Measuring Sharedness of Mental Models and its Relation to Team Performance

Maaike Harbers, Birna van Riemsdijk, Catholijn Jonker

TU Delft, Mekelweg 4, 2628 CD, Delft, The Netherlands
`{m.harbers,m.b.vanriemsdijk,c.m.jonker}@tudelft.nl`

Abstract. Members of a team have a mental model about the team's goal, the tasks, other team members, and other aspects related to teamwork. The extent to which these mental models are shared is often used to understand and predict performance of human teams. This paper describes a compositional measure for the sharedness of mental models, and investigates whether this measure can be used to predict the performance of agent teams. In a simulation experiment with different agent teams it was found that the proposed measure for sharedness of mental models can be used to predict team performance, and that different components contribute to team performance to different extents.

Keywords: Teamwork, shared mental models, measurement, team performance, BW4T testbed, BDI agents, multi-agent system.

1 Introduction

Human teamwork has extensively been studied in psychological literature. The concept of shared mental models (SMMs) is often used to explain and understand teamwork [2,3,9,10]. Mental models refer to humans' internal representations of the world around them, that help them to understand, explain and predict the systems in their environment [13]. In the context of teamwork, mental models can help individuals to understand and predict the behavior of other team members, which allows them to anticipate their own actions to the expected behavior of others. To coordinate actions well and thus achieve good team performance, it is important that the members in a team have similar mental models. Several experiments in human teams show that higher sharedness of team mental models lead to higher team performance [9].

Increasingly often, humans and intelligent software agents work together in mixed human-agent teams [14]. In order to use SMMs for understanding and predicting the performance of teams that contain agents (agent teams or human-agent teams), the SMM measures developed for human teamwork need to be adapted for teams with agents. Namely, humans are often asked to make judgements about concepts to determine their mental model, which agents are not always capable of in a similar fashion. Jonker et al. [8] proposed a formal measure for the sharedness mental models. The approach is based on literature about human teamwork, but the measure can be used for human teams, agent teams and human-agent teams.

In this paper, we will investigate whether the concept of SMMs developed for human teamwork, can be used to predict the performance of agent teams. For that, we will propose a more general version of Jonker et al.'s SMM measure. One of the generalizations is that the new approach is compositional [7]. This means that a mental model can consist of different components with their own weights attached to it, reflecting the component's predictive power for team performance. Our first hypothesis is that the compositional version of Jonker et al.'s measure for sharedness of mental models can be used to predict team performance. Our second hypothesis is that the sharedness levels of different mental model components contribute to team performance to different extents, dependent on the context. To test these hypotheses, we will perform an experiment in BlocksWorld for Teams (BW4T), a testbed for team coordination [6]. We will run a simulation with different agent teams in different contexts, compositionally measure the sharedness of their mental models, and relate these sharedness values to team performance.

The paper is organized as follows. In section 2 we will provide a background on measuring shared mental models in teams. In section 3 we will describe the setup of the experiment. In section 4 we will provide the results of the experiment. In section 5 we will discuss the results. Finally, in section 6 we will provide a conclusion and give suggestions for future research.

2 Measuring sharedness of mental models

In section 2.1 and 2.2 we discuss the what and how of measuring the sharedness of mental models, respectively. In both sections, we discuss the differences and similarities between human teams and teams with agents in this regard. In section 2.3 we introduce the generalized version of Jonker et al.'s compositional measure for sharedness of mental models, which we will use in this paper.

2.1 What to measure?

To achieve a good team performance, it is not necessary, and often even undesired, that the mental models of the team members are completely shared. Therefore, an important question when using the sharedness of mental models to predict team performance is which parts of the mental models should be shared. In other words, what should be measured? We will use the term team mental models to refer to those parts of mental models that are relevant for team performance.

Mohammed et al. [9] define a team mental model as "an organized understanding of relevant knowledge that is shared by team members". Team mental models can be classified into two broad dimensions: task-related knowledge and team-related knowledge (e.g. [3]). Task-related knowledge concerns knowledge, for instance, about how to achieve the task, and the current status of task achievement. Team-related knowledge concerns knowledge about e.g. the knowledge and capabilities of other team members, what they are currently intending or doing.

Cannon-Bowers et al. [2] proposed a more detailed categorization of team mental models. They distinguish the following four categories: equipment-related knowledge, task-related knowledge, team members-related knowledge, and team interaction-related knowledge. Relating this to the two broader categories, the first two are task-related and last two are team-related.

The components of team mental models that are distinguished in human teamwork seem to be relevant for human-agent teams as well. However, some knowledge that can be assumed to be shared in human teams, cannot be assumed to be shared between humans and agents. For example, humans have a lot of world knowledge that agents do not have, and humans know more or less what other humans are able to understand, whereas humans do not always know what other agents are able to understand. A measure for sharedness of mental models in human-agent teams should incorporate this additional knowledge.

Jonker et al. [7] proposed a more agent-oriented categorization of team mental model components. In addition to the components distinguished in literature on human teamwork, they add information and knowledge about the domain, and the underlying ontologies. This results in four components: ontologies, world state model, agent model, and organizational specification. The usual components referring to task, team interaction, and team members are part of the organizational specification. In section 2.3 we will discuss how sharedness regarding these components is measured.

2.2 How to measure?

To use SMMs for understanding and predicting team performance, the sharedness of team mental models needs to be measured. In psychological research, several measures for assessing sharedness have been proposed, e.g. concept mapping, paired ratings, card sorting and causal mapping [9,10]. Usually in these measures, first a set of concepts is determined, and then the team members individually have to make judgements about these concepts, e.g. to what extent the concepts relate or how they relate to each other. The judgements of different team members are compared to each other, and a sharedness value is calculated. The set of concepts is either determined by experts, in which case all team members work with the same set and their results are more easy to compare, or the set is determined by the team members individually, in which case the results are less influenced by choices of experts.

A general agreement on how to measure sharedness of team mental models is lacking. Moreover, in a comparison study it was found that different SMM measures measure different things [12]. In conclusion, the field is divided on how to measure sharedness of team mental models.

Assessing mental models of agents differs from assessing mental models of humans. On the one hand, agents are not capable of understanding and answering some of the questions that are asked to humans. For example, an agent may not have a mechanism to determine to what extent two concepts relate to each other, and it is usually not possible to ask an agent to what extent it had the feeling that the team members had a shared understanding of the task. Moreover, it is also hard to derive the answers to such questions from the agent's programming code.

On the other hand, whereas mental models of humans can only be assessed indirectly by asking questions and observing their behavior, it is possible to look directly inside the 'head' of an agent, i.e. the programming code. Depending on the way the agent's behavior is represented, this gives more or less insightful information. If an agent's behavior is represented by a neural network, for example, knowing the weights of the connections between nodes will not give much insight in what the agent 'knows'. When the agent's behavior is represented in a BDI-based language (Belief Desire Intention) [11], there is access to the agent's goals, beliefs and plans at any time.

Another advantage of assessing an agent's instead of a human's mental model is that it does not interrupt task execution. Mental models of humans are usually assessed by questionnaires. First, it takes time to fill in these questionnaires. Second, a human's mental model may have changed after filling in a questionnaire, e.g. certain task-related knowledge is forgotten. This is the reason that questionnaires are usually administered to humans only once and after task execution. Making a backup of an agent's program state, in contrast, neither takes time, nor changes the agent's mental model. It is thus possible to obtain insight in the development of the mental model and the sharedness of mental models of a team of agents over time.

2.3 A compositional measure

Jonker et al. [8] proposed an approach for measuring the sharedness of team mental models. This SMM measure is appropriate for our goals since it is applicable to teams with humans and/or agents instead of only human teams. However, the measure has certain limitations (see below). Therefore, in this section we present a more general and extended version of the measure proposed in Jonker et al. [8]. The new measure differs from the original one with respect to the following aspects.

- The original measure could be applied to only two models (m1 and m2), whereas the new measure can be applied to a set of mental models M.
- The original measure was restricted by the assumption that the mental models correspond to a finite ground model, whereas the new measure is not. This was accomplished by considering all beliefs in the models instead of all possible beliefs.
- In the original measure questions with variables were not allowed, whereas in the new measure they are. This was also accomplished by considering all beliefs in the models instead of all possible beliefs.
- In the new measure it is possible to assign different levels of relevance to different components of a mental model, which was not possible in the original measure.

The compositional sharedness measure is defined in terms of a definition of model subject overlap and a definition of model agreement. Subject overlap provides a measure for the extent to which models can provide answers to a set of relevant questions Q. Model m can answer question q iff $q \in L(Ont(m))$, where Ont(m) is m's ontology containing predicates, functions and constants, and L(Ont(m)) is the language defined over Ont(m).

Definition 1 Model Subject Overlap. Let the set of questions Q for which the set of models M provide answers (not necessarily similar answers) be $OverAns(M,Q) = \{q \in Q \mid \forall m \in M : q \in L(Ont(m)\}$. The level of subject overlap between the set of models M with respect to the set of questions Q is defined as $SO = |OverAns(M,Q)| / |Q|$.

Model agreement defines the extent to which models provide equivalent answers to questions.

Definition 2 Model agreement. Let *ans(m,q)* be the answer of model m with respect to question q. Then the agreement between the set of models M with respect to the question q is defined as

$$Ag(M,q) = \frac{|\bigcap_{m \in M} ans(m,q)|}{|\bigcup_{m \in M} ans(m,q)|} \qquad \text{if } |\bigcup_{m \in M} ans(m,q)| \neq 0;$$

$$Ag(M,q) = 0 \qquad \text{otherwise.}$$

Then, the level of agreement between the set of models M with respect to the set of questions Q is defined as $Ag(M,Q) = \frac{1}{|Q|} \sum_{q \in Q} Ag(M,q)$.

With these two definitions, a Shared Mental Model can be defined as follows.

Definition 3 Shared Mental Model. A model m is shared to the extent θ by the set of agents A with respect to a set of questions Q, denoted by $Sh(M,A,Q,\theta)$iff

1) $SO(M_A \cup \{m\}, Q) = 1$;
2) $Ag(M_A \cup \{m\}, Q) \geq \theta$.

We now define the sharedness of the compositional model in terms of a function that takes the sharedness per component to the extent it is relevant, and follows that same composition relation as the team mental model itself. A compositional linear additive function suffices for the case study, using weights to model the relevance of each component.

Definition 4 Compositional Sharedness. The function CS determines the extent of sharedness of a set of agents, their compositional models, and a set of questions. For any set of agents A, set of models M, and set of questions Q, CS is defined by:

$$CS(M,A,Q) = \max\{\theta \mid Sh(M,A,Q,\theta)\}, \quad \text{if } M \text{ is not composed.}$$

$$CS(M,A,Q) = c(\{CS(m,A,Q) \mid m \in M\}), \quad \text{if } M \text{ is composed.}$$

Where c is a composition function over the components m of M that combines the results of $CS(m,A,Q)$ for all $m \in M$ in such a way that the combined results is a value in the range [0, 1].

For example, c could be a linear additive function: $\sum_{m \in M} w_m \, CS(m,A,Q)$ and by demanding that each weight $w_m \in [0,1]$ and setting $\sum_{m \in M} w_m = 1$, the combined value is again in the range [0,1]. Note that the compositional structure of the model can be reflected in a compositional structure of the set of questions Q, e.g., the subset of questions on the current situation (domain model), the subset of questions on the organizational specification.

3 Experimental setup

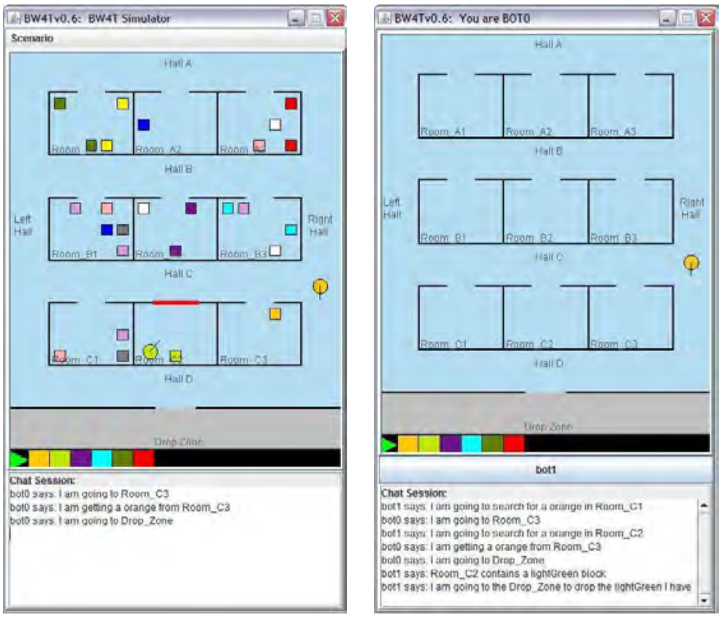
To test our hypotheses that the compositional SMM measure can be used to predict team performance, and that different components contribute to different extents to team performance, we conducted a simulation experiment. The experiment was performed in the BlocksWorld for Teams (BW4T) environment, a testbed for team coordination [6]. We ran several simulation runs in BW4T with different teams of agents in different scenarios. In this section we will describe the BW4T environment, the behavior of the agents, the scenarios, and the measures that we used in the experiment.

3.1 The BW4T environment

The goal of the BW4T task is to search and deliver a particular sequence of colored blocks as quickly as possible. A variable number of players can participate, and players can be humans, agents, or a mix of humans and agents. When the task is performed by more than one player, the players need to coordinate their actions in order to achieve a good team performance. The task is simple to learn, but due to the many interdependencies among the different players, the coordination of actions becomes complex.

Figure 1 displays a screenshot of a BW4T game session, showing the environment in which the players have to search for blocks. The left picture displays all blocks and players in the game, and the right picture shows what a single player can see. A player cannot see other players, and it can only see the blocks in a room when he is inside that room. The status bar below the Dropzone shows which blocks need to be delivered, and in what order they should be delived. To deliver a block successfully, a player has to go to a block of the right color, pick it up and drop it in the Dropzone. A player can only carry one block at a time. When a player drops a block of the wrong color in the Dropzone or any block in a hall, the block disappears from the game.

A team's performance on the BW4T task is measured by the speed of completing the task. BW4T is designed such that the task involves a large amount of interdependence among the players, and requires coordination to achieve a good performance. For instance, it is inefficient when one player is searching in a room that has just been checked by another. Furthermore, if a player is going to deliver a particular block, the others should not do that as well. To coordinate, players can send messages to each other, which appear in the chatbox below the Dropzone. Players can inform others about what they do, where they are and what they see. Furthermore, players can see the same status bar. So when a player delivers a block of the right color, the other players will know. Finally, only one player can be inside a room or the Dropzone at the same time. When a player tries to enter a room that is occupied, a red bar appears indicating that someone is inside.



```
Chat Session:
bot0 says: I am going to Room_C3
bot0 says: I am getting a orange from Room_C3
bot0 says: I am going to Drop_Zone
```

```
Chat Session:
bot1 says: I am going to search for a orange in Room_C1
bot0 says: I am going to Room_C3
bot1 says: I am going to search for a orange in Room_C2
bot0 says: I am getting a orange from Room_C3
bot0 says: I am going to Drop_Zone
bot1 says: Room_C2 contains a lightGreen block
bot1 says: I am going to the Drop_Zone to drop the lightGreen I have
```

Fig. 1. Overview of a BW4T session (left) and the view of a single player (right).

3.2 Agent types

We first developed a basic agent that was able to perform the BW4T task, and then made four variations of that agent. The basic agent's strategy is that if it knows about a block of the color that needs to be delivered, it will go to that room, pick up the block and deliver it. If it does not know about such a block, it will go to a room it has not checked yet to search for that block. The agent is capable of processing messages of other agents. So if another agent tells that there is a blue block in room A1, it will use that information if a blue block needs to be delivered. When another agent tells that it is going to deliver a block of the next color in the goal sequence, the agent will deliver or search for a block with the second-next color.

To test if sharedness of mental models predicts team performance, we needed to manipulate the sharedness of the agents' mental models. We did that by creating agents that communicate different amounts and types of information. Namely, when team members update each other on what they see, what they do and what they plan to do, sharedness of team mental models is increased. Because we use only agents and no humans, we can assume that all sent messages are perceived. We made the following four variations on the basic agent, which only differed from each other with respect to their communication behavior.

- Agent A – no communication at all
- Agent B – communication of world knowledge
- Agent C – communication of intentions
- Agent D – communication of world knowledge and intentions

World knowledge includes information about positions and colors of blocks, and information about the agent's own position and state, e.g. whether it is holding a block. Intentions include information about where the agent is going and which blocks it is going to deliver.

The agents were implemented in GOAL [5], a BDI-based agent programming language. A BDI agent is specified by its beliefs, desires and intentions, and its behavior is determined by a reasoning process on these mental concepts. BDI-based programming is based on folk psychology, and simulates the way humans think [1]. Because BDI agents have explicit representations of their intentions and beliefs, it is easy to communicate them to other team members [4].

3.3 Scenarios

To investigate whether the contribution of different shared mental model components to team performance is context dependent, we introduced two scenarios, one with a *high* level and one with a *medium* level of interdependence among the players. In the high scenario, the agents need to deliver six blocks which each have a different color. Consequently, there is only one possible order in which the blocks can be delivered, and for an agent to select a useful action highly depends on what the other agents do. In the medium scenario, the agents also need to deliver six blocks, but among those six blocks there are two pairs of subsequent blocks with the same color. In this scenario, the selection of an action depends less on what the other agents do. For instance, if there are two agents and two blue blocks that need to be delivered, both agents can go search and deliver a blue block, independent of what the other agent does. This is not the case when first a yellow and then a blue block need to be delivered.

3.4 Measures

To investigate the relation between sharedness of mental models and team performance, we needed to measure them. Team performance was measured by taking the speed of completing the BW4T task, i.e. the time it took to deliver all colors in the goal sequence.

Sharedness of mental models was measured by the compositional measure described in section 2.3. Table 1 provides an overview of the set of relevant questions Q that we distinguished for the BW4T task. The components and subcomponents in the table reflect the compositional structure of the model. The two main components are world knowledge and knowledge about intentions, where world knowledge contains seven subcomponents and knowledge about intention contains one subcomponent. The term 'imp' in the subcomponent of knowledge about intentions stands for imperative, so *imp(Player, Goal)* expresses that the player *Player* intents to achieve the goal *Goal*. Possible goals are *in(Player, Room)*, and *deliver(Player, BlockId)*. Each subcomponent contains one question. The questions were chosen such that for these agents the subject overlap is 1, i.e. the questions are part of the agents' ontology.

Component	**Subcomponent**	**Description**
world knowledge	allBlocksAt(Room,[BlocksId])	the list of block-ids in a particular room
	block(BlockId)	the id of an observed block
	color(BlockId,Color)	the color of a particular block
	gone(BlockId)	a particular block that was picked-up
	holding(BlockId)	a block that is being held by a player
	in(Player,Place)	the location of a particular player
	at(BlockId,Room)	the room where a particular block is
knowledge about intentions	imp(Player, Goal)	the intention of a player to reach a particular goal[1]

Table 1. Components and subcomponents of the team mental models in BW4T.

To calculate a team's sharedness θ on a subcomponent, the number of ground beliefs (e.g. block(17)) shared by all agents divided by the total number of ground beliefs of the agents. We will illustrate the measure by a simple example. The example involves a team of three agents with the following beliefs.

- Agent 1: block(2), block(3), block(4)
- Agent 2: block(1), block(3)
- Agent 3: block(2), block(3)

The number of shared beliefs in this example is 1, only the belief 'block(3)' is held by all agents. The total number of beliefs is 4, i.e. 'block(1)', 'block(2)', 'block(3)' and 'block(4)'. Thus, according to definition 2, the agreement on the subcomponent block(BlockId) is 1/4, and according to definition 3, the sharedness value θ is also 1/4. Note that the belief 'block(2)' is only shared by two of the three agents, and therefore it does not contribute to the sharedness value of the team.

3.5 Procedure

We performed eight BW4T simulation runs. In each run, a team of three agents of the same type played a particular scenario. There were four teams: team A, team B, team C and team D, where team A consisted of three agents of type A, team B consisted of three agents of type B, etc. The four teams each played two scenarios (medium and high), resulting in eight BW4T sessions.

---

[1] Note that players can have beliefs about the intentions of other players, but not about their own beliefs. Therefore, other players' beliefs about the goals of an agent should be compared to its goal base instead of its belief base.

We made logs of all the agents' mental states, i.e. their beliefs and goals at several time points in the BW4T sessions. More specifically, we stored the mental states of all agents immediately after a block was delivered. Since the agents had to deliver six blocks in both scenarios (high and medium), this resulted in three (agents) times six (block deliveries) mental state logs per session. Besides mental states, we also logged the time of each block delivery and the time it took to complete the task.

4 Results

In this section we will present the results of the eight BW4T simulation sessions. First, we will present the results regarding task performance. Second, we will present the results regarding sharedness of mental models. Third, we will analyze the predictive power of the shared mental model components on task performance.

4.1 Task performance

Figure 2 shows at what time each of the blocks were delivered in the four teams for both scenarios.

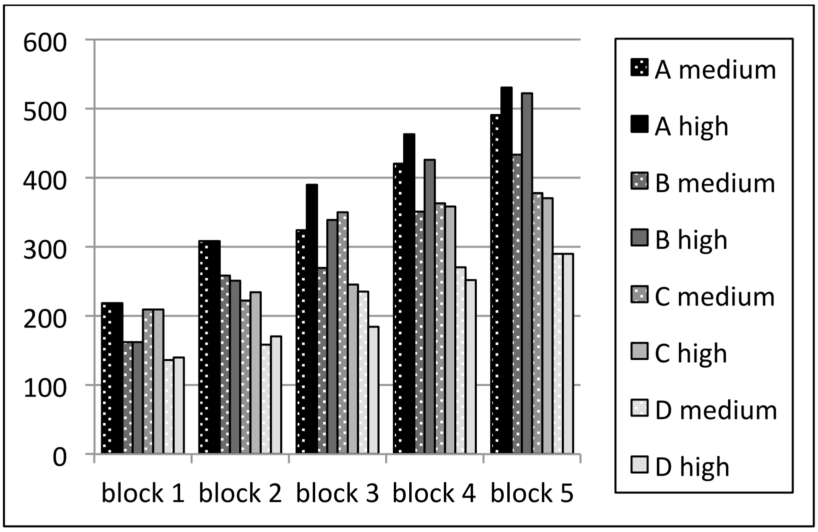


Fig. 2. Time of block delivery.

The results in figure show that the team with full communication (team D) took the least time to deliver all blocks, teams with partial communication (team B and C) took intermediate time to deliver all blocks, and teams with no communication (team A) took the most time to deliver all blocks. Of team B and C, which both involved partial communication, team C (communication of intentions) took less time to complete the task than team B (communication of beliefs).

For teams C and D there was almost no difference between performance in the medium and in the high scenario, but team A and especially team B performed much better in the medium scenario than in the high scenario. This can be explained by the fact that team C and D communicate about intentions and team A and B do not. In team C and D, agents know which blocks the other agents are going to deliver, so they will search for or deliver the right block in most cases, independent of whether there are pairs of blocks with equal colors in the goal sequence. In team A and B, however, it will not be unusual that two or three agents are trying to deliver a block of the same color, since they do not know that another agent is already intending to deliver that block. In that case, performance is not negatively affected when two blocks of the same color need to be delivered after each other (as in the medium scenario), but performance does go down when all blocks in the goal sequence have a different color (as in the high scenario).

The effect of not communicating about intentions in the high scenario is considerable. Namely, the overall performance (time of 5th block delivery) of team B in the high scenario is even worse than team A's performance in the medium scenario.

4.2 Sharedness of mental models

Figure 3 shows the compositional sharedness values at the time of each block delivery, for all eight conditions. The values were calculated according to the definitions in section 2.3. The composition function *c* was defined such that all (sub-)components were given an equal weight w_m. Thus, world knowledge and intentions both received a w_m of 1/2, and the subcomponents of the world knowledge component all received w_m 's of 1/7.

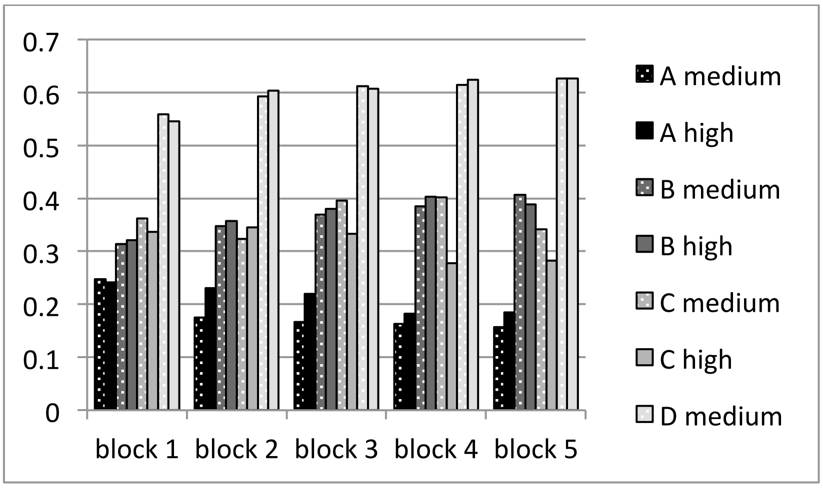


Fig. 3. Sharedness value (θ) at the time of each block delivery.

The results show that team A (no communication) has the lowest sharedness values, team B and C (partial communication) have intermediate sharedness values, and team D (full communication) has the highest sharedness values. The shared mental model in team A (θ is more than zero) did not develop through communication, but because all team members individually made the same observation, e.g. that there is a yellow block in room C1.

The total sharedness values are relatively stable over the course of a session. Though there are some fluctuations, most conditions do not show a clear trend in sharedness development. One may have expected an increase in sharedness for teams B, C and D, since their team members communicate during the session. First, however, most communication occurs at the beginning of the session. Second, the world state and the agents' mental model states change, which could lead to a decrease of sharedness. It seems that the communication that takes place after the first block has been delivered makes up for the loss of sharedness due to changes in the environment.

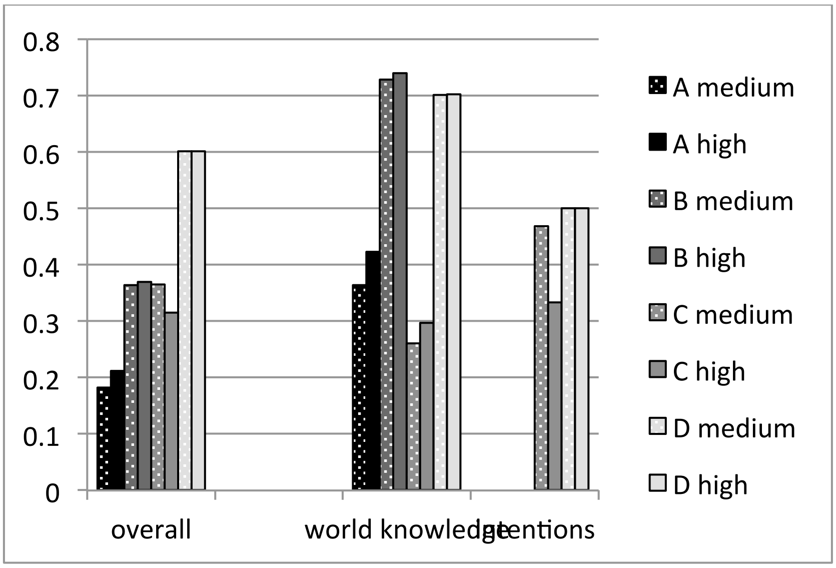


Fig. 4. Average sharedness values (block 1-5) in total and per components.

The total sharedness value is composed of the sharedness values of the two components *world knowledge* and *knowledge about intentions*, which are in turn composted of the sharedness values of their subcomponents. Figure 4 shows the sharedness values of these two components in comparison with total sharedness. The values are averages of all five block deliveries. The results show that team B (communication of world knowledge) has high sharedness values regarding world knowledge and a value of zero regarding knowledge about intentions. Team C (communication about intentions), in contrast, has high sharedness values for knowledge about intentions and relatively low values for world knowledge.

4.3 Sharedness of mental models and team performance

To calculate composed sharedness values, we assigned weights to the sharedness values of the (sub-)components. To investigate for which components the sharedness values best predict team performance, we correlated the average (block 1-5) sharedness values obtained in the eight conditions with the time to complete the task in the respective condition. The results are displayed in table 2. The results for the subcomponent *holding* are omitted since sharedness values always were zero in all conditions.

Overall	**Component**	**Subcomponent**	**Pearson correlation (r)**	**Significance (p)**
overall			-0.83	0.01
	world knowledge		-0.15	0.10
		allBlocksAt(Room, [BlocksId])	-0.29	0.48
		block(BlockId)	-0.07	0.87
		color(BlockId, Color)	0.07	0.87
		gone(BlockId)	-0.32	0.43
		in(Player,Place)	-0.16	0.71
		at(BlockId,Room)	0.08	0.86
	intentions		-0.92	0.04
		imp(Player, Goal)	-0.92	0.04

Table 2. Pearson correlations between task performance and sharedness values (overall, per component, and per subcomponent).

The results show that the correlation between the overall sharedness values and time to complete task is significant (p=0.01) and rather high (r=-0.83). The correlation is negative, which means that a higher sharedness value correlates with a smaller amount of time to complete the task. In other words, higher sharedness correlates with a faster task completion. Regression analysis indicates that overall sharedness explains 64% of the variance in time needed to complete the task (R^2=0.69, F(1,7)=13.4, p=0.01). The correlation of the component *knowledge of intentions* and time to complete task is even higher than that of the overall sharedness. So most likely, the overall predictive power of sharedness on team performance can be increased by assigning more weight to the component *knowledge of intentions*. The calculation of an optimal assignment of weights to the (sub-)components is beyond the scope of this paper, but in general the predictive power of sharedness on team performance can be optimized by selecting proper weights.

5 Discussion

In this section we will first discuss to the two hypotheses that we stated in the introduction. Subsequently, we will make some general remarks on the compositional measure for sharedness of mental models.

Our first hypothesis stated that Jonker et al.'s measure for sharedness of mental models can be used to predict team performance. The experiment yielded a significant result showing that overall sharedness predicted team performance in agent teams. Thus, the first hypothesis is supported by the results of the experiment.

Our second hypothesis was that the sharedness of different mental model components contributes to team performance to different extents, dependent on the context. The results show indeed that sharedness of knowledge about intentions contributes more to team performance than sharedness of world knowledge. Therefore, the first part of the hypothesis is supported by the results. From these results, however, it cannot be concluded that the contribution of a component depends on context since knowledge about intentions contributes more to team performance in all scenarios.

To summarize, from the results we can conclude that the compositional measure for sharedness of mental models can be used to predict team performance, and the components in the measure may differ in the extent to which they contribute to team performance. A potential difficulty in using the measure is how to determine the weights of the components in the measure. We distinguish two approaches addressing this problem that can be used separately or in combination. The first approach is an engineering approach, and involves making an analysis of the task, e.g. the dependencies, and based on that assign weights to the components. The second approach is an empirical approach in which the team has to discover during task execution which components are more and less important. The empirically determined weights can then be used to predict team performance in other teams.

6 Conclusion

In this paper, we investigated whether shared mental models can predict team performance in a team with members that are agents, and we investigated to what extent different components of a shared mental model contribute to team performance. For that, we performed an experiment in the BW4T testbed for team coordination. Namely, we ran several simulations with different teams of agents in different scenarios, and measured team performance and the sharedness of their mental models. We found that for the BW4T task, sharedness of knowledge about intentions better predicts team performance than sharedness of world knowledge.

In this paper we adopted the concept of SMM based on research on human teams to study agent teamwork. Our future aim is to use SMM to understand and predict the performance in human-agent teams. Therefore, our next step will be to redo the experiment performed in this paper with humans in the loop. So, instead of using teams of only agents, we use mixed human-agent teams. In such an experiment we can test whether the results of this experiment are indeed transferrable to human-agent teams.

Furthermore, we can ask additional questions to humans that involve their subjective experience of working in a team with different types of agents. We will, however, have to take into account that communication cannot be assumed to be perfect when there are humans in the loop. Another factor that we will have to take into account in an experiment with human-agent teams is trust. Humans may not have complete trust in agents, whereas the agents in our experiment assumed that the other team members were trustworthy.

After testing the compositional measure in human-agent teams, a next step will be to use these findings in a real world application in which human and agents work together in a team. Currently, we are studying crisis management, in which human workers such as police, fire-fighters and operators, collaborate with and are supported by intelligent agents. We aim to use insights about human-agent teamwork obtained in a simplified domain to better understand human-agent teamwork in real world applications as well, and based on that, improve team performance.

References

1. Bratman, M.: Intention, Plans and Practical Reason. Harvard University Press, Cambridge, MA. (1987)
2. Cannon-Bowers, J.A., Salas, E., Converse, S.: Shared mental models in expert team decision making. Individual and group decision making 39(3-4) (1993) 221-246
3. Cooke, N. J., Salas, E., Cannon-Bowers, J. A., Stout, R.J.: Measuring team knowledge. Human Factors 42(1) (2000) 151-173
4. Harbers, M., Bradshaw, J., Johnson, M. Feltovich, P., Van den Bosch, K., Meyer, J.-J.: Explanation and Coordination in Human-Agent Teams: A Study in the BW4T Testbed. In Proceedings of the COIN workshop (2011) 17-20
5. Hindriks, K.: Multi-Agent Programming: Languages, Tools and Applications. Chapter Programming Rational Agents in GOAL, Springer (2009) 119-157.
6. Johnson, M., Jonker, C., Van Riemsdijk, B., Feltovich, P., Bradshaw, J.: Joint activity testbed: Blocks world for teams (BW4T). In Proceedings of ESAW09, Springer (2009) 254-256
7. Jonker, C., Van Riemsdijk, B., Van de Kieft, I., Gini, M.: Towards Measuring Sharedness of Team Mental Models by Compositional Means, IEA, to appear (2012)
8. Jonker, C., Van Riemsdijk, B., Vermeulen, B.: Shared mental models - a conceptual analysis. In Proceedings of COIN@AAMASMALLOW (2010) 132-151
9. Mohammed, Ferzandi, Hamilton: Metaphor no more: A 15-year review of the team mental model construct. Journal of Management 36(4) (2010) 876-910
10. Mohammed, Klimoski, Rentsch: The measurement of team mental models: We have no shared schema. Organizational Research Methods 3(2) (2000) 123-165
11. Rao, A., George, M.: BDI-agents: From theory to practice. In Proceedings of ICMAS95 (1995)
12. Ross, S., Allen, N.: Examining the Convergent Validity of Shared Mental Model Measures. to appear
13. Rouse, W., Morris, N. M.: On looking into the black box: Prospects and limits in the search for mental models. Psychological Bulletin 100(3) (1984) 349-363
14. Sycara, K., Lewis, M.: Integrating intelligent agents into human teams. In Proceedings of the Human Factors and Ergonomics Society (2002)