that might have been overlooked in a text-based or static-image prototype.

Challenges and Limitations:

- Although video prototyping facilitated idea generation, it also risked constraining creativity by concretizing certain elements too early in the process.
- Some abstract elements of MBCT that rely on personal introspection may not be fully conveyed in a visual format, potentially limiting the interpretative flexibility of the game’s design.

Mental Health Professional 2 suggested an alternative approach: incorporating more player-driven choices and emergent scenarios rather than relying on predefined sequences. This feedback highlights an important consideration—although video prototyping improves clarity, it may also inadvertently narrow the creative possibilities by cementing a specific design direction too soon.

5. DISCUSSION

Videos created using SORA cannot be precisely controlled. It was impossible to make the results of the game's interactions appear as if they were recorded gameplay footage. However, by showing the videos as if the game were functioning, the experience was conveyed more intuitively and stimulated imagination. Nevertheless, interaction is still not possible, falling short of the experience of actually playing the game. By simulating real-time input, a visual novel-style choice was provided at critical points to offer a proxy gaming experience. Through the video prototype, the results of the interaction could be sufficiently conveyed. Although game designers can communicate without loss using generative images (Lee et al., 2023), videos have higher fidelity because they can confirm the specific implementation leading to user input.

High fidelity is important for prototypes, but game developers expect quick modifications and iterative testing during the prototyping process and use tools that facilitate this (Kasurinen et al., 2013; Schell, 2008). In this regard, video prototyping requires engineering prompts and repeatedly generating videos until they match the intended vision. In many cases, the repeatedly generated videos had to be blended together. Acquiring the know-how for engineering prompts and the process of iterating videos demanded more time than initially expected. However, because prompt engineering time continues to decrease, its efficiency is hard to evaluate.

In our project, the game designer directly created the videos without loss of ideas or intentions, minimizing communication loss, and this process gradually became faster. When game episodes were explained using images, text, or examples of similar games, misunderstandings occurred. However, when shown through video, the concepts became clear. Moreover, videos were excellent for showcasing not only the game mechanics but also the look and feel of the game art, indirectly presenting the gaming experience.

Play is an experience delivered through player control and participation, yet there remains a question as to whether simply watching can convey that experience with high fidelity. If the ultimate purpose of a game prototype is to simulate the interactive experience, then like documents or still images, moving videos may still fail to deliver this interaction. In such a case, claiming that a video prototype has high fidelity could be an illusion.

In addition, although this study used a widely accessible video generator, the rapid advancement of technology suggests that with the emergence of controllable world-generation tools such as Genie 3 by DeepMind—which can instantly create navigable and interactive spaces from text prompts—the usefulness of video prototypes may quickly diminish or may never become widely adopted (Genie 3, DeepMind).

6. LIMITATIONS

Fidelity assessments of this AI-powered video-prototyping process from a theoretical perspective were conducted by our research team, including two social work and mental health professionals, whose members are not in the gaming industry. The collected insights, although valuable, may reflect a narrower scope of experience than would be ideal for broader generalization. Moreover, the assessments were conducted with written semistructured guiding questions in a unidirectional and nonreflective manner. Future studies could conduct in-depth assessments in an interactive manner. The effectiveness of the studied method should be interpreted with caution, especially regarding generalizing the research results to a game design professional group or industry.

7. CONCLUSION

This study demonstrated that AI-generated video prototyping serves as a powerful communication tool in game development, particularly in the context of interdisciplinary team collaboration, in which domain experts might have limited familiarity with games. By providing a more immersive and clear visualization of game concepts than static images or written documents, this approach enables even domain professionals (nongame experts) to grasp the design intentions quickly and offer feedback. The high fidelity of video prototypes is especially effective when applying specialized theories such as MBCT, because it clarifies how abstract principles can be translated into tangible gameplay experiences, thereby reducing potential miscommunication.

Video-based prototypes also facilitate smoother interdisciplinary collaboration. Whereas traditional documentation often struggles to convey the essence of gameplay, the use of video prototypes gives psychologists and other nongame professionals an intuitive understanding of the intended player experience. This shared visual language allowed different stakeholders to engage in faster, more precise discussions about how to combine abstract theory with game mechanics, leading to more efficient and refined teamwork.

However, the current stage of AI video prototyping does have limitations. Generating the desired results requires iterative prompt adjustments and intricate fine-tuning, which can be time-consuming, and it remains challenging to control specific details such as character expressions, animations, or user interface elements. These constraints make quick prototyping difficult, but considering the pace of technological advancement, it is reasonable to expect that more accurate and speedy video generation will become available in the future.

Considering these findings, AI-generated video prototyping may have its value as a means of delivering design intentions with high fidelity, minimizing misunderstandings in the design process, and supporting more effective collaboration across fields. As AI evolves to improve efficiency and potentially allow for interactive elements in video prototypes, its role in a broader range of game development scenarios is likely to grow, positioning video prototyping as additional methodology for communicating complex experiential designs.

In addition, although ethical considerations regarding the use of generative AI such as copyright ownership, intellectual property dilution, and potential labor displacement are increasingly relevant in creative industries, the scope of this study focuses solely on prototyping for internal communication in a game development team. These prototypes are not distributed to end users as final products, which means that the typical copyright concerns associated with commercial release do not apply in this context. Nonetheless, organizations adopting this method should remain mindful of the provenance of AI-generated assets, the terms of use of the AI tools employed, and the importance of maintaining transparency in how AI-assisted prototypes are created to ensure that both legal compliance and ethical integrity are upheld.

ACKNOWLEDGEMENTS

This research was supported by the Culture, Sports and Tourism R&D Program through a Korea Creative Content Agency grant funded by the Ministry of Culture, Sports and Tourism in 2025 (Project Name: Development of game-based digital therapeutics technology for adolescent mental health (psychological and behavioral control) management, Project Number: RS-2025-00344893, Contribution Rate: 100%).

REFERENCES

Baer, R. A. et al., 2006. Using self-report assessment methods to explore facets of mindfulness. *Assessment*, Vol. 13, No. 1, pp. 27-45. doi:10.1177/1073191105283504

Bernard, P., Cullen, C., and Kuyken, W., 2020. *Mindfulness-Based Cognitive Therapy for Life Teacher Manual*. Oxford, UK, Oxford Mindfulness Centre.

Bolchini, D., Pulido, D. and Faiola, A., 2009. FEATURE "Paper in screen" prototyping: An agile technique to anticipate the mobile experience. *Interactions*, Vol. 16, No. 4, pp. 29-33. doi:10.1145/1551986.1551992

Buchenau, M. and Fulton Suri, J., 2000. Experience prototyping. *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*. New York City, USA, pp. 424-433. doi:10.1145/347642.347802

Camburn, B. et al., 2017. Design prototyping methods: State of the art in strategies, techniques, and guidelines. *Design Science*, Vol. 3, e13. doi:10.1017/dsj.2017.10

DeepMind, 2025. *Genie 3: A new frontier for world models.* Available at: https://deepmind.google/discover/blog/genie-3-a-new-frontier-for-world-models/ (Accessed: 16 August 2025).

DeepMind, 2025. *Veo 3*. Available at: https://deepmind.google/models/veo/ (Accessed: 16 August 2025).

Feldman, C., and Kuyken, W., 2019. *Mindfulness: Ancient Wisdom Meets Modern Psychology* (1st ed.). New York, USA, Guilford Press, pp. 135-136.

Fullerton, T., 2008. *Game Design Workshop – A Playcentric Approach to Creating Innovative Games* (2nd ed.). Burlington, MA, USA, Elsevier.

Fullerton, T. et al., 2006. That cloud game: dreaming (and doing) innovative game design. *Proceedings of the 2006 ACM SIGGRAPH Symposium on Videogames*. Boston, MA, USA, pp. 51-59. doi:10.1145/1183316.1183324

Israel, J. H., Bähr, B. and Exner, K., 2016. *Perspectives on Future Prototyping—Results from an Expert Discussion*. Cham, Springer International Publishing, pp. 11-22. doi:10.1007/978-3-319-24439-6_2

Kabat-Zinn, J., 1994. *Wherever You Go, There You Are: Mindfulness Meditation in Everyday Life*. New York, Hyperion.

Kasurinen, J., Strandén, J.-P. and Smolander, K., 2013. What do game developers expect from development and design tools? *Proceedings of the 17th International Conference on Evaluation and Assessment in Software Engineering*. Porto de Galinhas, Brazil, pp. 36-41. doi:10.1145/2460999.2461004

Kultima, A. et al., 2012. *Games and Innovation Research Seminar 2011 Working Papers*. Available at: https://urn.fi/urn:isbn:978-951-44-8705-7 (Accessed: 24 January 2025).

Lauff, C. A., Kotys-Schwartz, D. and Rentschler, M. E., 2018. What is a Prototype? What are the roles of Prototypes in companies? *Journal of Mechanical Design*, Vol. 140, No. 6, 061102. doi:10.1115/1.4039340

Lee, J., Eom, S.-Y. and Lee, J., 2023. Empowering game designers with generative AI. *IADIS International Journal on Computer Science & Information Systems*, Vol. 18, No. 2, pp. 213-230.

Lee, J. J. and Lee, J., 2023. Generative AI for the Super Game Designer. *Proceedings of the International Conferences on Interfaces and Human Computer Interaction 2023 (IHCI 2023); Computer Graphics, Visualization, Computer Vision and Image Processing 2023 (CGVCVIP 2023); and Game and Entertainment Technologies 2023 (GET 2023)*, pp. 203-212. Available at: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85181821189&partnerID=40&md5=de31ebab30010675c438851e2ab3ceef (Accessed: 24 February 2025).

Ling, L. et al., 2024. Sketchar: Supporting Character Design and Illustration Prototyping Using Generative AI. *Proceedings of ACM Human-Computer Interaction*, Vol. 8(CHI PLAY), Article 337, 28 pages. doi:10.1145/3677102

Manker, J., 2011. Game Prototyping – The Negotiation of an Idea. *Proceedings of DiGRA 2011 Conference: Think Design Play*. Tampere, Finland: DiGRA. Available at: https://dl.digra.org/index.php/dl/article/view/598 (Accessed: 24 February 2025).

Menold, J., Jablokow, K. and Simpson, T., 2017. Prototype for X (PFX): A holistic framework for structuring prototyping methods to support engineering design. *Design Studies*, Vol. 50, pp. 70-112. doi:10.1016/j.destud.2017.03.001

OpenAI, 2024. *Sora*. Available at: https://openai.com/sora/ (Accessed: 24 January 2025).

Otto, K. and Wood, K., 2001. *Design for the Environment, Product Design; Techniques in Reverse Engineering and New Product Development*.

Reyno, E. M. and Carsí Cubel, J. Á., 2009. Automatic prototyping in model-driven game development. *Comput. Entertain.*, Vol. 7, No. 2, Article 29, 9 pages. doi:10.1145/1541895.1541909

Riedl, M. O. and Zook, A., 2013. AI for game production. In: *2013 IEEE Conference on Computational Intelligence in Games (CIG)*, pp. 1-8. doi:10.1109/CIG.2013.6633663

Schell, J., 2008. *The Art of Game Design: A Book of Lenses* (1st ed.). Boca Raton, USA, CRC Press. doi:10.1201/9780080919171

Segal, Z. V., Williams, J. M. G., & Teasdale, J. D., 2013. *Mindfulness-Based Cognitive Therapy for Depression* (2nd ed.). New York, Guilford Press.

Signer, B. and Norrie, M. C., 2007. PaperPoint: A paper-based presentation and interactive paper prototyping tool. *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*. Baton Rouge, LA, USA, pp. 57-64. doi:10.1145/1226969.1226981

Snyder, C., 2003. *Paper prototyping: The fast and easy way to design and refine user interfaces*. San Francisco, USA, Morgan Kaufmann.

Ulrich, K. T. and Eppinger, S. D., 2016. *Product design and development*. New York, USA, McGraw-Hill.

APPENDIX

A. Playable Prototype (Mind Trekking Demo)

https://1drv.ms/f/s!Ahutx5nfyiKuwJ0ksWGIGjNkoieH0w?e=61uun8

B. SORA Video-Generation Prompts (Full Text)

Week	Image Seed	Video Seed	Prompt
1	□		The tree sways gently in the wind, with its branches and leaves moving fluidly and naturally, while the trunk remains rooted and stable. The camera remains fixed throughout the scene.
2			A collection of colorful 3D cubes slowly orbits around a hiking backpack placed at the center of the scene. Each cube moves smoothly in a circular path, creating a dynamic yet calm visual. Suddenly, a 3D red cube pops out from inside the hiking backpack, emerging with a gentle upward motion. As the red cube hovers in mid-air, a white circular marker briefly appears around it, highlighting the cube as if it has been selected. The marker remains visible for a moment, then fades away. After the marker disappears, the red cube begins to transform seamlessly into the shape of a 3D apple, retaining its position in mid-air. The camera remains fixed, capturing the entire sequence clearly, focusing on the movement of the cubes, the marker interaction, and the transformation.
3		□	Create a 5-second bright-themed 3D animation featuring a man hiking in a scenic mountainous environment. The camera starts with a full-body back view of the man and continuously follows him, keeping the man centered in the frame throughout the video. As the man progresses forward, the mountain in front of him gradually becomes more prominent and closer in view. At the 4-second mark, a sleek, interactive UI element appears on the screen, showing a touchable icon or indicator on the mountain. The animation emphasizes smooth character movement, realistic hiking gestures, and a visually engaging bright atmosphere. Maintain a consistent focus on the man, avoiding any shifts to other elements or locations in the scene.
4		□	*Strong remix prompt, based on result 'Week 3'* Create a 5-second bright-themed 3D animation featuring a man walking on a flat trail in a scenic mountainous environment. The camera starts with a full-body back view of the man and continuously follows him, keeping him centered in the frame at all times. The character's entire body remains visible throughout the video. As the man walks, a bus with 'RETURN' written on its front display approaches along the same flat trail from a distance. The bus gradually gets closer and comes to a stop in front of the man. At the moment the bus stops, two flat, 2D white circular UI markers dynamically appear on the bus, indicating that the bus is interactive and selectable. These markers should be touch-interactive UI elements. The animation emphasizes smooth movement of the bus and the character,

			a visually appealing bright atmosphere, and a clear focus on interactivity.
5		□	Create a 5-second bright-themed 3D animation featuring a man walking slowly on a flat trail in a scenic environment. The camera starts with a full-body back view of the man and continuously follows him, keeping him centered in the frame at all times. The character’s entire body remains visible throughout the video. As the man walks slowly, several groups of people approach from the opposite direction, passing by him. Among them is a large retriever dog, initially walking alongside one of the groups. The dog begins to bark actively at the man while wagging its tail excitedly. After a moment, the dog breaks away from its group, running toward the man in an energetic and playful manner. The animation ends with the dog jumping up slightly as if trying to engage with the man, disrupting his peaceful walk. The focus remains on the man throughout, with the dog’s playful yet disruptive behavior emphasized as a contrasting element. The animation should highlight the interaction between the man and the approaching groups while maintaining a bright and engaging atmosphere.
6			Create a 5-second bright-themed 3D animation featuring a squirrel in a forest environment. The scene shows the squirrel reacting angrily after noticing that the acorn it was about to eat has disappeared. The animation focuses on the squirrel’s expressive body language and facial expressions, emphasizing frustration and anger. Maintain a bright and engaging forest setting, but keep the squirrel’s reaction as the central focus of the animation.
7	□		Create a 5-second bright-themed 3D animation featuring a bulletin board with the title ‘D’ prominently displayed at the top. The animation begins with an empty bulletin board positioned for a clear, front-facing view. Three game-like, stylized photos fly in one by one from various directions and attach to the board in the following order: A ‘RETURN’ bus photo. People with a dog photo. A squirrel sitting on a stump photo. Each photo smoothly attaches to the board with a satisfying pin or stick motion, creating a dynamic and engaging animation. The animation should emphasize the vibrant, colorful, and stylized aesthetics, with the bulletin board and its title ‘D’ as the central focus throughout the sequence.
8			Create a 5-second bright-themed 3D animation featuring a wooden decorative display case housing seven crystal orbs. The orbs are arranged in a structured layout: three on the first row, three on the second row, and one on the third row, creating a balanced and visually pleasing composition. Each orb is uniquely colored and contains a stylized photo, representing:

A dog.
A bus.
A squirrel.
A tree.
A backpack.
A bird.
A cloud.

The wooden display case is polished, bright, and stylized to enhance a cheerful and natural aesthetic. Each orb glows faintly in its unique color, and the stylized photos inside are vivid and engaging. The camera remains stationary, focusing on the entire arrangement, emphasizing the layout of the orbs, their embedded photos, and the detailed wooden case. The overall tone is bright, natural, and visually appealing.

C. Key Questions to Assess Fidelity from a Theoretical Perspective

Question 1:

Until now, we have explained the decisions made in our meetings using text and simple images. In comparison, does viewing the information in video form result in a different level of understanding?

Question 2:

After watching the video, were we able to assess whether we were all thinking along the same lines? Alternatively, if there were differences in perspectives, did the video provide an opportunity to offer more specific feedback?

Question 3:

If discrepancies between the interdisciplinary team members' views were not apparent from the text and image-based documents, did the video make these differences more noticeable?

Question 4:

This research is based on the hypothesis that video enhances the fidelity of prototyping. Please share your thoughts on this hypothesis.

Additional Comments:

Beyond the questions above, please feel free to share any further insights or feedback.

D. Feedback on Fidelity from a Theoretical Perspective (Full Text)

Feedback from a Mental Health Professional in the Interdisciplinary Research Team:

"The visuals in the video were almost exactly what I had imagined. Previously, I could not know how the characters or landscapes would be implemented, so being able to see them in the video greatly enhanced my understanding.

It confirmed that we were thinking along the same lines—it was very similar to what I had envisioned. However, as mentioned in the first point, I noticed that the graphics, landscapes, and even the appearances of people differed from my expectations. I had seen many of the scenes outlined in our proposal and had imagined them as animated sequences, but the actual video had more of a documentary-like quality.

I believe that the video naturally increases the fidelity of the prototyping process. As someone without specialized knowledge in games, I initially had no clear sense of how the game's scenario would be realized in video form or which age group it might suit. After watching today's video, I felt that the content might be more appropriate for middle or high school students—or even children—rather than college students. This is because the scenario and graphics themselves do not seem to offer significant stimulation or appeal for adults. For example, as a parent of a first-grader, I think that even elements like reboarding a bus can help elementary school children reconsider their perspectives."

Feedback from a Mental Health Professional (Specializing in MBCT) in the Interdisciplinary Research Team:

"Watching the video, as opposed to reading text and viewing static images, provided a much higher level of understanding. Although not every detail described in the text and images was fully realized in the video, I believe that the underlying theoretical intentions were effectively communicated and could be applied to the video format. Additionally, the video stimulated more expansive ideas. For instance, the scene where an apple emerges from a bag seems to introduce the concept of encountering an object during a walk. The theory suggests that one should observe an object without immediate judgment—merely noting its attributes such as 'red', 'square', 'dotted', or 'with a stem'—and spend time perceiving it as it is, rather than instantly categorizing it. This approach allows for choices that may involve adding judgment or value, but the important point is to see the object as it truly is, rather than immediately assigning it a predetermined label. Such an approach could offer an opportunity to learn the mindful, automatic adjustment process inherent in human cognition.

In another example, the choice involving the bus represents a training scenario in mindfulness-based cognitive therapy, where one must decide between reacting automatically and responding skillfully. Although taking the bus to go faster might provide short-term relief, it could be counterproductive in the long run. Experiencing this scene may help illustrate that a more precisely designed game could serve as an effective mindfulness exercise. This concept is also explained in programs like 'Vicious Flower', where it is shown that when people face difficulties, they often resort to short-term behaviors—be it shopping, gaming, gambling, binge-watching, or oversleeping—to alleviate their distress, even though such actions rarely offer long-term benefits. Mindfulness, on the other hand, encourages individuals to pause these automatic reactions and instead respond in a more skillful manner.

Furthermore, viewing the video allowed me to more clearly discern the differences in understanding between theory providers and developers. The concrete implementations shown in the video help identify gaps between theoretical intentions and practical execution, thereby enabling a more advanced integration of the two. I accept the hypothesis that video prototyping enhances fidelity, as the video's tangible implementations allow us to assess whether the final product aligns with or deviates from the theoretical framework. While other methods may exist, since games are fundamentally visual, the video naturally serves as a preview of the game to be developed, providing a basis for feedback that can help ensure the prototype remains closely aligned with the underlying theory."