An Adaptive Hint System for Puzzle Games: A Multimodal-based Approach

Hao He
School of Science and Engineering
The Chinese University of Hong Kong, Shenzhen
Shenzhen, China
haohe@link.cuhk.edu.cn

Yulin Zhu
School of Data Science
The Chinese University of Hong Kong, Shenzhen
Shenzhen, China
116020372@link.cuhk.edu.cn

Wei Cai*
School of Science and Engineering
The Chinese University of Hong Kong, Shenzhen
Shenzhen, China
caiwei@cuhk.edu.cn

Abstract—For a puzzle game, the embedded hint system is a vital medium to improve the quality of the gameplay. A well-designed hint system can help balance the difficulty of the puzzle game and offer the player a better game experience. However, according to the questionnaire-based survey and the industrial investigation we made for puzzle games, two significant problems with the current hint systems are discovered: a) Most hint systems are rule-based. Many players consider that the timing of the hints cannot match the actual states of their game flows. b) In addition, the presentation modes of the hints are sometimes too abstract or too straightforward to follow and even leak the upcoming puzzles. Hence, to improve the game experience for the players when involved in puzzle games, we design and develop a multimodal-based adaptive hint system to help players get through the puzzles more smoothly with timely hints that are presented in proper modes. To test the system's effectiveness and usefulness, we conducted a control experiment based on three puzzle games, with a control group playing the original games and an experimental group playing the games with the multimodal-based adaptive hint system. The experimental results showed that our system has a significant effect in helping players improve their game experience and balance the difficulty of the game.

Index Terms—Game Analysis, Multimodal Sensing, Game Design, Deep Learning

I. INTRODUCTION

The basic rule of a puzzle in a video game is to collect clues, acquire the related items, interact with the target area with the right approach, and then get through it [7]. However, players often get stuck [14] in the puzzle games when their game flows are out of the baselines of the designer's intentions. If stuck for a while, the players may gain much sense of achievement after successfully solving the puzzles. If stuck for an exceeding duration, the player may accumulate negative feelings as time goes by and finally causes the emergence of frustration, which will significantly impact the player's experience.

However, "being stuck" is not always harmful to the players. A dialectic perspective is required to analyze this phenomenon.

Most games contain a series of obstacles that hinder the player's progress, and the player focuses on overcoming those obstacles. Many of them enjoy games where the majority of players may fail multiple times and feel extremely frustrating (such as the Dark Souls series, which is known for its difficulty). Games that lack difficulty and have few obstacles are often considered "boring" if they have nothing else to keep the player interested. In order to balance the experience of the puzzle games, designers often provide hints to offer the players valuable tips to solve the puzzles. However, controlling when and how those tips should be revealed is a tricky problem. Okopny [4] pointed out that it can have a negative impact on the player if the puzzle-solving processes are constantly provided with hints, which will reduce motivation and the sense of accomplishment. Wolterink [8] and others, in his study, points out that the game needs to balance the stuck issue, skill, and challenge in order to keep players immersive and gain a better game experience.

As described above, the hint system plays an essential role in balancing the player's "lost" and "gain" over the game flow. To study how to optimize the hint system to improve the game experience of players, we focus on the following three user-centered [12] research problems: 1) Players' subjective experience when playing puzzle games, including the questions of when the stuck issue occurred, how they responded to stuck issue, was there any hints private in the game to improve the stuck issue, and how did they think of those hints, did they work, did they affect the experience, and did they make the game experience better or worse? 2) What representational behaviors will the players perform when they are over stuck. And how to perceive and understand the associated indicators of these behaviors? 3) What can we do to offer the players better experience based on the perception of their negative feelings combined with the optimization of the hint systems in puzzle games?

Based on the above three questions, we have carried out four principal works: 1) We conducted an open questionnaire-

based survey, which contained the questions of how much the subject likes the puzzle games, how many of them the subject has played, self-assessment of the subject's expertise in solving puzzles, what measures does the subject take when stuck, how the subject thinks about the embedded hint systems, and how the subject considers the impacts of hint system on the aspects of in-game immersion and gameplay. Based on the survey results, we comprehended that the majority of the subjects treated the hint system as an essential medium for improving their game experience. 2) In addition to the survey, we conducted an industrial investigation for puzzle games to categorize the hint systems and summarize different types of hints. We also concluded the drawbacks of each category of hints to provide insights to optimize the hint systems. 3) To understand when the player requires the help of the hint system and how they prefer the presentation mode of the hints. We conducted a user study involving 18 participants based on three puzzle games. By recording the players' behavioral performances during the game and their corresponding screen records, we analyzed the external patterns of the players when they were stuck. We found that when stuck over a long-time, the players would reveal specific behavioral patterns in facial action, head posture, and eye gaze state. With those indicators, we trained a multimodal-based LSTM model to detect the stuck state of the players in real-time by collecting the data via a paid crowdsourcing that involved 30 volunteers. 4) Based on the survey and industrial investigation, players believe that concise, accurate, implicit, text-based, and without revealing the whole game flow is the best way to present the hints. Therefore, we optimized the hint systems for the three games we applied in the previous user study with the obtained results. Supported by the above works, we built a multimodal-based adaptive hint system for puzzle games, which could detect the stuck state of the player in real-time and provide them with optimized hint content during the game.

II. RELATED WORK

Historically, tutorials, the forerunners of hints, had been provided primarily through text documents accessed outside the context of the application [3]. While documents required very few resources to generate, they presented several challenges for users, who had difficulty keeping track of instructions as they switched between tutorials and applications. As a result, developers have designed more diversified tutorials, including teaching and interactive documentation in the context of the game. Despite these improvements, many researchers believe that contextual tutorials have tremendous potential to improve the learnability of applications [5]. As emphasized by game expert James Gee [2], players need to be taught the importance of new mechanics before using them, rather than showing them without context. On-demand access to the tutorials is most effective after the user is familiar with the essential interface elements and has asked specific questions. In this study, we argue that the hint system should also follow the theory that the hints should be displayed only when the player needs them.

However, most of the hints are rule-based, which cannot match the actual needs of the players in multiple aspects [1]. The drawbacks of these hints have inspired innovations from game companies and researchers [13] that want to improve the current situations. The "Social Engineer Game" introduced an adaptive hint system for games. The system balanced the player's motivation by using built-in algorithms that intelligently gave the right hint at the right time. This work led to the concept of the adaptive hint system for games. Adaptive systems can be used to help players solve problems when they encounter difficulties, balancing frustration and motivation. Therefore, our study hopes to apply the same concept of an adaptive hint system to puzzle games.

III. PRELIMINARY WORK

In order to understand the current situation of the players' attitudes towards puzzle games, what representational behaviors they will perform during the game, and how the game industry recognizes puzzle games, we carried out three preliminary works to figure out the actuality.

A. Questionnaire-based survey

We conducted a questionnaire based on 108 college students from a university. A total of 104 players who had played at least one puzzle game were selected as data sources. According to the survey, 68.3% of users had turned to the in-game hint system while stuck. 52.8% of users thought hints made the game more playable, and 20.4% thought tips improved the game quality. However, 66.3% of users felt that the hint systems did not meet their in-game needs well in the perspectives of inappropriate timing, presentation mode, and the way to acquire.

B. Industrial investigation

According to our industrial investigation for puzzle games, we list the existing categories of hint systems and their drawbacks:

- Hint on request: distract the players and may encourage them to request whenever a little obstacle occurs.
- Hint for payment: when the payment is useless, the players may suffer from more severe frustration.
- Hint on delay: may ignore the behaviors that are done by the players on purpose.
- Hint on failure: directly reveal the fact that the players failed multiple times. This may make them feel more upset.
- Hint on loading: the time is rather short for the players to catch up, and the hints may be random and uncorrelated with the context.

Combined with the survey results, we can learn that the hint system is crucial in affecting the experience and quality of puzzle games. However, the exiting hint systems needed to be improved to achieve a better experience for the players.

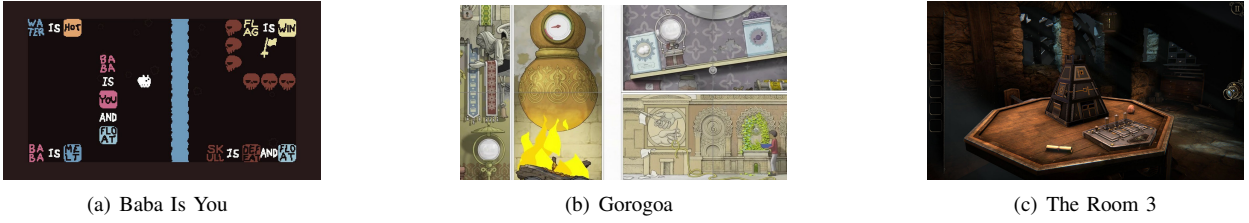| (a) Baba Is You | (b) Gorogoa | (c) The Room 3 |

Fig. 1. Screenshots of The Three Games

C. User study

To understand in what situations the player needs the help of the hint system and how they want to get the hint content presented. We conducted a user study with 18 participants. Subjects need to experience three puzzle games shown in Figure 1.

"Baba Is You" is a puzzle game where the puzzles are presented as word blocks that players can interact with. Word blocks can be used to give objects different properties. By pushing word blocks in a way similar to SOKOBAN (a game of pushing boxes), players can change how the game works, repurpose things in the levels and cause surprising interactions. The difficulty of the levels gradually increases, and the puzzle elements in the later stage of the game become more complex. This game has no embedded hint system.

"Gorogoa" is comprised of lavishly illustrated panels that players arrange and combine in imaginative ways to solve puzzles. Impeccably simple yet satisfyingly complex, expressing itself through soulful, charming illustrations and distinguished puzzle mechanics. This game has an embedded hint system in simple animation that shows the correct area to interact with, which is driven by the time of delay.

"The Room 3" is a point-and-click room escape game. Players need to find the correct items to solve the puzzle and solve the traps in multiple rooms. Players switch rooms through different clicking methods, view collected items, and interact with organs. There are various forms of organs, and the connection with puzzle-solving items is very creative. In the later stage, the puzzles are distributed in different rooms, and the clues are complex and challenging. There is a hint system in the game. According to the game's progress, the hint will be unlocked after a fixed period.

We recorded the subject's representational behaviors during the game and the corresponding game images on the screen. After the experiment, we conducted semi-structured interviews with the subjects. We marked the video clips according to the feedback from the interviewers (the points when the subject stuck, whether the subject encountered difficulty and required hints, whether a hint was revealed by the embedded hint system when stuck, and the subject's opinion on the presentation modes of hints). Finally, we manually observed the records to obtain the features of the user's representational behaviors when stuck and figured out that specific movements would be performed with the subjects' facial action, head posture, and eye gaze state.

D. Preliminary conclusion

Based on the above works, we summarized the following takeaway facts:
1) Players consider the hint system a critical medium to help them go through the game level when they encounter stuck issues and improve the game experience.
2) The exiting hint systems have different drawbacks respecting their categories. Most players criticize the hint systems from the perspective of the time of revealing and the mode of presenting.
3) Players consider the presentation mode of hint that is concise, accurate, implicit, text-based, and without revealing the whole game flow as the most promising.
4) Players will reveal specific representational behaviors when they are stuck in the dimensions of facial action, head posture, and eye gaze state.

To address the facts mentioned above, we determined to design and develop a hint system that can adaptively detect the stuck state of the players and offer them the proper hint contents to help achieve a better game experience and improve the quality of the puzzle games.

IV. THE MULTIMODAL-BASED ADAPTIVE HINT SYSTEM
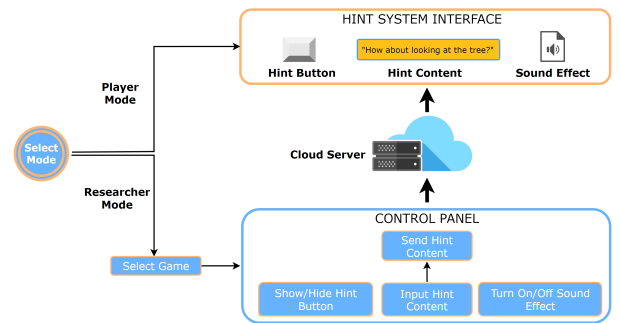
A. System overview

Fig. 2. Structure of Multimodal-based Adaptive Hint System

As shown in Figure 2, we adopt Unity 3D as the tool to develop the system. The system is divided into two modes: player mode and researcher mode. The player mode is running on the player's computer, in which the player does not notice any changes to the game content but only displays a controlled text-based hint corresponding to the aesthetic style of the game scene. In researcher mode, the researcher can use the control

panel to manipulate the hint system on the player's computer. The researcher determines whether the hint button should be displayed and the hint's content. The two sides are connected through a cloud server. Each pair of players and researcher forms a room. The server can accommodate multiple groups of experimenters at once.

To run the system, it requires a module to determine when should reveal the hint from the researcher side to the player side. According to the *preliminary conclusion 1*, players will require the assistance of the hint system when they encounter the stuck issue. Hence, in the following subsection, we will explain how we realize this module.

B. Stuck state detection

Based on the indicators obtained in *preliminary conclusion 4*, we made an academic investigation to seek the related methodologies to recognize them. As a result, we found that the works done by Alghowinem [10] could provide a matching inspiration for the case in this paper. She studied the impacts of facial action, head posture, and eye gaze state in helping the diagnosis of frustration by capturing the point clouds of the target areas. As shown in Figure 3, for facial action, the key points from the entire face were mapped into an AAM (Active Appearance Models) [9], then was used to detect the facial actions of the patients. After statistical analysis, the patients with frustration might perform **the facial actions of the mouth slightly open, pouting lips, and frowned upon**. For head posture, the Euler angles, including yaw, pitch, and row, were the essential indicators to detect frustration. After statistical analysis, it was found that patients with frustration had **slower head movements, more minor head position changes, longer time to look right, and longer time to look down**. For the eye gaze state, the eyelid distance of the open eye (c: the length of the line connecting the upper eyelid and the lower eyelid, and d: the length of the line connecting the center of the eye fold and the center of the outer border of the lower eyelid), and the gaze directions were the key indicators. After statistical analysis, it was found that **the average distance between the eyelids was significantly smaller, and the average blink time was significantly longer when the patients with frustration opened their eyes**. What is more, **their gaze directions would be offset from the object**.

To detect the indicators, we applied a data acquisition tool called OpenFace [6], which could detect the key points of the face, the Euler angles of the head posture, the openness of eyes, and the gaze direction in real-time with a RBG based webcam. With the support of OpenFace, we conducted a paid data collection crowdsourcing. 25 volunteers were asked to play the games "Baba Is you," "Gorogoa," and "The Room 3" for one hour each. During the data collection process, we asked volunteers to click a button when they were stuck and needed help so that we could label the data. The labeled time window was set as 1 minute before and after the stuck moment. With the labeled data, we built a LSTM model [11] to learn and train the features of each indicator and classify the stuck

(a) Facial key points   (b) Eyelid distance
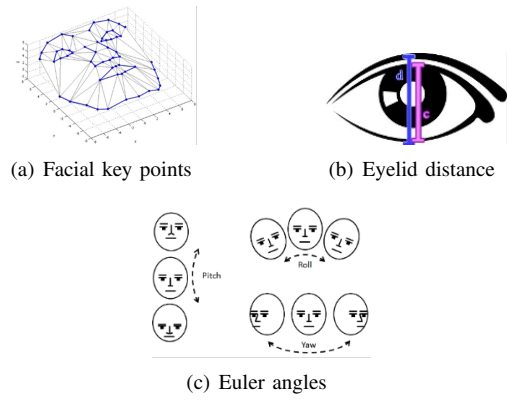
(c) Euler angles

Fig. 3. Indicators for Stuck State Recognition

state. We divided the dataset into a training dataset and a test dataset in a ratio of 8:2. The input layer was connected to the LSTM layer, a dropout layer followed the LSTM layer, and then two fully connected layers (100 neurons in the first layer and 2 neurons in the second layer). The classification accuracy was 93% on the training set and 78% on the test set. Hence, with the guaranteed accuracy, we embedded the LSTM model into the system to help detect the stuck issue of the player and remind the researcher side to reveal the hint to the player side.

V. EXPERIMENT

In order to verify the effectiveness and usefulness of the multimodal-based adaptive hint system in puzzle games, we conducted a control experiment.

A. Participants

We recruited 30 participants between 18 and 28 years old within the campus of a university, 13 of them were male, and the others were female. All the participants were in healthy physical and mental conditions. The participants had acknowledged the process of the experiment and agreed with the measurements we took during the whole procedure, including visual and audio materials recording and the collections of their answers respecting the prepared questionnaire.

A group of subjects would play the original puzzle game, while another group would play the puzzle game with the support of our system. Consistent with our former user study, we chose "Baba Is You," "Gorogoa," and "The Room 3" as the experimental scenes. Subjects in the experiment would be randomly assigned to the two groups with one of the three games. Both groups would play the same game levels and contents. All the subjects should have never played the game before. As a result, we assigned 5 participants for each game in each group.

The irrelevant variables involved in the experiment were the game ability of the subjects, the difficulty of the game level, and whether the subject had played the game level before. In order to eliminate the influence of irrelevant variables, we investigated the related background of the subject. By random

(a) Baba Is You
(b) Gorogoa
(c) The Room 3

Fig. 4. Examples of The Optimized Hints

assignment, the experimental group and the control group would not be affected by any of those irrelevant variables.

B. Hint optimization

Based on the feedback from the players in our preliminary work, we optimized the presentations of the hints for the three games. As Figure 4 shows, the text behind the question mark represents an example of a hint that we may provide for the players. The optimized hint system is deeply adapted for each game, such as the style of the hint button, the font, and the color of the text. Those features can ensure that players will not be affected by inconsistent game content, which may reduce their immersions. We set up text-based hints that are concise and enlightening. We designed a two-step tip for the rather challenging puzzles that the players may encounter obstacles more likely. The first hint would give the player some clues and keep the fun of thinking, while the second hint would give the answer and keep the player from getting frustrated.

C. Apparatus and procedure

All the subjects from the two groups were presented with an AOC screen with a resolution of 1920*1080. A keyboard and a mouse that connected to an Alienware PC were provided for them to interact with the game contents. For the experimental group, as Figure 5 illustrates, we utilized a RGB based webcam to collect the multimodal information from the subject, which was transmitted to the LSTM model for the prediction of each state of the features, including facial actions, head posture, and eye gaze state. After predicting the features, we used the prediction model from the stuck state detection module to determine whether the subject should be given a hint at this moment. When the researcher noticed a positive answer, he or she could input the optimized hint content and play a tip sound for the player. Then the subject could choose to click the button and read the hint content. For the control group, the subject experienced the original game contents without the support of the system.

After the experiment, a questionnaire-based semi-structured interview would be conducted to collect objective game experience for each subject using multidimensional evaluation indicators (a score that ranks with a Likert Scale from 1 to 5). The questions were: *How interesting is the game topic? How interesting is the gameplay? How enjoyable you are during the game? How helpful is the hint system? How frustrated you are when stuck? How difficult is the game?* We analyzed and visualized the collected data to explain the differences between the two groups.
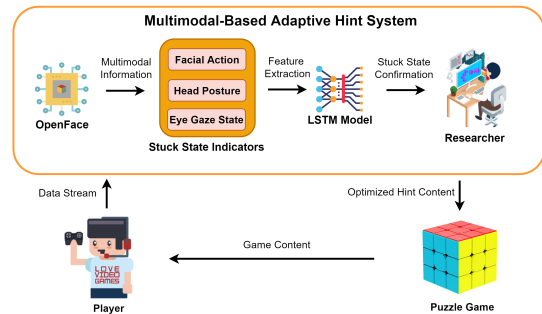

Fig. 5. Interaction Illustration of Multimodal-based Adaptive Hint System

VI. RESULT AND DISCUSSION

As Figure 6 presents, we visualized the multidimensional feedback from all the participants in the form of a radar map. The results were averaged and rounded off to make them easy to read. From the figure, we can see that the dimensions related to the game experience and the usefulness of the hint system significantly improved in the experimental group compared to the control group. Besides, the difficulties of the three games were balanced to an intermediate level in the experimental group (between 2 points and 3 points). While the ones in the control group were consistent with the actual difficulty levels of the games, which were "Baba Is You," "The Room 3", and "Gorogoa," from difficult to easy. Most importantly, the participants from the experimental group had experienced fewer frustration feelings compared to the ones in the control group. The indicators that were irrelevant to the hint system, including the topic of the game and gameplay, were similar between the two groups. Hence, we can confirm the effectiveness and the usefulness of the multimodal-based adaptive hint system in helping improve the game experience for the players.
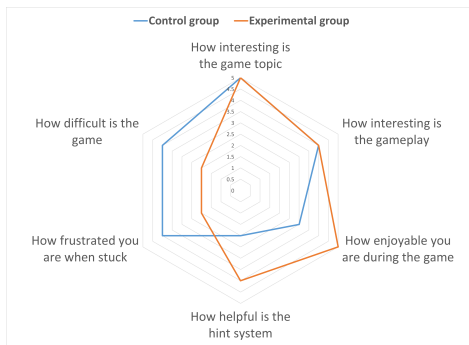
VII. LIMITATION AND FUTURE WORK

In this work, we adopted a half-automatic method to present the multimodal-based adaptive hint system. The stuck state of the player was detected automatically, but the corresponding hint had to be manually revealed by the researcher, which
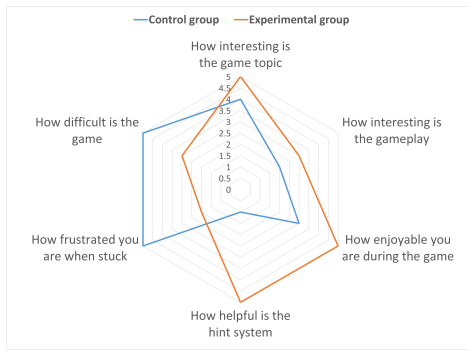
(a) Baba Is You
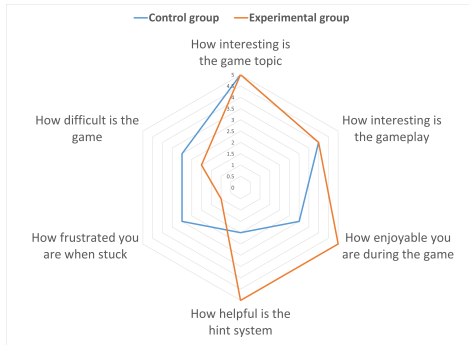
(b) Gorogoa

(c) The Room 3

Fig. 6. The average multidimensional evaluation between the control group and the experimental group for the three games. The blue line represents the control group, and the orange line represents the experimental group.

held back the extension of the scale of the experiment and the future utilization in the industry of puzzle games. To handle this issue, we need to collaborate with the game manufacturer to obtain an in-game API to track where the player is during the game. The second limitation was the optimization of the hints. We pre-prepared the corresponding hint at each point that we thought the player might encounter the stuck issue. However, this step was made based on our subjective opinion. The future version of the hint should be made with the NLP model according to the player's game flow in real-time. The third limitation was that a deeper discussion about the impact of the proposed system was unfulfilled due to the limited space of the paper. We will continue to dig into insights from the user study in the future.

VIII. CONCLUSION

In this paper, we conducted a series of studies to understand the current situation of puzzle games and how to improve the player's experience. We identified our research subject from a user-centered perspective by conducting questionnaires, user studies, and interviews and provided optimized hints to players by detecting their stuck states. We applied the related techniques of affective computing and deep learning, combined with the optimization of hints to design and develop a multimodal-based adaptive hint system for puzzle games. Via comparative experiments, it was proved that players who experienced the proposed system gained better game experience in multiple dimensions.

IX. ACKNOWLEDGEMENT

REFERENCES

[1] Aleven, Vincent , and K. R. Koedinger . "Limitations of Student Control: Do Students Know When They Need Help?" Springer-Verlag (2000).

[2] J. P. Gee. Learning by design: games as learning machines. interactive Educational Multimedia, 8:15-23, 2004.

[3] Andersen, Erik, et al. "The impact of tutorials on games of varying complexity." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012.

[4] Okopny, P. , I. Musabirov , and D Alexandrov. "Informal In-Game Help Practices in Massive Multiplayer Online Games." International Conference on Social Informatics Springer International Publishing, 2014.

[5] L. Bergman, V. Castelli, T. Lau, and D. Oblinger. Docwizards: a system for authoring follow-up-me documentation wizards. in UIST '05 Proceedings of the 18th annual ACM symposium on User interface software and technology, New York, NY, USA, 2005. acm.

[6] OpenFace 2.0: Facial Behavior Analysis Toolkit Tadas Baltrušaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency, IEEE International Conference on Automatic Face and Gesture Recognition, 2018

[7] Gabriela, G. "Serial sketching: visual problem solving in designing". Cybernetics and System.(1992). 191-219.

[8] Max Wolterink and Sander Bakkes. 2021. Towards explainable prediction of player frustration in video games. in The 16th International Conference on the Foundations of Digital Games (FDG) 2021 (FDG'21). Association for Computing Machinery, New York, NY, USA, Article 33, 1-10. DOI:https://doi.org/10.1145/3472538.3472566.

[9] Haase, D., Rodner, E., & Denzler, J. (2014). Instance-weighted transfer learning of active appearance models. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1426-1433).

[10] Joshi, J., Goecke, R., Alghowinem, S., Dhall, A., Wagner, M., Epps, J., ... & Breakspear, M. (2013). Multimodal assistive technologies for depression diagnosis and monitoring. Journal on Multimodal User Interfaces, 7(3), 217-228.

[11] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. Neural computation, 31(7), 1235-1270.

[12] Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications, 37(4), 445-456.

[13] Helen Wauck and Wai-Tat Fu. "A Data-Driven, Multidimensional Approach to Hint Design in Video Games". In Proceedings of the 22nd International Conference on Intelligent User Interfaces, pp. 137-147, 2017.

[14] Drey, T., Fischbach, F., Jansen, P., Frommel, J., Rietzler, M., & Rukzio, E. (2021). To Be or Not to Be Stuck, or Is It a Continuum? A Systematic Literature Review on the Concept of Being Stuck in Games. Proceedings of the ACM on Human-Computer Interaction, 5(CHI PLAY), 1-35.