Sharing human mental model with an AI agent to achieve team effectiveness

Ziad Nawar[1] , Ruben Verhagen[1] , Carolina Jorge[1] , Dr Myrthe Tielman[1]
[1]TU Delft

Abstract

The collaboration between AI agents (Artificial Intelligence) and human is an essential part of achieving complex goals more efficiently. Many aspects are influential in achieving effective teamwork. One of them is trust. In addition, sharing the mental model would improve the understanding of the other's behavior and the prediction of their actions. In this paper, we will analyze the influence of sharing the mental model on team performance. We will consider the human side's trust in an AI agent under various shared mental model structures.

1 Introduction

Humans and digital computers have worked together for decades to get things done efficiently. These tasks varies from simple calculations and data storage to self-driven vehicles, intelligent robots, and many other applications with intelligent algorithms. Matthew Johnson and Alonso Vera demonstrate the importance of combining artificial intelligence (AI) agents with humans to achieve high team effectiveness [1], focusing on human-AI agents team and the factors that affect their performance. An important factor influencing team performance is a shared mental model [2]. Many researchers have presented various mental models [3; 4]. Intuitively, team performance increases with higher sharedness in the shared mental model about their beliefs and intentions[3].

In addition, Jonker et al. show that under certain circumstances, some components of the shared mental model are more important than others for higher team performance[4]. That paper focused on the AI agent - AI agent teaming. However, in this paper, a similar experiment will be conducted to analyze the effect of different shared mental models on the human-AI agent teaming. We hypothesize that the more information in the mental model that is exchanged between the human-AI agent team, the higher the performance.

In our study, we focus on the research question that is how the human should share his mental model with the agent to achieve high team effectiveness. Section 2 is devoted to the subdivision of the research problem and the analysis of each sub-problem. The experimental design is shown in Section 3.In section 4 we show the results and the statistical analysis we ran on the data. Section 5 will reflect the ethical aspects of this research. A discussion that reflects on the experiment outcomes and future work recommendation is in section 6. Section 7 will include the conclusion of this paper.

2 Methodology

This research is focused on determining the impact of sharing the human mental model with the AI agent on the team performance. We will start by formulating a shared mental model, then how to determine the team effectiveness. In addition, we will look into how to measure the trust of the human in the agent with the same type of mental model.

2.1 Mental Model

As per the hypothesis, the focus is on the mental model of the human inside the human and AI agent teaming. As per Jonker et al. mentioned a definition for the mental model of a human which is ”internal representations of the world around them, that help them to understand, explain and predict the systems in their environment” [3]. As a human in any situation would collect information and use it to take decisions or understand the situation. In this paper, the focus is on the teaming of the human and AI agent which leads simply to the next paragraph about shared mental models.

Shared Mental models can be defined as ”knowledge structures held by members of a team that enable them to form accurate explanations and expectations for the task, and, in turn, coordinate their actions and adapt their behavior to demands of the task and other team members” [5]. Previous research has described different mental models and their effect on the performance of a team of AI agents [3; 5]. The mental model that is going to be used in this paper is inspired from [3] which is a task-based mental model that focuses on finishing a task while sharing information. This information can be classified into two main parts, world knowledge information and intentions information. World knowledge information focuses on the state of the world that the team is working in, on the other hand, intention information focuses on the intentions of the players that will change the world state. These two components will be used in the experiment which has been also used before in the experiment conducted by Jonker et al. [3]. We are aiming to see if the results from our experiment would have the same trend as in the experiment Jonker et. al. conducted.

2.2 Team effectiveness

Measuring the actual effect of the shared mental model on the team effectiveness is quite challenging as multiple confounding factors can affect the time taken to finish the task. A few of these factors could be the participant's background knowledge about the software that will be used in this case MATRX and the human's focus during the task. In order to avoid these problems, the participants will be asked about their familiarity with the software in the questionnaire and for the latter, the participant will be asked to team up with the agent to finish the game as fast as possible, which we believe will increase the participant's focus. The best indicator to evaluate the performance is the time taken to finish the task. In addition, we will use the number of messages sent across the communication channel between the human and the agent. We will include both the average time taken to finish the game and the average number of messages per experiment in our team effectiveness evaluation. The best shared mental model should have the lowest time taken and the lowest number of messages sent.

2.3 Trust

Research about trust between humans and AI agents has been developed in multiple resources as in [6; 7]. The trust of the human in automated agent affects the way the human would deal with the agent which affects the team effectiveness [7]. Measuring trust has been an ongoing research process, Lyons et al. have come up with antecedents that relate to the trust of the human in AI agents, all of these researches are summarised in [6]. There seems to be an overlap of trust factors among researchers which has been developed in multiple questionnaires shown by Hoffman et al. as there can be multiple scales to measure trust [8]. The Trust scale for XAI in that paper would identify the amount of trust of the human in the agent after finishing the experiment. This scale was created by the authors on two main parts which are validity and reliability and we believe it is the most reliable scale that is in line with the experiment setup. Since it contains questions that are related to the AI agent behavior more than other scales that have been shown. This trust scale will be used in a questionnaire shown in appendix A after the experiment which the user will fill in to reflect on how much did the participant trusted the agent. There are 5 different answers for each question which are strongly disagree, disagree, neutral, agree and strongly agree that range from 1-5 respectively. The trust of the participant in the agent will be the average of all questions. However, question 11 in the questionnaire asks about the wariness in the agent, which is if the participant's response was strongly agree then the participant doesn't trust the agent, therefore, the values of the results for this question will be reversed.

3 Experimental Setup

In this section, we will show the experimental setup and the software that we will use to test our hypothesis. It will include the process of implementing the agent and their strategy also the questionnaire used to test for trust.

3.1 Game setup

The hypothesis we propose involves testing the effectiveness of a team of human-AI agents under four different conditions related to human-agent communication. The four different settings are first, silence, in which the human and agent do not exchange information, secondly human and AI agent only share intentions, thirdly, human and AI agent only share the knowledge of the world, and lastly, both of them share both the intentions and knowledge of the world. Experimenting with the four suggested configurations would result in us knowing how much the human needs to share to achieve high team effectiveness, as explained in section 2.2. The measurements that will be analyzed are going to be the time, the number of messages sent by the AI agent, and the number of messages sent from the human. Moreover, the questionnaire in appendix A will be used to test the human trust in the agent.

The game is executed in the MATRX software with the customized configuration of the block world for the team theme. The game contains 6 different blocks with shapes of a square and a triangle and three different colors red, blue, and green. It also contains twelve rooms. A room has a door that can be opened by pressing '*open the door*' button on the screen or the key 'r' on the keyboard. The blocks are initially in these rooms and some of the rooms are empty. To finish the game the players need to deliver the blocks shown in the goal block area in the order from bottom to top to the endpoint of the map. Several pilot tests were run with the AI agent to determine the number of rooms for the map. The results of these tests showed that the map initially was too small, in which the game would finish in few ticks and the difference won't be big enough to see a trend in different setups. So we chose twelve rooms to scatter the blocks away from each other. In addition, the number of target blocks that were also chosen was four blocks to further extend the playing time and to use the information of the shared world knowledge as part of the full information sharing.

Any player in this game can use chat to communicate with the other players in the game to send and receive messages. This chat allows only text-based communication and avoids typing errors. Messages sent from the participant's side were predefined for all the possible messages needed during this experiment. For more information on the messages sent from the participant's side, see subsection 3.3

3.2 AI agent implementation

The AI agent is implemented in Python and can handle some cases that may arise while playing with a human participant. Dealing with is all possible situations would take much more time than the duration of this research project. Some of the situations the agent can handle is to periodically check the endpoint after a four-room search when the intentions cannot be shared between it and the participant. The mentioned strategy was developed to help the AI agent to know which block needs to be collected after a certain amount of time, as the participant and the AI agent don't share whether or not they delivered a block to the endpoint. Another situation that the agent implementation has covered is if the agent has a block that is needed for the endpoint, it would check whether the previous blocks were placed or not. If the blocks have not been placed, the agent waits until these blocks are in the endpoints. This case can occur when the participant sends the agent a message that they will deliver a certain block to the endpoint that is not the last block. The agent will pick the block that comes after that block and also moves to the endpoint before the human place his own block down. The AI agent waits for the participant to first place his block before placing the block it holds.

The AI agent follows a strategy of searching all rooms in order, starting from room 1 in the upper left corner to room 12 in the lower right corner. A human participant would find this strategy convenient as they would start their search at the place closest to where they started. In the case where the human and the AI agent share their intention to send the room to search, the AI agent receives those messages and save the rooms searched by the human. If the next room that the AI agent will be searching has been already been searched by the human, it will skip this room and pick a room in the sequence the human hasn't sent the message for that they have searched.

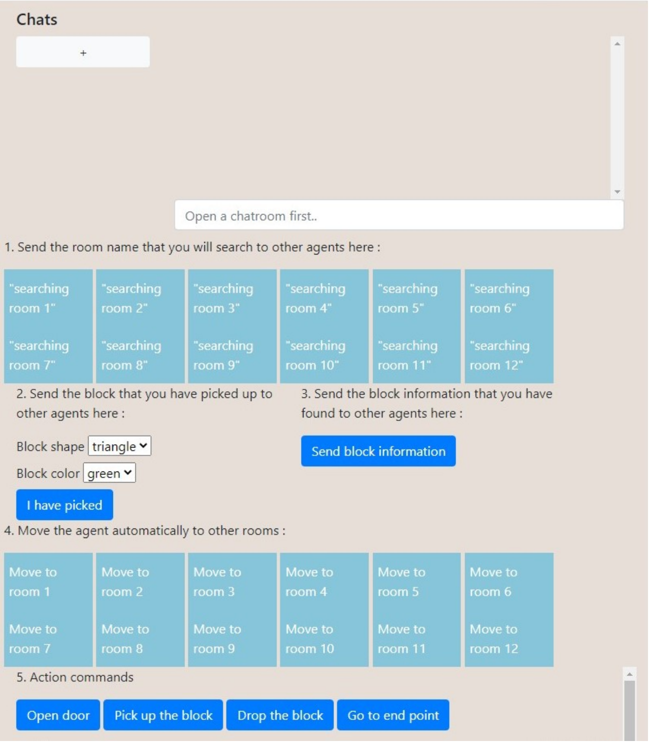
The AI agent can send information about the blocks it sees. The block information includes a block ID with the room location, a specific map location, shape, and color. It only sends information from blocks that have not been placed at the endpoint. After a block has been delivered, the agent no longer sends information about a similar block.

The AI agent would send a message about the next room it will be searching exactly when it has fully finished searching the room. The agent would move directly towards the room and open the door then search each tile of the room as this searching algorithm is dynamic and can be used with any room size.

Also, It can send a message indicating that it has picked up the block that needs to be delivered next and that it would put it down in its place in the endpoint. This message is sent immediately after the AI agent has picked the block. Then it moves towards the endpoint and checks whether the blocks preceding this block have been placed down or waits until they are being placed down.

3.3 Participant side implementation

The participant could only control its own agent and not the AI agent through a set of buttons as shown in figure 1. These buttons have predefined messages that are attached to them. In case the participant is allowed to share intentions as sending the room they will be searching next, they click on one of the buttons in section 1 of the control panel and it will automatically send a message to the agent in the global channel in the chat saying they will be searching this room. In addition, the participant can select which block colour and shape they have picked. The label on the button associated with this actions section will change indicating exactly the contents of the messages would be before sending it and then the participant can click on this message to send it to the agent through the global communication channel. Lastly, It could also send the block information of the blocks visible to the human's agent in the map by pressing on the button *sending block information* and it will send a message with the block id, room location, exact location and shape and colour of a block.



1. Send the room name that you will search to other agents here :

2. Send the block that you have picked up to other agents here :

Block shape triangle

Block color green

3. Send the block information that you have found to other agents here :

4. Move the agent automatically to other rooms :

5. Action commands

Figure 1: Participant control panel. The control panel is divided into 5 sections. Section 1 and 2 are used to send intentions messages. Section 3 is used to send world knowledge messages. Section 4 and 5 are used to control the actions of the agent such as moving around the map, opening a door, dropping and picking up a block.

3.4 Experiment guide

The test participants went through four different stages: First, the participants were asked to fill in the first page of the questionnaire, which contained an information sheet, name, age group, gender, and familiarity with the MATRX software. Second, the experimenter introduces the game to the participant and explains all the rules of the game and the controls , and then the participant is allowed to play a game without artificial intelligence in the world to adapt to the game itself and ensure that the participant understands the game and the agent they are playing with. Third, another game is initialized with a different map where the AI agent presents, as shown in figure 2. Participants are required to play with the AI agent, but at this point. The experimenter should not talk to participants, in order not to provide help or advice to avoid any biases in the results. Lastly, after the game ends, the participant is asked to fill in the rest of the questionnaire.

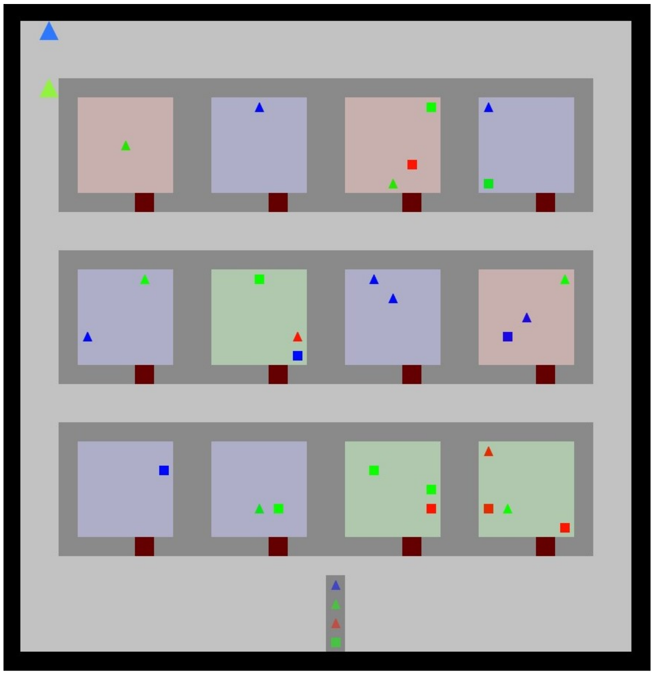
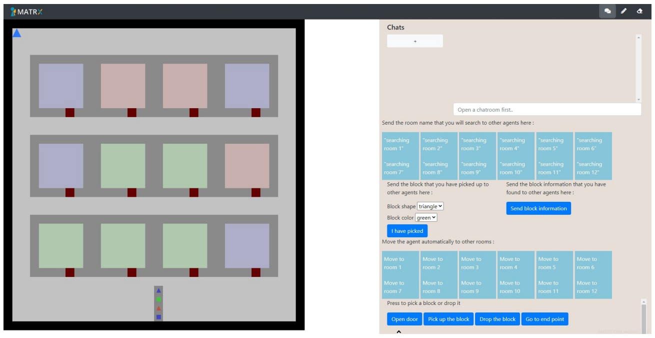


Figure 2: The map which the participant played with the AI agent.

In the second phase of the experiment, the participants' views are shown in Figure 3. Participants run it for the first time to get used to the buttons that control the agent they use. This is necessary for participants to get used to the agents and the controls, and since the experiment focuses on communicating of the human with artificial intelligence agents, rather than bringing the complexity of the game into the formula. Since agent control could be a problem, participants are allowed to play a single-player game to get used to the agent until they knew all the commands and game rules.



Open a chatroom first..

Send the room name that you will search to other agents here :

Send the block that you have picked up to other agents here :

Block shape: triangle

Block color: green

Send the block information that you have found to other agents here :

Move the agent automatically to other rooms :

Press to pick a block or drop it

Figure 3: Participant view

The results collected at the end of the game are partly provided by the MATRX structure, which includes the total number of last ticks and moves, as well as the number of messages sent by participants and agents.

3.5 Participants

The experiment was conducted remotely for 24 participants, with 6 participants in each sub-experiment. The experiment was run in a between-subject design. Each participant has conducted only one type of experiment. The age of the participants ranged from 18 to 30 years old. The ratio between male - female participants were 71% and 29% respectively. Only 17% of participants were familiar with MATRX software. Almost all participants are university students. All data shown in the next section are anonymous and unidentifiable.

4 Results and analysis

4.1 Team effectiveness results

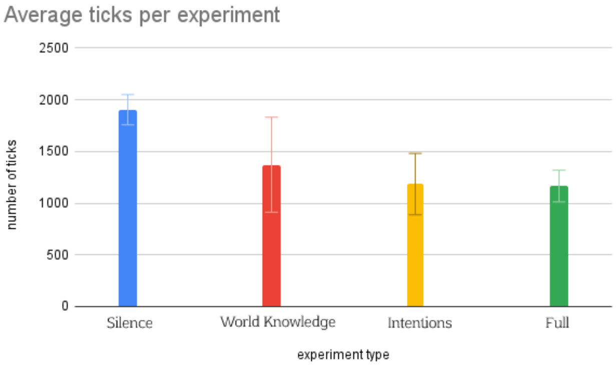


Figure 4: Average ticks

Shared Mental model	Average ticks	Standard deviation
Silence	1904.17	145.95
World knowledge	1372.17	459.43
Intentions	1184.83	296.71
Full	1165.33	153.13

Table 1: The last tick values for the graph in figure 4

According to the results in figure 4, It can be shown that the time taken to finish the game is lower when information is shared. Therefore the team performance increases when information is shared. Silence, in this case, had an average of 1904 ticks but the other three experiments which included information sharing have a lower number of ticks.

Most importantly, full information exchange and intention exchange, on average, outperform world knowledge information sharing. This interesting observation shows that some components of the shared mental model have a greater impact than others. This trend was also noticed in the experiment, because the unique division of knowledge about the world only led to confusion in team coordination, because the participants were not sure whether the agent actually took meaningful action in this situation. The world knowledge information sharing setting forced the participants to spend more time checking the endpoints and figuring out that some blocks haven been already delivered. So they would discard the blocks they have and search for the next goal block.

Moreover, also it has been noticed that the number of messages on average for intentions only and full information exchange were higher than world knowledge information exchange. As shown in figure 5. It can be seen that also in this case specifically for intentions users tend to send more messages than world knowledge however they shared less than the full information exchange by nearly 40%.

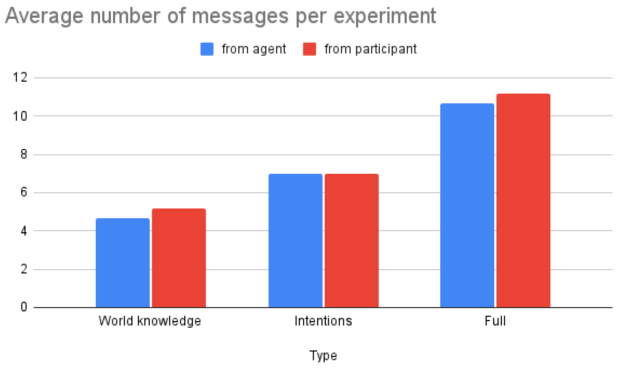


Figure 5: Average number of messages per experiment

Type	World knowledge	Intentions	Full
From agent	4.67	7	10.67
From user	5.17	7	11.17

Table 2: The average number of messages values for the graph in figure 5

Figure 6 shows a box plot of 4 values in table 3. From the shown figure, we can see that the minimum number of ticks for world knowledge exchange is better than full exchange and intentions exchange values. This value shows that some participants were lucky enough as they searched for a room with the desired block first. This investigation shows that some unintended luck has helped in finishing the task earlier than the other participants within the same experiment type. Also, there can be seen that the maximum value for intentions seems to be higher than the average for intentions sharing which was 1184.83, which occurred due to errors from the participants by sending wrong messages or not dropping off a block which leads to more time to finish the game.

Type	Min	Quartile 1	Quartile 3	Max
Silence	1644	1897	1993	2061
World knowledge	862	1013.5	1772.25	1934
Intentions	884	964	1356	1641
Full	931	1123.75	1281.75	1343

Table 3: The values in this table is an analysis of the data for the last tick value.

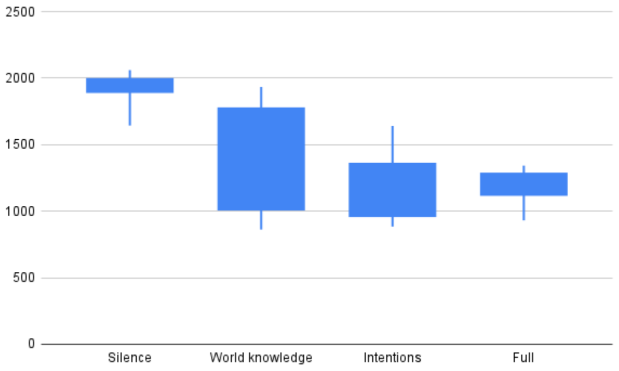


Figure 6: Box plot of the last tick values for all experiments from values in table 3

Two statistical analyses were run on the last tick data. First, we started with the null hypothesis that all four experiments' averages are exactly equal. ANOVA (Analysis of Variance) was conducted on the last tick data and the results from this analysis are shown in table 4. The ANOVA results show that the $p - value = 0.000885$ which is lower than 0.05. So we can reject the null hypothesis which shows that our four experiments' averaged are very unlikely to be equal. In Addition, the effect size - partial eta squared (n_2) has been calculated with the following formulae [9].

$$\text{ETA squared} = \frac{\text{SS between}}{\text{SS between + SS within}} \quad (1)$$

For now, suffice to say that $n_2 = 0.554$ for our data. This effect size explains why our F-test is statistically significant despite our very tiny sample sizes of n = 6 per group.

Source of variation	Between Groups	Within Groups
Sum square (SS)	2136889.792	1719343.833
Degrees of freedom (df)	3	20
Mean Square (MS)	712296.5972	85967.19167
F	8.285679495	
P-value	0.0008854194146	

Table 4: ANOVA (Analysis of Variance) results on the **last tick** data set from all the experiments.

Furthermore, the Tukey HSD test has been conducted on the original data to show the relation between the means of each group and whether there is a significant difference between them. Results for the Tukey HSD are shown below in table 5.

Comparison	Mean diff	P-adj	Significant diff
Full vs Intentions	19.5	0.9	NO
Full vs Wk	206.8333	0.61	NO
Full vs Silence	738.8333	0.0016	YES
Intentions vs Wk	187.3333	0.6728	NO
Intentions vs Silence	719.3333	0.0021	YES
Wk vs Silence	532	0.0243	YES

Table 5: Multiple Comparison of Means of **last tick** data set - Tukey HSD, alpha=0.50. (Wk = world knowledge)

The results show that there is a significant difference between the mean of silence and three other experiment types.

4.2 Trust results

As mentioned previously the trust was evaluated as the average of the 8 questions. These values were averaged for each experiment type creating the histogram in figure 7.

The lowest trust average was around 3 as can be shown in figure 7 for world knowledge participants to the AI agent. That average was lowest due to the fact participants didn't the agent's actions as shown in appendix B. A factor to this result is that the agent delivers the block without telling the human, which have might lowered the trust for this type of information sharing as shown in appendix B. However, the agent's output didn't affect the participants' trust in case of the silence experiment. Most of the time the agent finished the game, and the participants weren't aware of the agent's actions. This interesting observation carries us to the conclusion that participants trusted agents that don't communicate more than agents that communicate with insufficient information to understand the change in the world state. The information would have been sufficient if the agent would share more information about its intentions.

Moreover, Intentions information sharing gained the highest trust by the participants which is higher than full information sharing as seen in figure 7. The factors that lead to the result that the participants trusted agent that shares only intentions more than Full information sharing agent on 6 different questions are shown in appendix B.

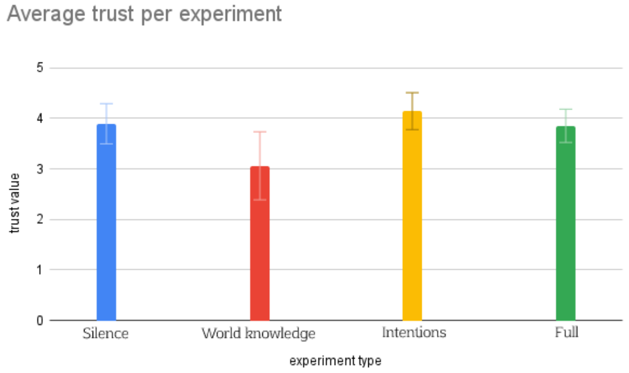


Figure 7: Average amount of trust per experiment

Shared Mental model	Average ticks	Standard deviation
Silence	3.90	0.40
World knowledge	3.06	0.67
Intentions	4.15	0.37
Full	3.85	0.33

Table 6: The average trust values for the graph in figure 7

The same statistical analysis has been ran on the trust data set. The results are shown in table 7 with an eta value of $n_2 = 0.481$. The results from the ANOVA test shows that the means of the four experiments are not likely to be the same and the eta value shows that the F-Test is significant.The values in table 8 clearly show that there is a significant difference between the world knowledge exchange and the three other experiments.

Source of variation	Between Groups	Within Groups
Sum square (SS)	3.966146	4.281250
Degrees of freedom (df)	3	20
Mean Square (MS)	1.322049	0.214062
F	6.175994	
P-value	0.003817	

Table 7: ANOVA (Analysis of Variance) results on the **trust** data set from all the experiments.

Comparison	Mean diff	P-adj	Significant diff
Full vs Intentions	0.2917	0.6809	NO
Full vs Wk	0.7917	0.0356	YES
Full vs Silence	0.0417	0.9	NO
Intentions vs Wk	1.0833	0.0032	YES
Intentions vs Silence	0.25	0.766	NO
Wk vs Silence	0.8333	0.0255	YES

Table 8: Multiple Comparison of Means of **trust** data set - Tukey HSD, alpha=0.50. (Wk = world knowledge)

5 Responsible Research

This experiment involved working with human participants which raised important ethical concerns. Participant data is sensitive data, which in this case must be well maintained and managed. We obtained consent from participants before starting the experiment to confirm that the data collected would only be used during this research study. In addition, the participants were allowed to leave at any time during the study, which was not the case, and they were informed with this information both orally and in the consent form. Data was stored on Google Drive which complies with the EU regulations for data protection regulations. The data was deleted after the end of the research project. This information was made available to the participant both in the information sheet and in the user's consent form in Appendix A. The information sheet outlined the purpose of the project and the type of data collected.

This article shows the experimental guidance on how the experiment was performed. We tried to make the guide easy to reproduce the experiment for future research that involve more participants. The agent's strategy that has been used in the game is described in the experimental setup section. The questionnaire which has been used in this experiment has been attached to this paper in appendix A that can be sent to get the results for the trust of the participant in the agent. The code is available to anyone upon request.

6 Discussion

The data suggest a positive correlation between team performance and information exchange. The analyses we did in the results section support this correlation. First, It can be seen from the statistical analysis results that there is a significant difference between the average means between the silence experiment and the other type of experiments. In addition, the high value of the effect size shows that the data is statistically representative. Secondly, the average for the silence experiment is the highest, and for the experiments with information sharing, the average means are lower. From the previous two points, we can conclude that when there is information shared in the mental model, the performance of the team increase, which in line with our hypothesis that mental models with more information sharing the higher the performance.

Furthermore, our experiment showed that the intention exchange achieved a lower number of messages in comparison to the full information exchange's average message count. However, the two experiments' means are relatively close. Therefore, we assume that intention sharing was the best shared mental model with low communication overhead and near-best performance.

Jonker et. al. showed that the shared mental model can be used to predict the team performance in the AI agent - AI agent team [4], and also we have reached a similar result in the Human - AI agent teaming. In addition, we have looked into the trust between the Human and the AI agent, and we concluded that the participants trusted the agent sharing intentions only more than the others. Participants trusted the AI - agent when the agent shared its intentions and world knowledge, or intentions only or no communication more than sharing world knowledge only. This observation shows that participants trusted the agents who communicate with useful information or nothing at all more than agents who would share world knowledge only. World knowledge information sharing can't allow the participant to predict the agent's actions. In the silence experiment, they didn't share any information. However, the participant trusted that the agent efficiently works.

We noticed that the participant's strategy affected the time needed to finish the game and we believe it is one of the confounding factors that could affect the end time of the game. As if it is in line with the agent's strategy, the task would end early most of the time. However, as the agent can't change its strategy depending on the various strategies that exist, better results would be seen if the participant and the AI agent could negotiate a strategy to finish the game before they start.

Our data is only a representation of 24 participants, and we expect to reach a more concrete conclusion with more participants. The low number of participants was a limitation in reaching a solid conclusion. The low number of participants was due to the limited duration of this study. The network connection was also another limitation and confounding factor to our experiment.

For future research, we would suggest having more participants. If the experiment would be conducted in an online setting, we advise that the whole game should be on the client-side rather than on the server-side, which would avoid the internet connection delay. Also, we highly recommend giving a demo of the game to the participants to introduce the game to them.

7 Conclusion

In this paper, we have shown four different mental model representations for a person to share his mental model with the AI agent. The level of information exchange in each mental model differs in both amount and type. We have run our experiment of human-AI agent teaming on BW4T. After analyzing the experiment results, we found a trend towards the importance of the mental model in improving the team's performance between the human and the AI agent. The results were in line with the hypothesis that more information shared would lead to better team performance. Also, we have investigated the trust of the participant in the agent in each experiment type and found out that participants in the intention exchange experiment had the highest trust value.

References

[1] M. Johnson and A. Vera, "No ai is an island: the case for teaming intelligence," *AI Magazine*, vol. 40, no. 1, pp. 16–28, 2019.

[2] B.-C. Lim and K. J. Klein, "Team mental models and team performance: A field study of the effects of team mental model similarity and accuracy," *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, vol. 27, no. 4, pp. 403–418, 2006.

[3] M. Harbers, M. v. Riemsdijk, and C. Jonker, "Measuring sharedness of mental models and its relation to team performance," in *Proceedings 14th International Workshop on Coordination, Organisations, Institutions and Norms*, pp. 106–120, 2012.

[4] C. M. Jonker, M. B. Van Riemsdijk, I. C. Van De Kieft, and M. Gini, "Compositionality of team mental models in relation to sharedness and team performance," in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, pp. 242–251, Springer, 2012.

[5] I. van de Kieft, C. M. Jonker, and M. B. van Riemsdijk, "Improving user and decision support system teamwork. an approach based on shared mental models," in *IJCAI 2011 Workshop ExaCt*, pp. 61–70, 2011.

[6] J. B. Lyons, K. T. Wynne, S. Mahoney, and M. A. Roebke, "Chapter 6 - trust and human-machine teaming: A qualitative study," in *Artificial Intelligence for the Internet of Everything* (W. Lawless, R. Mittu, D. Sofge, I. S. Moskowitz, and S. Russell, eds.), pp. 101–116, Academic Press, 2019.

[7] M. Lewis, K. Sycara, and P. Walker, "The role of trust in human-robot interaction," in *Foundations of trusted autonomy*, pp. 135–159, Springer, Cham, 2018.

[8] R. R. Hoffman, S. T. Mueller, G. Klein, and J. Litman, "Metrics for explainable ai: Challenges and prospects," *arXiv preprint arXiv:1812.04608*, 2018.

[9] R. G. van den Berg, "ANOVA – Super Simple Introduction." https://www.spss-tutorials.com/anova-what-is-it/#effect-size.

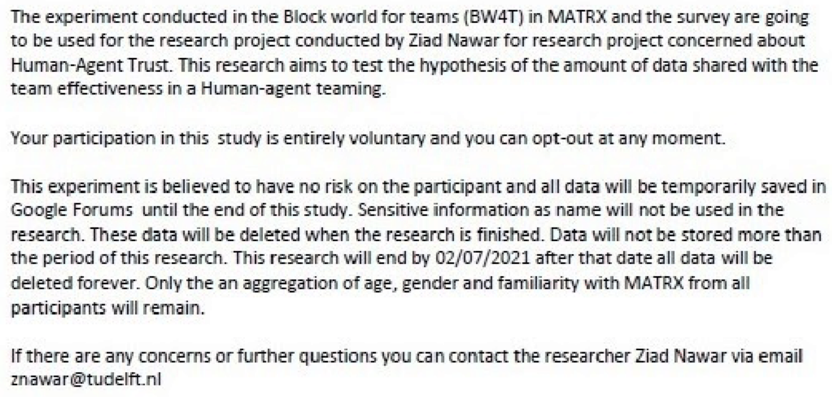
A Questionnaire

Research survey

This research is conducted by Ziad Nawar. Any data that could be identifiable for the user such as name will be deleted after the research is done.

* Required

Information sheet



The experiment conducted in the Block world for teams (BW4T) in MATRX and the survey are going to be used for the research project conducted by Ziad Nawar for research project concerned about Human-Agent Trust. This research aims to test the hypothesis of the amount of data shared with the team effectiveness in a Human-agent teaming.

Your participation in this study is entirely voluntary and you can opt-out at any moment.

This experiment is believed to have no risk on the participant and all data will be temporarily saved in Google Forums until the end of this study. Sensitive information as name will not be used in the research. These data will be deleted when the research is finished. Data will not be stored more than the period of this research. This research will end by 02/07/2021 after that date all data will be deleted forever. Only the an aggregation of age, gender and familiarity with MATRX from all participants will remain.

If there are any concerns or further questions you can contact the researcher Ziad Nawar via email znawar@tudelft.nl

1. Name *

____

2. What is your age? *

Mark only one oval.

○ 18 - 30
○ 31 - 40
○ 41 - 50
○ 50 +

3. What gender do you identify as? *

Mark only one oval.

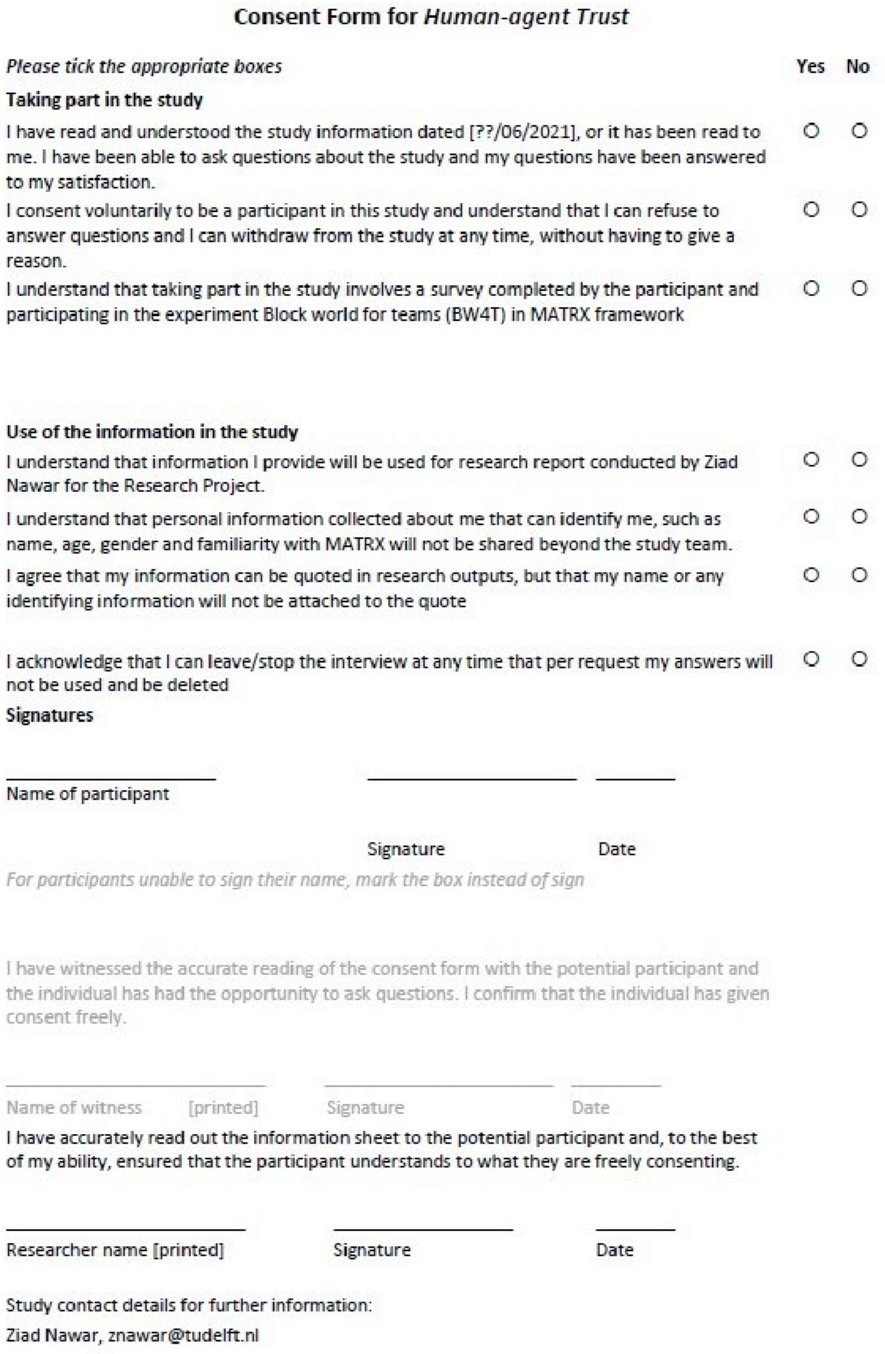
○ Female
○ Male
○ Prefer not to say

4. Have you used MATRX before ? (Implementation or played a game) *

Mark only one oval.

○ Yes
○ No

5. By pressing yes you are accepting the consent form *



Consent Form for *Human-agent Trust*

Please tick the appropriate boxes	**Yes**	**No**
Taking part in the study		
I have read and understood the study information dated [??/06/2021], or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	○	○
I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.	○	○
I understand that taking part in the study involves a survey completed by the participant and participating in the experiment Block world for teams (BW4T) in MATRX framework	○	○
Use of the information in the study		
I understand that information I provide will be used for research report conducted by Ziad Nawar for the Research Project.	○	○
I understand that personal information collected about me that can identify me, such as name, age, gender and familiarity with MATRX will not be shared beyond the study team.	○	○
I agree that my information can be quoted in research outputs, but that my name or any identifying information will not be attached to the quote	○	○
I acknowledge that I can leave/stop the interview at any time that per request my answers will not be used and be deleted	○	○

Signatures

____________________ ____________________ ________

Name of participant

Signature Date

For participants unable to sign their name, mark the box instead of sign

I have witnessed the accurate reading of the consent form with the potential participant and the individual has had the opportunity to ask questions. I confirm that the individual has given consent freely.

____________________ ____________________ ________

Name of witness [printed] Signature Date

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

____________________ ____________________ ________

Researcher name [printed] Signature Date

Study contact details for further information:

Ziad Nawar, znawar@tudelft.nl

Mark only one oval.

○ Yes, I understand

○ No I don't want the data I produce to be used in the research

Experiment 1

6. The outputs (messages and actions) of the agent are very predictable *

Mark only one oval.

○ Strongly disagree
○ Disagree
○ Neutral
○ Agree
○ Strongly Agree

7. I like working with the agent for decision making. *

Mark only one oval.

○ Strongly disagree
○ Disagree
○ Neutral
○ Agree
○ Strongly Agree

8. The agent is efficient in that it works very quickly. *

Mark only one oval.

○ Strongly disagree
○ Disagree
○ Neutral
○ Agree
○ Strongly Agree

9. I am confident in the agent. I feel that it works well. *

Mark only one oval.

○ Strongly disagree
○ Disagree
○ Neutral
○ Agree
○ Strongly Agree

10. The agent is very reliable. *

Mark only one oval.

○ Strongly disagree
○ Disagree
○ Neutral
○ Agree
○ Strongly Agree

11. I am wary of the agent. *

Mark only one oval.

○ Strongly disagree
○ Disagree
○ Neutral
○ Agree
○ Strongly Agree

12. I feel safe that when I rely on the agent I will get the right responses. *

Mark only one oval.

○ Strongly disagree

○ Disagree

○ Neutral

○ Agree

○ Strongly Agree

13. The agent can perform the task better than a novice human user. *

Mark only one oval.

○ Strongly disagree

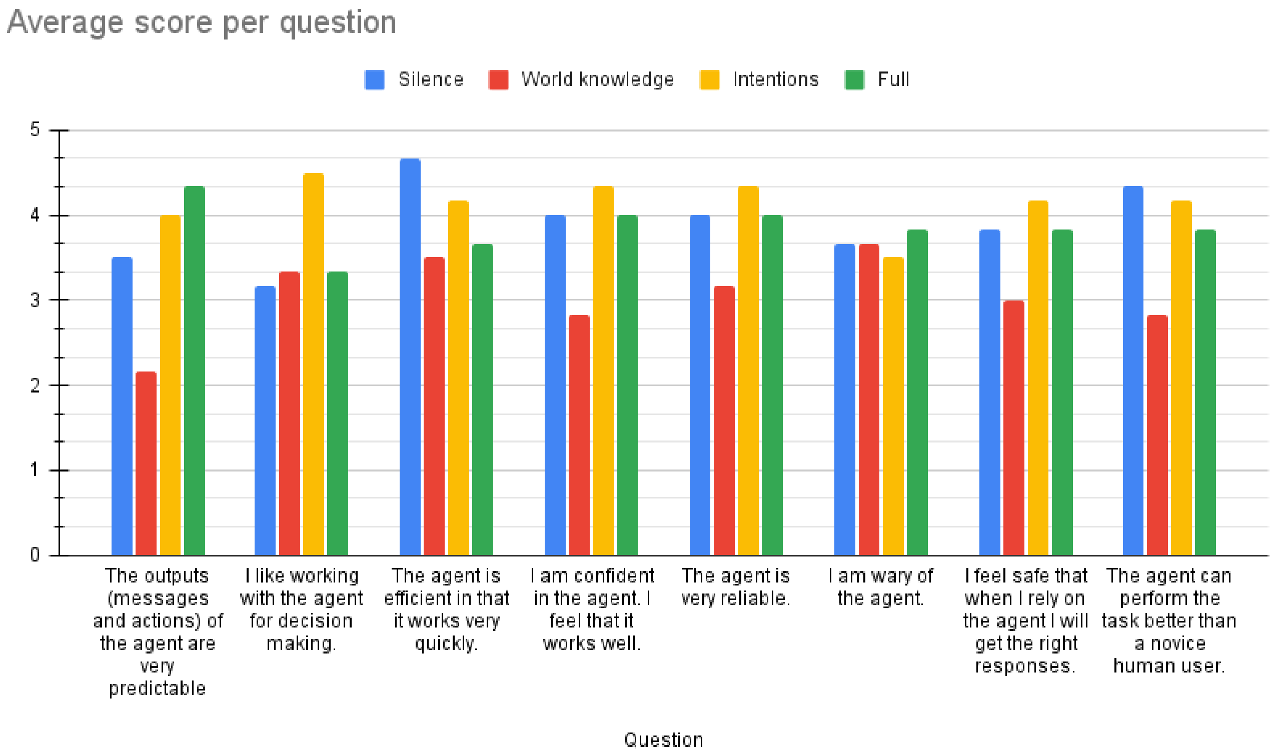
○ Disagree

○ Neutral

○ Agree

○ Strongly agree

B Questions result



Average score per question for the trust questionnaire section after applying the conversion for the question about the wariness from the agent