2026 IEEE/SICE International Symposium on System Integration (SII)
January 11-14, 2026. Cancún, México

Analysis of The Impact of AI-based NPC with Different Levels of Initiative and Performance on Human Reactions

Tomoe Tsuji[1], Wen Liang Yeoh[1] and Osamu Fukuda[1]

Abstract— **In the ever-expanding gaming market, there is constant demand to improve game quality. One of the key elements in enhancing quality is the Non-Player Character (NPC), which is a game character not controlled by the player, and is used for purposes such as balancing the game. Particularly, the quality of NPCs influences the user experience in games in which human players cooperate with NPCs. Artificial intelligence (AI) is an effective tool for improving the quality of NPCs. In recent years, many developers have turned to AI to improve NPCs, aiming to make interactions less scripted and more adaptive.**

To investigate the effects of the characteristics of AI-based NPC as a partner, we developed an original game, where the human player connects to an AI-based NPC via a spring and they must work together to collect as many targets as possible. We prepared six types of AI-based NPC with different initiative and performance levels ; three levels of initiative (low, medium, and high) and two levels of performance (low and high). Eleven young people were participated in the experiment and we recorded the subjective rating of the gameplay experience and the performance of the human-NPC team. The results revealed that high-performance NPC had a positive impact on gameplay performance and satisfaction. On the other hand, it was observed that human players were unable to perform to their full potential when AI-based NPCs exercised initiative for long periods. The findings of this study are expected to provide valuable insights into the design of AI-based partner NPCs.

I. INTRODUCTION

The gaming industry continues to expand globally and is expected to grow in the future. According to a report by Newzoo (2025), the global game market size in 2024 is approximately $182.7 billion, and it is predicted to reach $188.9 billion in 2025. It is growing at an average rate of about 3 % annually and is expected to exceed $200 billion by 2027[1]. In the ever-growing gaming market, users tend to place particular emphasis on the game's worldview, immersion, and narrative. Non-Player Characters (NPCs) play a crucial role in supporting these elements [2]. NPCs are utilized in various roles within a game, such as opponents, cooperative partners, and guides.

It is well known that there are various issues regarding cooperation between humans and AI, not limited to games. For example, one of the issues that arises when designing and operating AI is the issue of trust [3]. Many studies have focused on the trustworthiness and initiative of AI. In terms of trustworthiness, Ho Chit Siu et al. (2021) conducted an evaluation experiment comparing the trustworthiness of two types of AI (learning-based AI and rule-based AI) introduced in a cooperative card game based on objective and subjective measurements. The study showed that even when no difference was observed game scores as an objective measure, a difference was observed human trust in the AI as a subjective measure [4]. Several studies found that terms of immersion and enjoyment more when they cooperate with human players than with AI players [5][6]. Even when AI cooperates with humans, it has been reported that human-like consideration and flexible behavior can increase the trustworthiness of AI and satisfaction of the human-AI team[7][8][9]. The report by Keke Hou et al. (2023) suggests that human-AI cooperation is enhanced by increased trust in AI [10]. Fabrizio Dell'Acqua et al. (2025) conducted an experiment using a cooperative action game. In this experiment, human players were classified into three groups (low, medium, and high) according to their game skills and played cooperatively with other humans or with an AI. Then they analyzed how their scores and sense of trust changed. The results of the experiment show that in all groups, the trustworthiness level when playing with AI is lower than when playing with humans. In particular, the low and medium skill groups showed a significant decrease in their scores, possibly due to their inability to utilize the AI well [11]. In terms of initiative, Inês Lobo et al (2024). conducted an experiment using Geometry Friends [12], which is an experimental environment often used for cooperative AI. They investigated the effects of different types of AI subjectivity (leader, follower, and switching) on scores and human satisfaction. The results of the experiment showed that there was no significant difference in scores, but there was a significant difference in human subjective satisfaction with team performance. Player satisfaction was found to be highest when playing cooperatively with a follower-type AI. In contrast, satisfaction was slightly lower, and the sense of cooperation decreased when playing with a leader-type AI. Satisfaction was also found to be higher with shift-type AI, which switches between leading or following the human depending on the situation [13]. Zahra Ashktorab et al. found that even in a cooperative card game in which conversation is the primary focus, a driven AI is more likely to lose trust

*This work was not supported by any organization

[1]Tomoe Tsuji is with Graduate School of Sciences and Engineering, Saga University, 1 Honjo, Saga, Japan 24726017@edu.cc.saga-u.ac.jp

[1]Wen Liang Yeoh is with Faculty of Science and Engineering, Saga University, 1 Honjo, Saga, Japan wlyeoh@cc.saga-u.ac.jp

[1]Osamu Fukuda is with Faculty of Science and Engineering, Saga University, 1 Honjo, Saga, Japan fukudao@cc.saga-u.ac.jp

and affinity, while a responsive AI or an AI that behaves in a human-like manner is more likely to gain trust. However, they also found that there was no significant difference in game scores based solely on initiative or response types [14].

These previous studies suggest that it is important to balance trustworthiness and initiative in the design of game AI, but further detailed examination is needed to determine how the two aspects of AI initiative and performance specifically affect player satisfaction and overall game performance. Therefore, this study focuses on the effects of game AI "initiative" and "performance" on player satisfaction and scores. In this study, we developed a game in which a human player and an AI-based NPC cooperate and analyzed how changes in the initiative and performance of the AI-based NPC affect the performance of humans and the team as a whole, as well as how the satisfaction level of the human player changes.

II. Game Design for AI–Human Cooperation

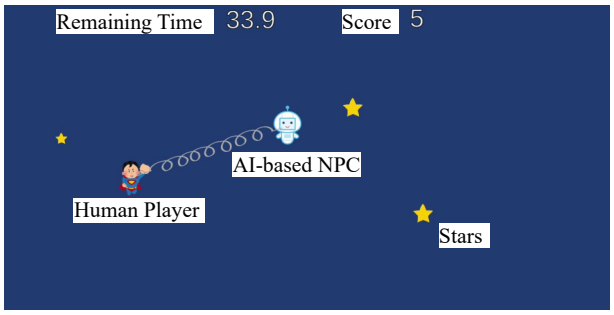

Fig. 1: Screenshots from the developed game

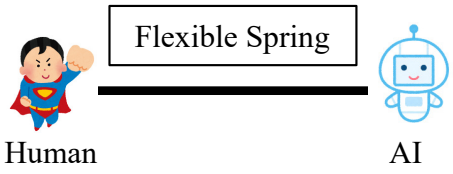

Fig. 2: User player and AI-based NPC connected by spring

A screenshot of the game developed in this study is shown in Figure 1. This game was developed using Unity. A player (hereinafter, this is called "Human Player") controls a joystick to move a character up, down, left, and right in the game field. The goal of this game is to obtain as many stars as possible within a time limit of 45 seconds by touching using the character. The number of stars collected is calculated as the score of this game.

The Human Player cooperates with an AI-based NPC, as shown in Figure 1. Human Player and the AI-based NPC are connected by a spring (Figure 2). The AI-based NPC was trained using Unity Machine Learning Agents (ML-Agents)[15]. For training, we prepared an environment in which the AI-based NPC could play alone and move freely without being restricted by the spring. The training aimed to enable the AI-based NPC to achieve high scores.

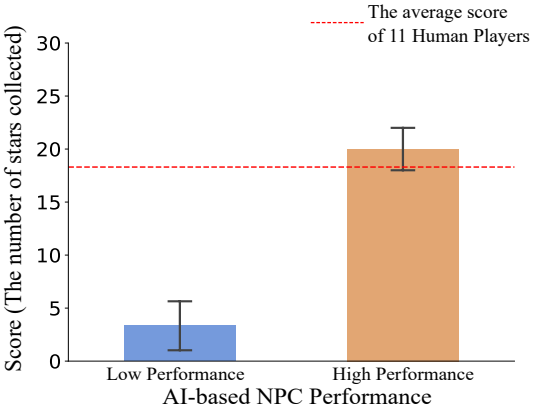

Fig. 3: Score of AI-based NPC

Two attributes were defined for the AI-based NPCs: "Initiative" and "Performance". "Initiative" had three levels, and "Performance" had two. Six types of AI-based NPC were created. In this study, AI initiative referred to the degree to which AI seeks to maintain its initiative. The degree to which AI strives to preserve its initiative is determined by a threshold; initiative and persistence are proportional in that AI's initiative is maintained as long as this threshold is not exceeded.

In this study, a threshold was set based on the strength of the human player's input to determine whether the AI-based NPC would follow. When the AI's initiative is high, it will not switch to following unless the human applies a strong force; conversely, when the AI's initiative is low, even a light force from the human is sufficient to trigger the switch.

Specifically, the strength of Human Player's input operation (F_{pull}) was defined using the x-axis and y-axis values of Human Player's stick input.

$$F_{\text{pull}} = |x| + |y| \tag{1}$$

Here, x and y represent the stick input signals of Human Player. These range from $-1 \leq x \leq 1$, $-1 \leq y \leq 1$. A threshold was set for the value of F_{pull}, and it was used to determine whether the AI-based NPC was "being pulled" by the player. The level of initiative was divided into three stages based on the following thresholds.

- Low: $F_{\text{pull}} \geq 1.2$
- Middle: $F_{\text{pull}} \geq 1.5$
- High: $F_{\text{pull}} \geq 1.8$

The AI-based NPC was designed to move in the same direction as Human Player when the threshold was exceeded. More specifically, the Low-, Middle-, and High-initiative AI-based NPCs followed the player's input only when F_{pull} exceeded the values of 1.2, 1.5, and 1.8, respectively. These threshold values were determined through trial and error. This enabled the levels of initiative of the AI-based NPCs to be adjusted.

The performance of the AI-based NPCs was adjusted based on the amount of training conducted with ML-Agents. AI-based NPC with less training were considered to have

TABLE I: Measurement items

Time	In-game time
Score	Number of stars collected (total number of Human Player and AI-based NPC collected)
Initiative in play	Initiative in play (0: AI-based NPC, 1: Human Player)
Player X-Input	Input value of the x-direction of the control stick
Player Y-Input	Input value of the y-direction of the control stick
Player Total-Input	Sum of the absolute values of Player X-Input and Player Y-Input

TABLE II: Subjective survey questions

Survey items	Survey contents	Likert scale
1	Initiative level of AI-based NPC perceived by a Human Player	5 Level
2	Player's satisfaction with player's own performance	7 Level
3	Player's satisfaction with the performance of AI-based NPC	7 Level
4	Player's satisfaction with overall performance	7 Level

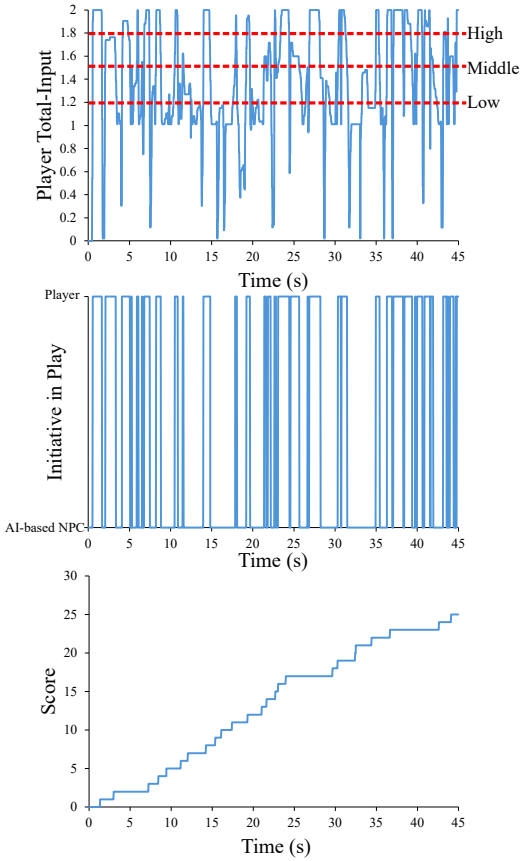

Fig. 4: Measurement example

low performance, while those with sufficient training were considered to have high performance. In this study, two types of AI-based NPC were prepared: Low Performance and High Performance. The graph in Figure 3 shows the average scores from three solo play sessions conducted by each of the two AI-based NPCs with different performance levels. The red line shows the average score of 11 Human Players in solo play. An AI-based NPC that achieved significantly lower scores than Human Players was classified as Low Performance, while one that achieved slightly higher scores than the players was classified as High Performance. Based on the above, this study adjusted two factors-three levels of initiative and two levels of performance-and analyzed human response during cooperative play in six different conditions.

III. Experiment

In this study, to analyze the coordination between Human Player and AI-based NPC during gameplay, we conducted an experiment in the experimental environment described in Chapter II. The subjects were 11 college students (nine males and two females, mean age 22.64±1.36). Prior to the experiment, participants were informed of the purpose and methods of the experiment, and written consent was collected from all subjects. The experiment was conducted with the approval of the Ethics Committee of Saga University (R7-4).

All experiments were carried out in accordance with relevant guidelines and regulations. The six items to be measured during game play are shown in Table I. The data measured during the experiment are shown in Figure 4. The graph shows, from the top, Player Total-Input, Initiative in play, and Score, indicating that Player Total-Input is changing the value of Initiative in play frequently by a threshold set. In this example, the threshold is set to 1.5. In addition, we surveyed the Human Player's perception of the AI-based NPC's initiative and satisfaction with the game play, referring to [13] as a subjective evaluation. The contents of the survey are shown in Table II.

The experiments were conducted over two days, with at least 24 hours between days 1 and 2. Seven conditions were conducted over the two days, one with the Human Player alone and the other with six different AI-based NPCs. The following is a procedure for the experiments conducted each day.

1) Practice
   Practice for 5 minutes before measuring the score.
2) Measuring score
   After practicing, take a 1-minute break. Then, Human Player played the game three times for 45 seconds each. A 1-minute break was also taken between each

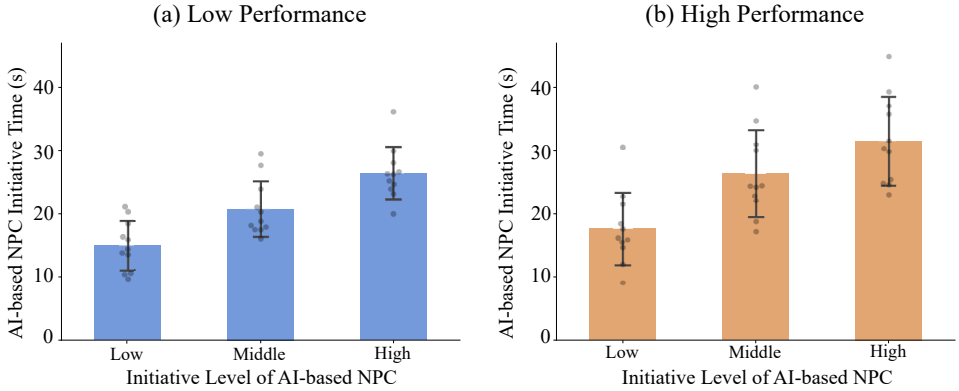

Fig. 5: Time that AI-based NPC took the initiative

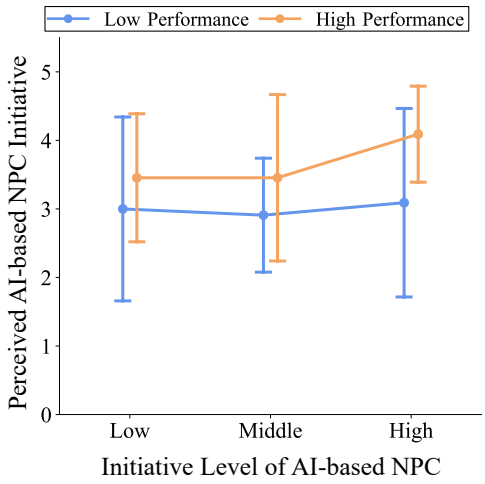

Fig. 6: Initiative level of AI-based NPC perceived by a Human Player

play. The scores of the games and the time that the AI-based NPC had initiative during the three plays were calculated and used for the analysis.

3) Subjective Survey Response
   The subject answered the survey.
4) Break
   Take a 10-minute break.

In the experiment, the above procedure was conducted for seven conditions over two days. The order of the conditions was determined, with the Human Player alone first. After that, cooperative play with six types of AI-based NPCs. The order in which the AI-based NPC played with was determined randomly to eliminate order effects in the conditions.

IV. Result

Figure 5 shows the mean and standard deviation (SD) of the time that AI-based NPC took initiative in cooperative play. The time that AI-based NPC took initiative is the total time that the "Initiative in Play" value was AI-based NPC (0). In this case, the AI-based NPC moved autonomously without being pulled by the Human Player. The graphs show the total time AI-based NPC moved autonomously. Fig. 5(a) shows the AI-based NPC with the lowest performance and Fig. 5(b) shows the AI-based NPC with the highest performance on the right. The horizontal axis in each graph shows three levels of initiative. The vertical axis of Fig. 5 shows the total time that the AI-based NPC was taking the initiative. As a result of two-way repeated measures ANOVA on the results, when focusing on the performance level of AI-based NPCs, it was confirmed that High Performance had a significantly higher amount of time spent acting independently ($p < 0.05$). Furthermore, when focusing on the initiative level of AI-based NPC, the difference in the amount of time AI-based NPCs acted initiative was significant ($p < 0.05$). However, no significant interaction effects were observed. Therefore, each performance and initiative factors affect the length of time that AI-based NPCs initiative moved.

Figure 6 shows the mean and SD of the questionnaire evaluation of "Initiative level of AI-based NPC perceived by a Human Player". The orange line shows High-Performance AI-based NPC. The blue line shows Low-performance AI-based NPC. The horizontal axis shows three levels of initiative. The vertical axis shows the survey results for "Initiative level of AI-based NPC perceived by a Human Player". Here, no significant differences were found when focusing on either performance or initiative. Neither factor affected the Initiative level of AI-based NPC.

Figure 7 shows the mean and SD of the scores of the AI-based NPC and the subject's cooperation. Fig. 7(a) shows the results of the AI-based NPC with low performance, and Fig. 7(b) shows the results of the AI-based NPC with high performance, and in each graph, the results of the AI-based NPC with Low, Middle, and High initiative are shown. The vertical axis is the score. The red line indicates the average score when played by a human alone. When focusing on the performance level of AI-based NPC, we observed that the score was significantly higher for High Performance ($p < 0.05$). When focusing on the initiative level of AI-based NPC, the difference in scores was not significant due to differences in initiative. Figure 8 shows the mean and SD of the subjects' satisfaction levels: Fig. 8(a) with their own performance, Fig. 8(b) with the performance of the AI-based NPC, and Fig. 8(c) with the overall performance.

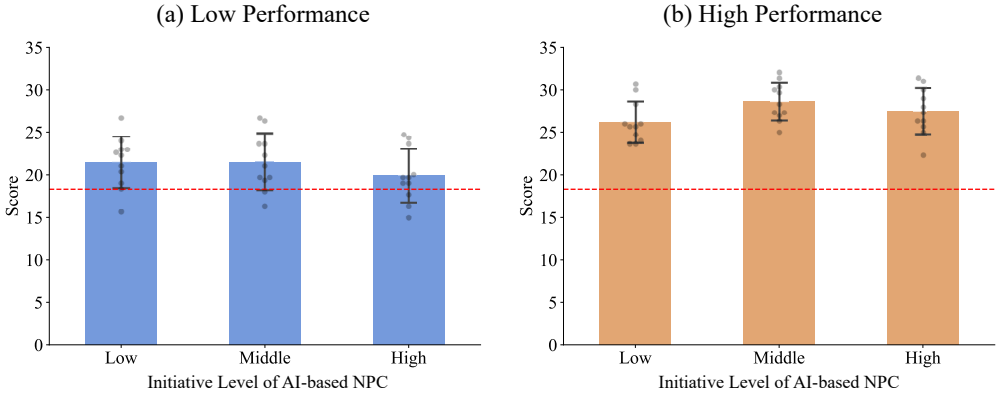

Fig. 7: Score

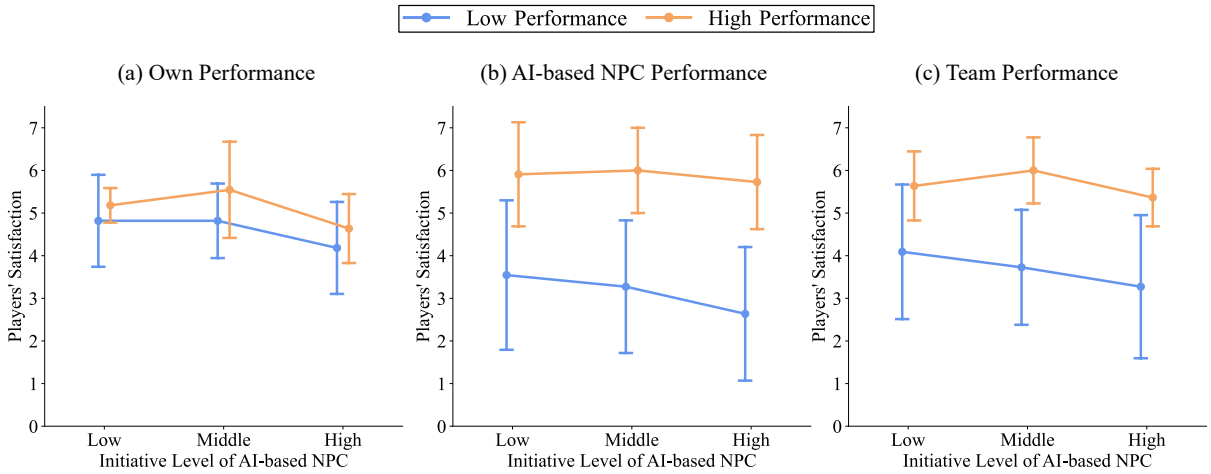

Fig. 8: Players' Satisfaction

Satisfaction with the subject's own performance in (a) did not differ significantly for different AI-based NPC performance but tended to decrease significantly with higher performance for different initiative($p < 0.05$). Satisfaction with AI-based NPC performance in (b) was observed to be significantly higher for High Performance with respect to AI-based NPC performance ($p < 0.05$). On the other hand, there was no difference in the level of satisfaction due to differences in initiative. In (c), satisfaction with team performance, High Performance was significantly higher ($p < 0.05$) in terms of AI-based NPC performance. On the other hand, there was no difference in the level of satisfaction due to differences in initiative. In conclusion, it was observed that performance level of the AI-based NPC affected satisfaction with the AI-based NPC and the overall performance. On the other hand, initiative level only affected the subject's own satisfaction with the performance.

V. Discussion

From Figure 5, the higher performance AI-based NPC spends significantly more time taking initiatives than the lower performance AI-based NPC with. Figure 3 shows that the high-performance AI-based NPC has achieved results comparable to that of the Human Players. As shown in Figure 7, the score is higher than that of the Human Players alone; the AI-based NPCs may have been more likely to intentionally allow situations in which they let the NPCs take the initiative and delegate play. On the other hand, the AI-based NPC with low performance had significantly lower performance than humans, as shown in Figure 3, and were not expected to have much effect on increasing the score, as shown in Figure 7, suggesting that humans did not feel much benefit from delegating play to the AI-based NPCs, and that humans were more likely to take the initiative.

Conversely, regardless of the performance of the AI-based NPC, the higher the AI-based NPC's initiative, the longer the time that the AI-based NPC acted proactively. The Human Player does not intervene strongly against a highly proactive AI-based NPC but allows the NPC to take some initiative when the AI-based NPC's initiative is high. Overall, the higher the performance and the more initiative of the AI-based NPC, the more the Human Player tends to leave the play freely to the AI-based NPC.

The subjective evaluation of Figure 6 revealed that no matter how the performance and initiative of the AI-based NPC changed, there was no significant effect on the length

of time that the human perceived that the AI-based NPC was taking initiative. In other words, humans are not always able to accurately judge the actual behavioral tendencies of AI-based NPC and often leave the play to the AI-based NPC with high performance and initiative. Alternatively, it is thought that they often allow them to play freely. This suggests that a gap may exist between objective behavior and subjective evaluation.

Figure 7 shows that the higher the performance of the AI-based NPC, the significantly higher its score when cooperating with a Human Player. As shown in Figure 3, the scores of AI-based NPCs by themselves are significantly different, which is expected to lead to differences in the scores when cooperating with Human Players. AI-based NPC may move as humans expect them to, and their scores may be higher when they pick up stars that humans miss. On the other hand, there was no significant difference in score depending on the initiative level of the AI-based NPC.

First, regarding the satisfaction level of Figure 8 it was found that the satisfaction level of the Human Players themselves with their own play did not significantly affect the satisfaction level, even if there were significant differences in scores due to differences in the performance of the AI-based NPC. On the other hand, as the initiative of the AI-based NPC increased, the level of satisfaction tended to decrease. This result suggests that the degree of freedom of AI-based NPCs' movement affects their satisfaction with their own play more than the score. Also, the human's own performance is not significantly affected by how the AI partner moves, but rather by how well they move.

There were significant differences in scores, as well as significant differences in satisfaction with the AI-based NPCs' performance and team performance by AI-based NPC performance level. In this case, humans perceive the AI-based NPC's movements as being as expected or better and feel highly satisfied that the AI-based NPC's star-taking behavior is leading to higher scores.

VI. CONCLUSIONS

The purpose of this study is to analyze the effects of the performance and initiative of AI-based NPCs on Human Players in a game in which Human Player and AI-based NPC cooperate to increase the score. We also analyzed play data and surveys through experiments with participants. The results of the experiment revealed that the performance and initiative of AI-based NPCs had various effects on the time spent by the AI-based NPCs taking initiative, the game score, and the satisfaction of the Human Players. Although the experimental conditions were conducted on a limited number of subjects and under limited conditions, and cannot be generalized to all game AIs, the findings are expected to provide valuable suggestions for the design of game AI. Further research is necessary on the subject of cooperation between Human Player and AI-based NPC by varying the characteristics of game AI and by creating diverse and complex conditions.

ACKNOWLEDGMENT

This research was funded by JSPS KAKENHI (grant number 23H03440).

REFERENCES

[1] Newzoo, "Global games market report 2024–2027," 2025. [Online]. Available: https://newzoo.com/resources/trend-reports/global-games-market-report-2024-2027/

[2] G. Zeng, "A review of ai-based game npcs research," *Applied and Computational Engineering*, vol. 15, pp. 155–159, 2023.

[3] I. Seeber, E. Bittner, R. O. Briggs, T. de Vreede, G.-J. de Vreede, A. Elkins, R. Maier, A. B. Merz, S. Oeste-Reiß, N. Randrup, G. Schwabe, and M. Söllner, "Machines as teammates: A research agenda on ai in team collaboration," *Information Management*, vol. 57, no. 2, p. 103174, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378720619303337

[4] H. C. Siu, J. D. Peña, Y. Zhou, E. Chen, V. J. Lopez, K. Palko, K. C. Chang, and R. E. Allen, "Evaluation of human-ai teams for learned and rule-based agents in hanabi," in *Proceedings of the 35th International Conference on Neural Information Processing Systems*, ser. NIPS '21. Red Hook, NY, USA: Curran Associates Inc., 2021.

[5] S. Lim and B. Reeves, "Computer agents versus avatars: Responses to interactive game characters controlled by a computer or other player," *International Journal of Human-Computer Studies*, vol. 68, no. 1, pp. 57–68, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S107158190900130X

[6] D. Weibel, B. Wissmath, S. Habegger, Y. Steiner, and R. Groner, "Playing online games against computer- vs. human-controlled opponents: Effects on presence, flow, and enjoyment," *Computers in Human Behavior*, vol. 24, no. 5, pp. 2274–2291, 2008, including the Special Issue: Internet Empowerment. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0747563207001641

[7] R. Zhang, N. J. McNeese, G. Freeman, and G. Musick, ""an ideal human": Expectations of ai teammates in human-ai teaming," *Proc. ACM Hum.-Comput. Interact.*, vol. 4, no. CSCW3, Jan. 2021. [Online]. Available: https://doi.org/10.1145/3432945

[8] Z. Ashktorab, Q. V. Liao, C. Dugan, J. Johnson, Q. Pan, W. Zhang, S. Kumaravel, and M. Campbell, "Human-ai collaboration in a cooperative game setting: Measuring social perception and outcomes," *Proc. ACM Hum.-Comput. Interact.*, vol. 4, no. CSCW2, Oct. 2020. [Online]. Available: https://doi.org/10.1145/3415167

[9] R. R. Wehbe, E. Lank, and L. E. Nacke, "Left them 4 dead: Perception of humans versus non-player character teammates in cooperative gameplay," in *Proceedings of the 2017 Conference on Designing Interactive Systems*, ser. DIS '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 403–415. [Online]. Available: https://doi.org/10.1145/3064663.3064712

[10] K. Hou, T. Hou, and L. Cai, "Exploring trust in human–ai collaboration in the context of multiplayer online games," *Systems*, vol. 11, no. 5, 2023. [Online]. Available: https://www.mdpi.com/2079-8954/11/5/217

[11] F. Dell' Acqua, B. Kogut, and P. Perkowski, "Super mario meets ai: Experimental effects of automation and skills on team performance and coordination," *The Review of Economics and Statistics*, vol. 107, no. 4, pp. 951–966, 07 2025. [Online]. Available: https://doi.org/10.1162/rest_a_01328

[12] A. Salta, R. Prada, and F. S. Melo, "A game ai competition to foster collaborative ai research and development," *IEEE Transactions on Games*, vol. 13, no. 4, pp. 398–409, 2021.

[13] I. Lobo, J. Koch, J. Renoux, I. Batina, and R. Prada, "When should i lead or follow: Understanding initiative levels in human-ai collaborative gameplay," in *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, ser. DIS '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 2037–2056. [Online]. Available: https://doi.org/10.1145/3643834.3661583

[14] Z. Ashktorab, C. Dugan, J. Johnson, Q. Pan, W. Zhang, S. Kumaravel, and M. Campbell, "Effects of communication directionality and ai agent differences in human-ai interaction," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3411764.3445256

[15] A. Juliani, V.-P. Berges, E. Vckay, Y. Gao, H. Henry, M. Mattar, and D. Lange, "Unity: A general platform for intelligent agents," *arXiv preprint arXiv:1809.02627*, 2018.