Physiological Synchrony and Collaboration: an empirical inquiry using the Overcooked game

October 23, 2024

Master Media Technology Thesis

By Fatima Mashood
Department of LIACS , MSc Media Technology, Leiden University

Primary supervisor: Maarten Lamers
Department of LIACS , MSc Media Technology, Leiden University

Secondary supervisor: Max van Duijn
Department of LIACS , MSc Media Technology, Leiden University

Physiological Synchrony and Collaboration: an empirical inquiry using the Overcooked game

Fatima Mashood (f.mashood@umail.leidenuniv.nl)
Department of LIACS , MSc Media Technology, Leiden University

Primary supervisor: Maarten Lamers
Department of LIACS , MSc Media Technology, Leiden University

Secondary supervisor: Max van Duijn
Department of LIACS , MSc Media Technology, Leiden University

Abstract

This research examines the relationship between physiological synchrony (PS) and collaboration within dyads engaged in overcooked-a virtual kitchen simulation game. For the study, 14 participants were paired to form 7 dyads based on whether they were familiar or unfamiliar with each other. PS was assessed by analyzing participants' heart rate data, while the dyad's game score assessed collaboration during gameplay. Findings presented a weak yet positive relationship between PS and collaboration, while statistically insignificant. Additionally, familiarity between participants did not significantly affect PS or collaboration. These findings contribute to understanding PS and collaboration in virtual gaming environments while highlighting the influence of familiarity on such research.

Keywords: physiological synchrony; collaboration; overcooked

Introduction

Synchrony manifests in various forms around us, ranging from starlings' coordinated and breathtaking murmurations to the subtle alignment of human behaviours, such as yawning. This coordination of observable actions is referred to as behavioural synchrony, where individuals unconsciously mirror each other's movements and actions during interaction (Denk et al.,). However, synchrony is not limited to external behaviours. It also manifests within us where physiological responses, such as heart rate, may align amidst social interaction. This alignment is referred to as PS and is typically defined as the temporal alignment of such responses between individuals during social interaction, over time (Jonathan L. Helm & Hastings,).

The development of various measurement techniques has facilitated research on PS across various contexts of social interaction (Nazarchuk,). Typically utilizing technologies that measure PS indicators such as heart rate, skin conductance, hormone levels, and brain activity over time (Mønster, Håkonsson, Eskildsen, & Wallot,). Among the various contexts within which PS has been researched, relationship types are a prominent theme observed (Palumbo et al.,). Within this theme, A majority of existing literature explores PS in parent-child relationships (Feldman, Magori-Cohen, Galili, Singer, & Louzoun,), romantic relationships (Levenson & Gottman,), between teammates (Müller & Lindenberger,), as well as in therapist-client (Stratford, Lal, & Meara,) relationships (Mønster et al.,).

While PS has been extensively studied in offline social interactions, an emerging interest has developed in understanding PS in virtual environments, such as virtual game-play environments. Within virtual environments, phenomenons of social interaction such as cooperation, collaboration and competition can be closely examined within controlled settings (Moharana, Keighrey, & Murray,). Additionally, certain games are designed to assess one or more of these phenomena within the game, using performance metrics. Overcooked (Github-HumanCompatibleAI/overcooked-demo,), a virtual kitchen simulation game, is an example of a gaming environment where players must cooperate and collaborate to complete tasks within a time limit (Meimandi, Bolton, & Beling,). In Overcooked the extent of collaboration between players is quantified by the total number of dishes the players serve together, presented as their game score. By studying PS in such contexts, we can assess whether PS between players correlates with the extent to which they collaborate during the game.

This study seeks to address gaps in the existing research by investigating the relationship between PS and collaboration among dyads in a virtual gaming environment. Additionally, the study examines the role of familiarity in influencing PS and collaboration. The research employs a within-dyad design, where heart rate data is gathered using PPG sensors, and collaboration within dyads is measured by their game scores during gameplay. The research hypothesizes that higher PS will be associated with higher game scores, and hence greater collaboration within dyads. The research also hypothesizes that familiar dyads may present with higher PS and higher collaboration during gameplay. The following section will review existing literature on PS within contexts of collaboration, to ground the current study within existing research.

Literature Review

This section presents a concise understanding of existing research on PS, how it manifests within our physiology, its role in cooperative and collaborative contexts, and the methods used to measure it. By examining these areas, this literature review seeks to contextualize the current study on PS in virtual gaming environments within existing literature.

Physiological Synchrony

The biological system within which PS is measured is the autonomic nervous system (ANS), responsible for regulating involuntary bodily functions (Palumbo et al.,). Within the ANS, the sympathetic nervous system (SNS) and parasympathetic nervous system (PNS), regulate stress and relaxation responses within individuals respectively (Palumbo et al.,). The SNS regulates the 'fight or flight' responses, while the PNS regulates the 'rest and digest' responses (Palumbo et al.,). Both these systems work together to regulate such responses within individuals (Palumbo et al.,). Monitoring and aligning these responses, such as heart rate, over time, can help measure PS within two or more individuals (Jonathan L. Helm & Hastings,). Understanding the biological system within which PS is measured, provides insights into how PS manifests within us. PS has been extensively researched within various contexts of social interaction (Palumbo et al.,). For the current study, we will focus on PS research in contexts of cooperation and collaboration.

While both cooperation and collaboration involve individuals working together, cooperation typically refers to working towards a shared goal that may not require mutual engagement (Camarihna-Matos & Afsarmanesh,). In contrast, collaboration is characterized by interaction, mutual engagement, and joint problem-solving (Camarihna-Matos & Afsarmanesh,). We focus on PS in the context of collaboration, specifically within the virtual gaming environment of Overcooked. The game requires not only cooperation, where players perform individual tasks, but also collaboration, where players must constantly communicate, strategize, and solve problems together.

Having reviewed the biological system within which PS manifests, we will now explore existing research on PS within various contexts of cooperation and collaboration. These contexts provide insights into understanding PS during social interactions. Research suggests that PS plays a critical role in social bonding and successful teamwork, making it a key area of interest when studying interpersonal dynamics in both cooperative and collaborative settings (Henning, Boucsein, & Gil,).

PS within cooperative contexts

Within existing research on PS in contexts of cooperation, a research examined the physiological responses of teams engaged in building origami sailboats in an assembly-line manner together (Mønster et al.,). Physiological responses such as participants' heart rate, facial muscle activity, and skin conductance data were gathered. The research sought to assess if higher PS was associated with increased cohesion or bond within the team during a cooperative task (Mønster et al.,). The cooperative task of building origami sailboats was divided among a team of three, with members working towards building the maximum number of boats without communicating with each other (Mønster et al.,). When comparing PS in within individuals who interacted in a team, and those who did not interact as part of a team- higher PS was found among team members (Mønster et al.,). The findings of this study suggest that PS can be a potential indicator of team emotions, and can provide insights into team dynamics (Mønster et al.,). These findings provide important insights for understanding PS within cooperative contexts, suggesting that higher PS is strongly associated with cooperative interaction in teams.

A research on PS in cooperative contexts, explored the relationship between PS and cooperative success, to assess if PS could help predict cooperative success between individuals (Behrens, Moulder, Boker, & Kret,). To assess cooperative success, and how well people work together to achieve a common goal, participant pairs (dyads) were engaged in playing a modified version of the prisoners' dilemma game (Behrens et al.,). During the study, heart rate and skin conductance data, measures of PS, were gathered (Behrens et al.,). Additionally, the researchers examined the influence of visual contact on PS and cooperation by comparing dyads that could see each other during the study with those that could not (Behrens et al.,). The findings revealed that PS, measuring skin conductance levels, predicted cooperative success, and this positive relationship was stronger when participants could see each other and non-verbally communicate (Behrens et al.,). This study provides important insights into the relationship between PS and cooperative success, suggesting a positive relationship between the two. Additionally, the research highlights the influence of communication, even when non-verbal, on PS and cooperative success (Behrens et al.,). Thus, highlighting the need for understanding PS within collaborative contexts, where participants can communicate, and strategize while working towards a shared goal.

PS within collaborative contexts

Within existing research on PS in collaborative contexts, a study was designed to assess PS, between computer science students within a classroom environment (Ahonen et al.,). The students worked in pairs (dyads) on programming exercises, while their heart rate data was collected using lightweight ECG (electrocardiography) sensors (Ahonen et al.,). When comparing PS within the paired and unpaired dyads, higher PS was observed in the paired dyads (Ahonen et al.,). The findings revealed higher PS between the paired students who collaborated, in comparison to those who did not (Ahonen et al.,). This study provides important insights into PS and collaboration, highlighting their positive relationship. This study helps present that working together towards a shared goal, within a collaborative environment can reflect higher PS (Ahonen et al.,). The study also presents the viability of using PS measurement techniques, such as lightweight sensors, within uncontrolled environments such as classrooms (Ahonen et al.,). While this research, focuses on assessing the relationship between PS and collaboration within an offline context, it highlights the need for understanding PS within virtual contexts of collaboration.

PS within virtual collaborative contexts

Research assessing PS within virtual collaborative contexts presents relevant insights into PS and collaboration. One such study examined PS between dyads engaged in a collaborative virtual reality (VR) task (Moharana et al.,). Paired participants were engaged in solving a puzzle within a Virtual Reality (VR) environment, while taking on either the role of a leader or follower, with participants interacting within the same virtual environment but from separate rooms (Moharana et al.,). The findings revealed that higher PS correlates with better team performance, indicative of enhanced mutual understanding and coordination (Moharana et al.,) between individuals. These findings suggest that measuring PS, specifically heart rate alignment, can provide insights into how well participants collaborate in virtual environments. Additionally, these findings highlight the potential of using physiological data to enhance team performance in virtual settings. This is relevant for understanding the role of PS in designing better team-based tasks, and improving collaborative strategies. Furthermore, the research presents that external factors such as participant demographics and familiarity between team members, play a role in determining how well participants synchronize, highlighting the importance of shared experiences and background for successful collaboration in virtual environments (Moharana et al.,). This study provides important insights for understanding PS within virtual collaborative contexts and highlights the need to understand the influence of factors such as familiarity amidst PS research.

Influence of familiarity on PS research

Although familiarity and its influence on PS research remain mostly unexplored, the limited existing research suggests that familiarity can potentially induce PS. In Research assessing PS in a group decision-making task, the lack of familiarity between participants prior to the experiment was addressed as an important limitation of the study (Gordon et al.,). Additionally, research examining the extent of interaction between participants, found that stronger social interaction induced higher PS between participants (Rezaei,). The researchers also observed a stronger correlation between intimate participants in comparison to strangers, presenting the importance of assessing familiarity within PS research. The next section overviews certain measures of PS and collaboration within existing research on PS to contextualize the research design for this study.

Heart rate as an indicator of PS

Within existing research on PS, various physiological indicators, such as heart rate, skin conductance, brain activity, facial muscle activity have been used to measure PS, as discussed above. Among these, heart rate is a common indicator, used to asses synchrony or alignment of physiological states during social interaction (Gordon et al.,). Existing research suggests that the alignment of the heart rate, specifically the inter-beat intervals (IBIs) alignment, is associated with positive elements of social interaction such as better communication, emotional connection (Gordon et al.,).

Various heart monitoring instruments have been used within PS research, including electrocardiography (ECG), and photoplethysmography (PPG). When using an ECG, electrodes need to be positioned, precisely, on the participants' skin. This can be invasive and additionally hinders movement between participants (?, ?). As a result, PPGs are more feasible, within research, as they are less invasive and can be easily worn as wearables (?, ?). Different approaches are utilized to process heart rate data. One such approach was used to measure PS in a group drumming tasks (Gordon et al.,). During the experiment, participants were randomly assigned either synchronized or asynchronized drumming tasks (Gordon et al.,). Heart rate data was collected using an ECG recording device, and processed to extract IBI data in time series (Gordon et al.,). The IBI time series data was aligned and analyzed utilizing the cross-correlation function (CCF) technique to measure group synchrony (Gordon et al.,). The cross-correlation function is frequently used to assess PS between two or more individuals (Gordon et al.,). By utilizing the CCF function to assess IBI data, the degree of synchrony between the heart rates of individuals over time, can be determined (Gordon et al.,). These findings, from existing research, provide important considerations for selecting the instrument, and data processing technique to measure PS.

Using Overcooked game to measure collaboration

Overcooked is a collaborative kitchen simulation game where players work together to prepare and serve dishes under time constraints (Meimandi et al.,). The game requires players to communicate, coordinate, and collaborate effectively to achieve high scores. Within the game, participants progress through game orders, completing various tasks ranging from cooking ingredients to serving dishes (Github-HumanCompatibleAI/overcooked-demo,). Figure 1 shows a game order (layout of the game) within overcooked, with objects within the game, described. The game is played by two players, presented as two chefs; the primary and secondary players of the game. Within the game, the chefs are required to collect ingredients, either onions or tomatoes,and serving plates from their distinct dispensers, cook the ingredients in the pot, plate the dish a subsequently deliver the dish to the serving area. The game data, shown below the game order, shows the players game score, the remaining time left to complete the game, to help the players track these metrics. For the purpose of this research, overcooked was selected to measure collaboration within a virtual gameplay environment. The game scores, quantified by the total number of dishes served by the players, is used as a metric for collaboration. By assessing the dyads game scores and PS during the gameplay, this research research seeks to explore the relationship between PS and collaboration within a virtual gaming environment such as Overcooked.

This literature review sought to overview existing research on PS, relevant to the context of this study, to contextualize

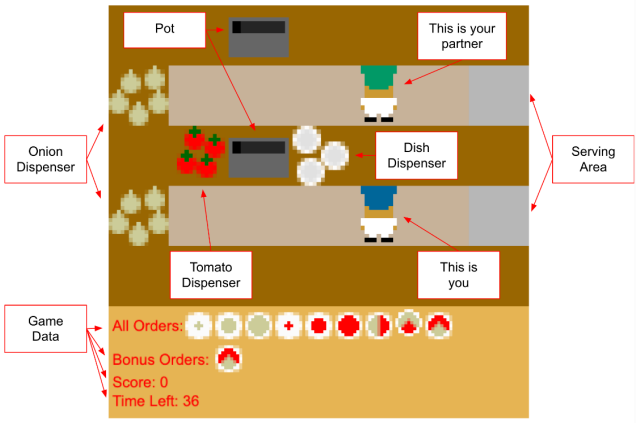

Figure 1: Instructions for the Overcooked game, within one of the game orders of the game, with game objects labeled

the currents study on PS in virtual gaming environments. We started with a brief understanding of PS within our biological system, followed by an overview of existing research on PS within contexts of social interaction. For the purpose of this research, we focused on PS within contexts of cooperation, collaboration, and virtual collaborative environments. The review highlights the emerging interest in understanding PS within virtual environments, like gameplay. Additionally, we highlighted the influence of familiarity within existing research on PS. The review then presents an overview of using heart rate as an indicator of PS, followed by an overview of using Overcooked as a measure of collaboration for the context of this research. The following section seeks to develop the research statement for this research.

Research statement

Existing research on PS, typically investigating PS within offline environments, proposes a positive correlation between PS and collaboration. Less is known about this relationship within virtual environments, specifically within virtual gaming environments. Additionally, the influence of familiarity on PS, especially within collaborative contexts is less researched. As a result, this research seeks to examine the relationship between PS and collaboration within dyads in a virtual environment using the game overcooked. The research question seeks to assess the correlation between PS, measured by heart rate alignment, and collaboration, measured by game scores. The study hypothesizes that a positive relationship between PS and collaboration will be observed, with higher PS associated with greater collaboration, indicated by higher game scores. Additionally, this research seeks to assess the influence of familiarity within dyads, on PS and collaboration, hypothesizing that familiar dyads will present with higher PS and greater collaboration within dyads.

Primary research question: Does physiological synchrony, measured by heart rate alignment, correlate with collaboration within dyads in a virtual gameplay environment?

Methodology

This section outlines the methodology designed and utilized to examine the relationship between PS and collaboration, within a virtual gaming environment. The methodology section seeks to present the research design, procedure and methods utilized for data collection and analysis. For this study, physiological indicators of PS utilized focus primarily on heart rate measures. Collaboration within dyads is assessed within Overcooked, presented as the dyad's game score for each game order.

Research design

The research design for this study employs a quasi-experimental design to examine PS and collaboration within a virtual gameplay environment. The study design engages dyads (pairs of participants) in playing Overcooked, while their heart rate data is gathered. The participants were paired to form dyads based on whether they were familiar or unfamiliar with each other before the experiment. During the experiment, each dyad engaged with 6 game orders with differences in the way the layout of the game order. For instance the number of ingredients (onions, or soup, or both), access to the serving area or pot, may vary within the layout of different game orders. The order at which the dyads played the game orders (1,2,3,4,5,6) was kept consistent for all dyads, and did not increase in difficulty as the game progressed. Collaboration is assessed within Overcooked as the dyadic game score during gameplay. While the participants are engaged in gameplay, heart rate data, specifically IBI data, is collected from each participant in the dyad using a PPG finger sensor. The data is processed and assessed, and cross-correlated to measure PS within dyads. The cross-correlated IBI data, and game scores are analyzed to assess the relationship between PS and collaboration within the dyads.

Participants

The participants for this study were recruited utilizing a convenience sampling method. While the sampling technique may induce bias, it was selected to ensure that a maximum number of participants were engaged in the study. Additionally, this technique was utilized to pre-determine the pairs based on whether they were familiar or unfamiliar with each other. A total of 14 participants were recruited for this study, formed into 7 dyad pairs, of 4 unfamiliar dyads, and 3 familiar dyads. Given the time constraints of a master's thesis project, a larger sample size could not be considered for this study.

Experiment setup

The game environment for Overcooked was built using Docker, referencing the overcooked demo code documentation (Github-HumanCompatibleAI/overcooked-demo,). Custom APIs were developed for the sensor server to control the start and stop of reading and logging sensor data, followed by minor modifications to the overcooked game for

integration. This approach was employed to enable participants to play on two different URLs while connected to the same network. For each experiment, start time, stop time, game order data, as well as game score, were logged by the sensor server, as seen in Figure 2, which was invoked by the Overcooked game's user interface. Along with this, the sensor server logged the heart rate and IBI data for each dyadic experiment and game order, making it convenient to perform consecutive game orders within the experiment.

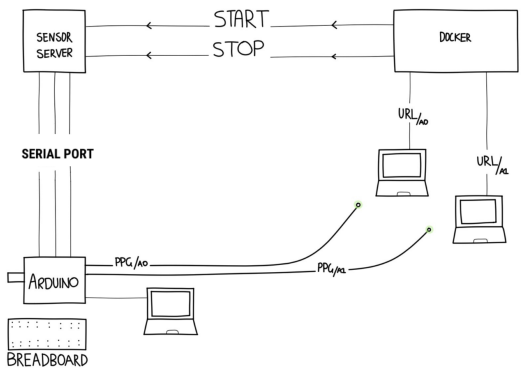

Figure 2: The figure presents an illustration of the setup of the gaming environment, showing the arrangement of participants' laptops, Arduino and breadboard, two ppg sensors, as well as the game server in docker.

Figure 3 shows a schematic illustration of the experiment setup. The PPG sensors, A0 and A1, are connected to the Arduino setup and are attached to the participants during the experiment. The Arduino communicates with the Sensor Server via the serial port, sending heart rate data. The Docker environment runs the game on two distinct URLs, with (URL A0) for the participant wearing PPG sensor A0 and (URL A1) for the participant wearing PPG sensor A1.

Procedure

When recruiting participants, an information sheet along with consent forms were shared with potential participants approached. Upon receiving signed consent forms, the participants were invited for the experiment and asked to bring their laptop. Upon their arrival, the dyads were seated across from each other and asked to set up their laptops. To be able to distinguish between the heart rate readings of the two participants within the dyad, identifier tags, were placed next to their laptop, identifying the participant as either A0 or A1. The dyads were introduced to each other and left alone for a few minutes, to ensure that the participants were acquainted enough to be able to coordinate and communicate during the experiment.

The participants were then given an introductory briefing, revisiting the purpose, design, procedure, and ethical considerations for the study. For each experiment within this study,

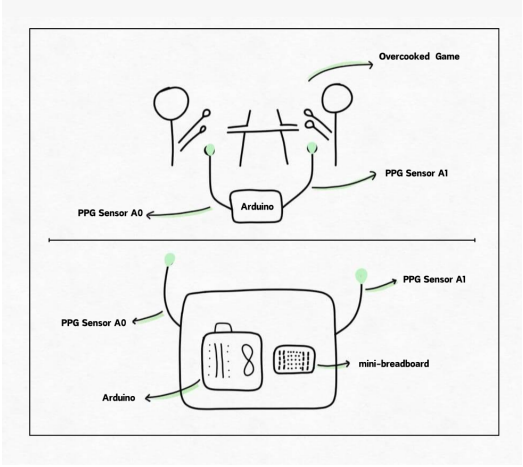

Figure 3: Illustrates the experiment setup showing the sensors and the participants connected to the Arduino

each participants was given a demonstration on how to correctly position and secure the PPG finger sensor, and the participants were asked to secure the sensor themselves, with assistance provided upon request. The heart rate readings for both participants were calibrated for 3-5 minutes, to ensure accurate baseline measurements before starting the experiment. While the data was being calibrated, the participants were given instructions on how to play overcooked. The participants were shown the instruction page within overcooked, as shown in Figure 1, to acquaint them with how to play the game. To ensure that the participants understood the instructions of the game well, the participants were engaged in a test round of one of the game orders, while their heart rate data was logged. Following the test session, the participants were approached to begin the experiment.

After completing all game orders, a post-experiment questionnaire, developed in Qualtrics, was shared with the participants. Typically completed within 3-4 minutes, the questionnaire obtained demographic details such as age, gender, as well as familiarity within the dyads. The dyad's heart rate, real-time IBI data, and game score for each game order were anonymized, stored, and, processed for analysis. To assess the relationship between PS and collaboration, heart rate data, game score data and questionnaire data were collected, the following section seeks to address the data analysis techniques utilized to analyze the gathered data.

Data Analysis

A number of steps were taken to process the data gathered. Missing values in the raw IBI data collected, were replaced with their next or previous valid value from the data frame, to ensure that the data is consistent. The data is filtered by setting a frequency range of of 0.5-5 Hz, removing noise from the data while retaining valid frequencies, as suggested by (Alhawari, Albelooshi, & Perrott,). The cross correlation function is applied to the now cleaned and pre-processed

heart rate data, correlating/aligning the IBI values of both participants within the dyads to assess PS, as suggested by (Gordon et al.,). This statistical measure, utilized within existing research, is used to assess PS between two or more individuals (Gordon et al.,). The cross-correlation function was applied, correlating the IBI values, at all possible time lags yielding cross correlated IBI values as well as the maximum cross correlated IBI values for each dyadic experiment.

Results

This research explored the relationship between PS and collaboration within dyads in a virtual gaming environment using the game overcooked. This section seeks to present the results, covering descriptive and inferential results obtained from the 7 experiments conducted for the study.

Descriptive statistics

A total of 14 participants, paired to form 7 dyads, participated in this study. Within the sample, the age of participants ranged from 23-28 years, with an average age of 26 years. When analyzing the age difference within dyads, a limited age difference was observed, with an average age difference of 2 years. Overall the maximum age difference within dyads was 4 years, observed in 3 dyads. The participant sample consisted of 7 male participants, and 7 female participants, presenting an equal distribution of male and female participants. When dyadic-gender was assessed, only 2 dyads were same-gender pairs, while 5 of the dyads were mixed-gender pairs. The participant sample consisted of 4 unfamiliar dyads, and 3 familiar dyads. There was no missing data within the experiment data, with a total of 42 data points.

The average maximum IBI correlation was observed to be 0.761 with a high maximum correlation of 1.873, suggesting variability between the dyadic IBI correlations. The lowest game score observed was 20, while the highest game score was 495, presenting high variation between the game scores of different dyads. The high standard deviation of 97.031 suggests that the participants performed quite differently from each other.

Inferential Results

Correlation of game order, game score, maximum IBI correlations To assess the correlation between game order, game scores and, the maximum IBI correlations, the Pearson's' correlation analysis was conducted. Table presents the results including the sample size (n), Pearson's r values, and p-values for each dyad. The findings for game order, show a weak yet positive correlation with game scores, with an r-value of 0.133, and a p-value of 0.401. Additionally, game order presents a weak and negative correlation with a maximum correlation of IBI, with an r-value of -0.077, and a p-value of 0.628. When assessing the correlation between collaboration and PS, quantified by game score and maximum IBI correlations respectively, a weak yet positive correlation was observed with an r value of 0.145. The p-value of 0.361 indicates that the correlation is not statistically significant. Observing this linear correlation between game score, a measure of collaboration, and maximum IBI correlations presents that as the game score increases, reflecting the direction, but modestly(weak), reflecting the strength of the relationship. These relationships are visually represented in Figures 4, 5, and 6.

			n	Pearson's r	p
game_order	-	game_score	42	0.133	0.401
game_order	-	max_corr_ibi	42	−0.077	0.628
game_score	-	max_corr_ibi	42	0.145	0.361

Table 1: Pearson's correlation coefficients between Game Order, Game Score, and Maximum Correlations of IBI

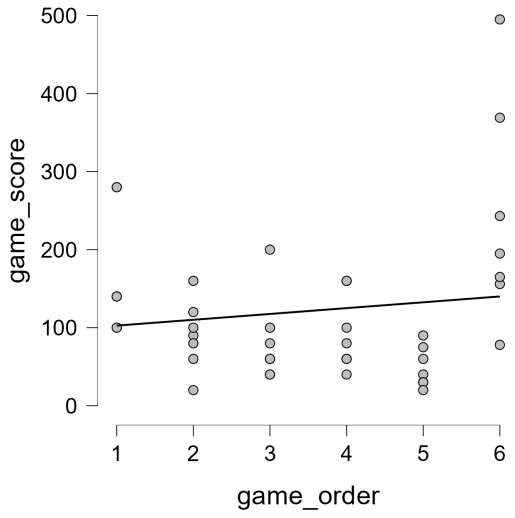

Figure 4: Distribution of Game Order vs Game Score

Effect of Familiarity on Game Score and Maximum IBI Correlations To assess the influence of familiarity on the game score and maximum IBI correlations, an independent samples t-test was conducted, the results for which are displayed in table . The results for the influence of familiarity on maximum IBI correlations present with a t value of 0.555, p value of 0.582, Cohen's d value of 0.173, and SE Cohen's d value of 0.313. Additionally, assessing familiarity and game scores, presents a t value of 0.493, p value of 0.624, Cohen's d value of 0.154 and, SE Cohen's d value of 0.313. These findings suggest that familiarity does not significantly affect the maximum IBI correlations or game scores. The shapiro-wilk test, results for which are displayed in table shows that the data for maximum correlation of IBI and game score are normally distributed for the "false" familiarity condition, while, for the "true" familiarity condition, the data is not normally distributed. Figures, s7 and 8, present bar plots comparing the maximum correlation of IBI and game scores based on familiarity, respectively.

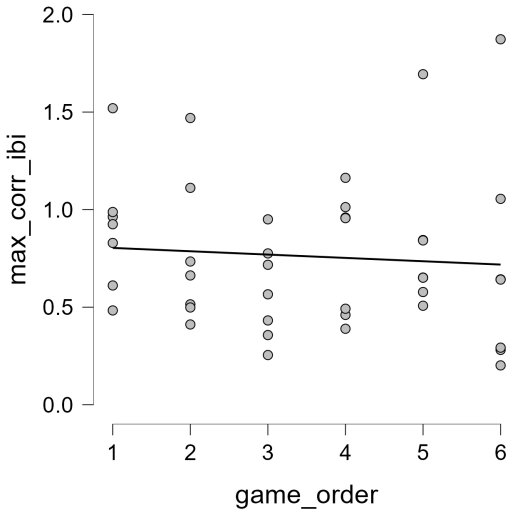

Figure 5: Distribution of Game Order vs Maximum Correlation of IBI

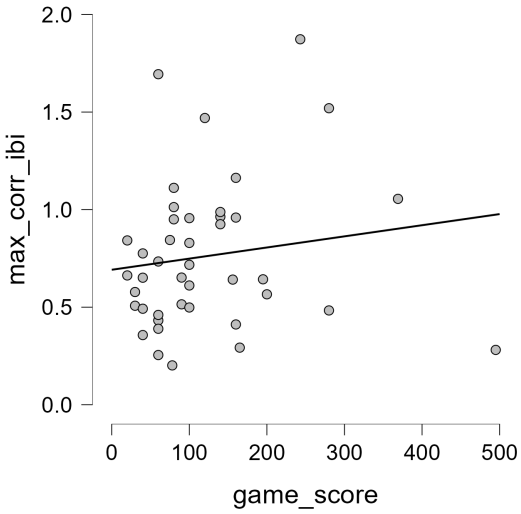

Figure 6: Distribution of Game Score vs Maximum Correlation of IBI

	t	df	p	C.d	SE C.d
max_corr_ibi	0.555	40	0.582	0.173	0.313
game_score	0.493	40	0.624	0.154	0.313

*C.d=Cohens' D

Table 2: T-test for Familiarity on Game Score and Maximum IBI Correlations

		W	p
max_corr_ibi	no	0.946	0.368
	yes	0.882	0.009
game_score	no	0.903	0.066
	yes	0.724	$< .001$

Table 3: Test of Normality (Shapiro Wilk) on Maximum IBI Correlations and Game Score

Effect of dyad gender on PS To assess the effect of the dyads' gender on PS, reflected by the maximum IBI correlation, an ANOVA test was carried out. The results, as shown in table , present an F value of 6.892, and a P value of 0.003. These findings indicate that dyad gender has a statistically significant effect on the maximum IBI correlations, PS. When comparing mixed-gender dyads and same-sex dyads, as shown in Figure 9, mixed-gender dyads present with the lowest maximum IBI correlations.

Effect of dyad gender on game scores The ANOVA test was also conducted to assess the effect of dyads' gender on collaboration within dyads, quantified as game scores. The results, as shown in present an F-value of 1.372 and a p-value of 0.266. This suggests that dyad gender does not have a statistically significant effect on game scores. This finding is further supported by the descriptive plot 10), which shows no substantial difference in game scores between different gender compositions within dyads.

Effect of Dyad Gender Equality on IBI and Game Scores To assess the effect of dyads gender equality on PS and collaboration, quantified as maximum IBI correlations and game scores respectively, a t-test was conducted. The results, as shown in , for maximum IBI correlations present a t value of 3.027, a p-value of 0.004, Cohen's d of 1.034. This is indicative of higher synchrony in same-gendered dyads. The results for game scores present a t-value of 1.247, p-value of 0.219, and Cohen's d-value of 0.426, revealing that no significant effect of gender quality was found on game scores. Additionally, the Shapiro-Wilk test was assessed to check for normality. The results confirmed normality for maximum IBI correlations but indicated deviations for game scores in one group. Findings presented by the bar plot 12 of game score and dyad gender equality (same-sex gender) presents as true, indicating a higher average game score than the same-sex gender dyads. Additionally, findings presented by the bar plot, 11 of maximum correlated IBI and dyad gender equality, (same-sex gender) present as true, suggesting that same-sex dyads have a significantly higher average maximum IBI correlations.

Discussion

This section seeks to discuss relevant insights from the results of the study while developing implications for future work.

Correlation of Game Order, Game Score, and Maximum IBI Correlations

Results from Pearson's correlation analysis presented a weak yet positive correlation with game scores. and a weak negative correlation with maximum IBI correlations. Additionally, the correlation between game scores and maximum IBI correlations was weak yet positive. These findings indicate

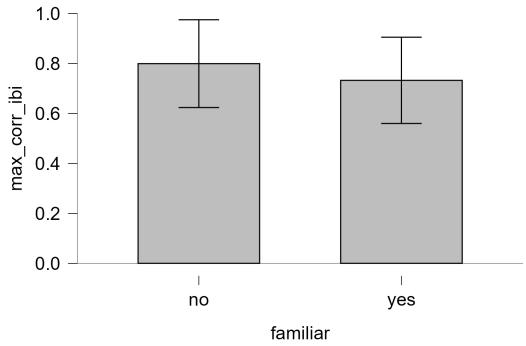

Figure 7: Bar plot showing comparison of Maximum Correlation of IBI based on Familiarity with lines depicting standard deviation, shows limited effect of Familiarity

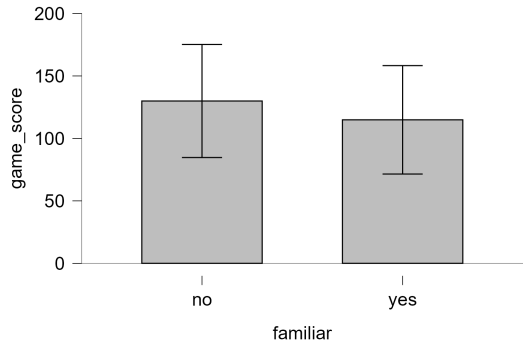

Figure 8: Bar plot showing comparison of Game Score based on Familiarity with lines depicting standard deviation, shows limited effect of Familiarity

that game order had limited influence on collaboration within dyads, and PS respectively. This weak yet positive relationship assessed in the context of this study, aligns with the hypothesis that higher PS is associated with greater collaboration, even though the relationship is not statistically significant.

Effect of familiarity on PS and collaboration

Results from the t-tests on the influence of familiarity within dyads do not significantly affect PS or collaboration, measured by maximum IBI correlations and game scores, respectively. However, this may be because assessing familiarity between participants as a binary (familiar, unfamiliar)limits how little much influence it may have on this study.

Effect of dyad gender on PS and collaboration

Although not a prior aim of the research effect of dyad gender on the maximum IBI correlation, and game scores were analyzed to measure PS and collaborative respectively. The results suggest a statistically significant effect of dyad gender on PS. Specifically, mixed-gender dyads exhibited lower PS compared to same-gender dyads. This suggests that dyad gender plays a role in influencing PS, with mixed-gender

Cases	Sum of Sq	df	Mean Sq	F	p
dyad_gender	1.570	2	0.785	6.892	0.003
Residuals	4.443	39	0.114		

Table 4: The ANOVA results for Maximum IBI Correlations, based on Dyad Gender (M = both male (N= 6), F = both female (N= 6), D = both different gender (N=30))

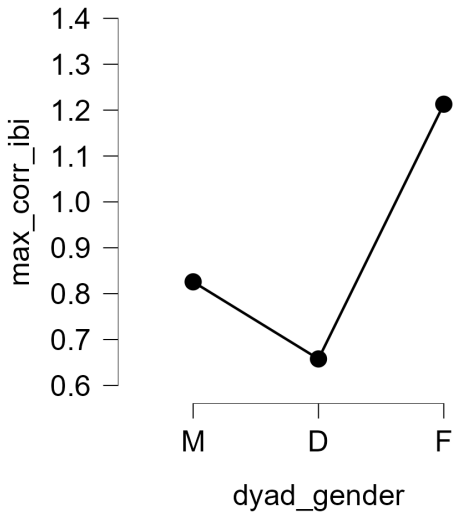

Figure 9: The ANOVA result plot for Maximum IBI Correlations, based on dyad gender (M = both male (N= 6), F = both female (N= 6), D = both different gender (N=30))

pairs showing lower PS. However, the effect of dyad gender on collaboration, measured by game scores, was not statistically significant. The ANOVA test results indicate that there is no substantial difference between game scores when comparing dyad genders within this study. This suggests that while dyad gender may influence PS, it does not seem to present a substantial difference in how the participants collaborate in this study. However, t-test results on dyad gender equality (same or different gender of participants within dyads) on PS and collaboration revealed that same-gender dyads presented with higher maximum IBI correlations compared to mixed-gender dyads. However, the effect of gender equality on game scores was not statistically significant, indicating that same-gender dyads did not collaborate as effectively.

Limitations and future work

This section seeks to overview the known limitations of the study for this research. The limitations section, presents certain limitations, regarding the study design, while indicating avenues of change for future work on related research.

The sample size utilized for the study was limited due to the time constraints of a master's thesis project. For this study, it was important to keep the game's UI simple and user-friendly game, so that no substantial experience with gaming

Cases	Sum of Sq	df	Mean Sq	F	p
dyd_gnd	25375.800	2	12687.900	1.372	0.266
Residuals	360639.533	39	9247.168		

Table 5: The ANOVA results for Game Score, based on Dyad Gender (M = both male (N= 6), F = both female (N= 6), D = both different gender (N=30))

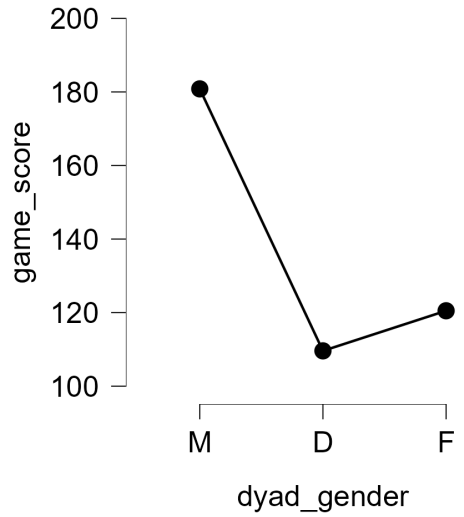

Figure 10: The ANOVA results plotted for Game Score, based on Dyad Gender (M = both male (N= 6), F = both female (N= 6), D = both different gender (N=30))

is required from participants. Additionally, utilizing a more complex version of the game, that includes more objects, can potentially induce stress-which would influence the heart rate data. As a result, the time duration of each game order was limited to 60 seconds each.

With regard to the measuring techniques used, PPG sensors may not be as precise as using LAB devices for measuring PS, as done in certain related studies. Although, the participant's heart rate data was calibrated before each experiment for 3-5 minutes, potential invalid data points could have hindered the reliability of the Heart rate readings. An important avenue for future work to explore is to assess multiple physiological responses, such as heart rate, skin conductance, eye-tracking, and neural activity, simultaneously to verify the data collected. Several limitations of the study design have been considered, which may have potentially influenced the results, along with certain suggestions for future work.

Conclusion

This research explored the relationship between PS and collaboration within dyads in a virtual gaming environment using the game overcooked. The findings present a weak yet positive correlation between PS and collaboration, although not statistically significant. Additionally, mixed-gender dyads presented lower PS compared to same-gender dyads, but dyad gender did not significantly affect collaboration. Familiarity between participants did not significantly impact PS or collaboration. These results highlight the need for further research with larger sample sizes and multiple indicators of PS to validate these findings.

	t	df	p	C.d*	SE C.d*
max_corr_ibi	3.027	40	0.004	1.034	0.402
game_score	1.247	40	0.219	0.426	0.352

Table 6: Independent samples t-test for Dyad Gender Equality on Maximum IBI Correlations and Game Scores. *C.d=Cohens' D

		W	p
max_corr_ibi	TRUE	0.929	0.373
	FALSE	0.957	0.252
game_score	TRUE	0.946	0.575
	FALSE	0.718	$< .001$

Table 7: Test of normality (Shapiro-Wilk) for Dyad Gender Equality

References

Ahonen, L., Cowley, B., Torniainen, J., Ukkonen, A., Vihavainen, A., Puolamäki, K. (2016). Cognitive collaboration found in cardiac physiology: Study in classroom environment. *PloS one*. doi: 10.1371/journal.pone.0159178

Alhawari, M., Albelooshi, N. A., Perrott, M. H. (2014). A 0.5 v lt;formula formulatype=”inline”gt;lt;tex notation=”tex”gt;*amp*;*lt*;4 μlt;/texgt;lt;/formulagt;w cmos light-to-digital converter based on a nonuniform quantizer for a photoplethysmographic heart-rate sensor. *IEEE Journal of Solid-State Circuits*, *49*, 271-288. doi: 10.1109/jssc.2013.2284349

Behrens, F., Moulder, R., Boker, S. M., Kret, M. E. (2020). Quantifying physiological synchrony through windowed cross-correlation analysis: statistical and theoretical considerations. *random*. doi: 10.1101/2020.08.27.269746

Camarihna-Matos, L. M., Afsarmanesh, H. (2008). Concept of collaboration. *random*. Retrieved from `https://api.semanticscholar.org/CorpusID:60871446`

Denk, B., Dimitroff, S., Meier, M., Wienhold, S., Gaertner, R., Klink, E. S. C., ... Pruessner, J. C. (2024, Jan). *Physiological synchrony of the autonomic nervous system - an analysis and comparison of different methods.* PsyArXiv. doi: 10.31234/osf.io/dws8c

Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., Louzoun, Y. (2011). Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behavior and Development*, *34*(4), 569-577. doi: https://doi.org/10.1016/j.infbeh.2011.06.008

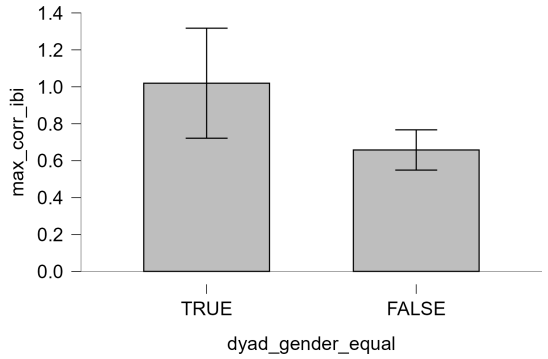

Figure 11: Bar plot representing Maximum IBI Correlations based on Dyad Gender Equality, with true representing same-gender dyads, and false representing different-gender pairs

Figure 12: Comparison of Game Scores Based on Dyad Gender Equality

Github-HumanCompatibleAI/overcooked-demo. (n.d.). *GitHub - HumanCompatibleAI/overcooked-demo: Web application where humans can play Overcooked with AI agents. — github.com.* https://github.com/HumanCompatibleAI/overcooked-demo. ([Accessed 26-09-2024])

Gordon, I., Gilboa, A., Cohen, S., Milstein, N., Haimovich, N., Pinhasi, S., Siegman, S. (2020, May). Physiological and Behavioral Synchrony Predict Group Cohesion and Performance. *Scientific Reports*, *10*(1), 8484. doi: 10.1038/s41598-020-65670-1

Henning, R. A., Boucsein, W., Gil, M. C. (2001). Social–physiological compliance as a determinant of team performance. *International Journal of Psychophysiology*, *40*, 221-232. doi: 10.1016/s0167-8760(00)00190-2

Jonathan L. Helm, S. K. N. R. T., Jonas G. Miller, Hastings, P. D. (2018). On measuring and modeling physiological synchrony in dyads. *Multivariate Behavioral Research*, *53*(4), 521–543. doi: 10.1080/00273171.2018.1459292

Levenson, R. W., Gottman, J. M. (1985). Physiological and affective predictors of change in relationship satisfaction. *Journal of Personality and Social Psychology*. doi: 10.1037/0022-3514.49.1.85

Meimandi, K. J., Bolton, M. L., Beling, P. A. (2024, April). Action over words: Predicting human trust in ai partners through gameplay behaviors. doi: 10.36227/techrxiv.171198000.01201027/v1

Moharana, B., Keighrey, C., Murray, N. (2023, June). Physiological synchrony in a collaborative virtual reality task. In *2023 15th international conference on quality of multimedia experience (QoMEX)* (pp. 288–293). IEEE.

Mønster, D., Håkonsson, D. D., Eskildsen, J., Wallot, S. (2016). Physiological evidence of interpersonal dynamics in a cooperative production task. *Physiology Amp; Behavior*, *156*, 24-34. doi: 10.1016/j.physbeh.2016.01.004

Müller, V., Lindenberger, U. (2011). Cardiac and respiratory patterns synchronize between persons during choir singing. *PLoS ONE*. doi: 10.1371/journal.pone.0024893

Nazarchuk, I. (2020). *Heart rate variability analysis within a data-enabled design research process* (Master's Thesis, Eindhoven University of Technology). Retrieved from https://research.tue.nl/en/studentTheses/heart-rate-variability-analysis-within-a-data-enabled-design-rese

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O. H. G., McGee, H., Liu, S., Goodwin, M. S. (2016). Interpersonal autonomic physiology: a systematic review of the literature. *Personality and Social Psychology Review*, *21*, 99-141. doi: 10.1177/1088868316628405

Rezaei, S. (2013, July). Physiological Synchrony as Manifested in Dyadic Interactions [Master's Thesis]. Retrieved from https://www.semanticscholar.org/paper/Physiological-Synchrony-as-Manifested-in-Dyadic-Rezaei/35c76ed8be79dc09081125209025aa0d17f833cf

Stratford, T., Lal, S., Meara, A. (2012). Neuroanalysis of therapeutic alliance in the symptomatically anxious: the physiological connection revealed between therapist and client. *American Journal of Psychotherapy*. doi: 10.1176/appi.psychotherapy.2012.66.1.1