Technical university of Dresden
Faculty of mathematics and science
Faculty of psychology
Chair of engineering psychology and applied cognitive research

MASTER THESIS

Communication during Fault Diagnosis as a Function of Shared Mental Models

Submitted by:

Rica Bönsel, born on November 17th 1991 in Ingolstadt

Student Number: 4539127

1st Corrector and supervisor: Dr. rer. nat. Dipl. Psych. Romy Müller
2nd Corrector: Prof. Dr.-Ing. Leon Urbas

Dresden, January 30th 2018

Content

I Abstract

In the process industries, control room operators (CROs) and field operators (FOs) have to closely work together to diagnose faults in the plant. There is a high risk of problems as the partners are physically separated, have access to different types of information, and must rely on verbal communication. Shared mental models are considered to facilitate team interaction, but previous research has only collected information on the similarity of mental models post-hoc. The present study directly manipulated the similarity of team members' task models and team models by varying the amount of knowledge participants received during an instruction phase. Three groups of ten teams with two participants had to jointly diagnose five technical faults in a simulated plant, and either shared a task model, a team model, or models of both team and task. Performance and communication data was evaluated. Teams that shared both mental models communicated less than teams that shared only one model. Additionally teams that shared both models violated work safety rules less often than teams that shared only the team model. Other performance measures (e.g., solution time, correctness score) did not differ between groups. These findings suggest that it can be beneficial to train all team members on the complete team task instead of teaching them only their own responsibilities.

I Zusammenfassung

In der Prozess Industrie müssen Mitarbeiter in der Leitwarte (control room operator) und Mitarbeiter im Feld (field operator) eng zusammen arbeiten um Störungen der Anlage zu diagnostizieren. Bedingt durch die physische Trennung der Arbeitsplätze, stehen den Mitarbeitern unterschiedliche Informationen zur Verfügung und sie sind auf verbale Kommunikation angewiesen. Diese Umstände erhöhen das Risiko für Probleme und Missverständnisse. Es wird angenommen, dass geteilte mentale Modelle die Interaktion innerhalb eines Teams verbessern. Allerdings hat sich die bisherige Forschung darauf beschränkt die Ähnlichkeit mentaler Modelle im Nachhinein auszuwerten. Die vorliegende Studie manipulierte die der Ähnlichkeit von Team- und Aufgabenmodel direkt mittels der Wissensmenge die den Versuchspersonen während der Instruktionsphase vermittelt wurde. In drei Versuchsgruppen erhielten jeweils zehn Teams aus zwei Versuchspersonen die Aufgabe zusammen fünf technische Störungen in einer Prozessanlage zu diagnostizieren. Dabei wurden Leistungs- und Kommunikationsdaten von Versuchspersonen, die entweder das Aufgabenmodell, das Teammodel oder beide Modelle teilten, analysiert. Teams in welchen beide mentale Modelle geteilt wurden kommunizierten weniger als Teams die nur ein mentales Model teilten. Zusätzlich verstießen die Teams die beide Modelle teilten seltener gegen Arbeitssicherheitsvorschriften als Teams die nur ein Modell teilten. Andere Leistungsmaße (z.B. Lösungszeit, Korrektheit) unterschieden sich nicht zwischen den Gruppen. Die Ergebnisse zeigen, dass es vorteilhaft sein kann alle Teammitglieder in der gesamten Teamaufgabe zu schulen, anstatt das jeweilige Training nur auf die Aufgaben der einzelnen Teammitglieder zu fokussieren.

1 Introduction

Effective teamwork and communication is an important topic in almost all professions, but it is especially important for operators who work in complex technical environments. In these domains the consequences of failures can be fatal and the work is highly susceptible for faults due to miscommunication. One such domain is the process industries. Bullemer and Laberge (2010) estimated that in 17% of recorded process industries accidents communication problems were one of the root causes. Particularly in unfamiliar and uncertain situations misunderstandings occur frequently (Hofinger, 2012). Additionally, operators in the process industries are often limited to verbal communication over phone and radio, which is especially susceptible to misunderstandings due to the lack of shared context and the absence of nonverbal cues. The dangers of poor communication become even more obvious when looking at a well-known example of a severe accident, the Texas Refinery explosion in March 2005 (U.S. Chemical Safety and Hazard Investigation Board, 2007). During the startup of a unit a raffinate splitting tower was overfilled and flammable liquid was released through pressure relief devices. The liquid was then ignited by a nearby car engine. The final report of the U.S. Chemical Safety and Hazard Investigation Board (2007) states, that the accident happened partly due to inefficient and insufficient communication among operations personnel. During startup procedures not all information is easily available and previously executed control actions are crucial to understand the current status of the plant. Yet, no clear communication guidelines that regulated the information exchange between shifts had been established. A misunderstood instruction of Lead Operator to the Board Operator during a phone call also complicated matters prior to the explosion, as it led to the closing of a valve which sent raffinate to storage, thereby deteriorating the condition of the raffinate tower. Thus, technical failures were exacerbated by miscommunication and resulted in an explosion that killed 15 people, injured another 180, and damaged equipment worth over $1.5 billion (U.S. Chemical Safety and Hazard Investigation Board, 2007). These huge losses clearly illustrate the potentially severe consequences of accidents in the process industries. With effective communication, it might have been possible to prevent the explosion or at least reduced the damage. It is therefore very important to extend research on the principles of effective teamwork and communication in complex technical environments. The present study aims to contribute to this subject by investigating the possible benefits of shared team and task mental models on remote communication within a team during fault diagnosis.

.

2 Shared Mental Models in the Process Industries

2.1 Process Industries - Working in a Complex Technical System

2.1.1 Characteristics of a Complex System

The operation of a complex system, like a process plant, places high demands on the quality of coordination and communication between workers. Process plants have been defined as causal systems that convert large quantities of raw materials into a product during a coordinated, continuous process (Lau, Jamieson, & Skraaning, 2012). A petrochemical plant for example takes crude oil (i.e. raw material) and converts it into various products (e.g., gasoline) by eliminating impurities (e.g., natural gas or salt water) and isolating the different components. A process plant displays the characteristic features of a complex technical system. Such systems are characterized by an interconnectivity of elements, dynamic effects, non-transparency, multiple goals, and social complexity (Brehmer & Dörner, 1993; Funke, 2010; Kluge, 2014; Sterman, 1994). These characteristics will now be explained in more detail,

Element interconnectivity describes that components of a plant are coupled and cannot be influenced independently from each other. This means that adjusting the setting of one device may influence the function of the whole plant (Brehmer & Dörner, 1993). If, for example, the flow rate of a pump is decreased, this will not only change the liquid level in the adjourning vessels, but also change the mixing ratio of materials in the next vessel, possibly altering the chemical properties of the process and thereby endangering product quality. Additionally, the influence of one parameter is not limited to one variable but can also influence all other variables (e.g., increasing the flow of hot liquid into a tank raises the liquid level of the tank, but also its pressure and temperature).

A process plant is a *dynamic environment,* meaning that all parameters change in real time, both through control actions of an operator and on their own (Sterman, 1994). One example for this is the influence of weather on the process. During winter the temperatures in the plant drop and slightly decrease the temperature of the material within pipes, thereby reducing its viscosity, slowing down the flow of materials through the plant, and raising the risk of clogged devices. Influences of the environment on the process are not immediately obvious and not all of them have been specified during the construction of the plant (Perrow, 1999). Many effects only become evident after a time delay, making it hard to identify the reason for a change in values. This system quality is called *non-transparency*. It indicates that the function and relation of devices within the plant and the monitored parameter changes are only comprehensible for operators to a certain extent, and the complete system cannot be fully understood by one person (Funke, 2010).

While supervising and operating a complex system like a process plant, an operator has to balance *multiple conflicting goals* (Funke, 2010). To satisfy management and customers, the need for cost- and time-efficient production must be balanced with concerns about product quality and longevity of equipment. A petrochemical plant that quickly delivers substandard fuel may suffer the loss of customers, while long delivery times may also lead to loss of customers, regardless of fuel quality. Yet, modifying equipment and processes to produce faster with constant quality may satisfy customers but increase the strain on the equipment and thereby lead to more process loss due to malfunctions and equipment failure. Consequently, it is virtually impossible to fully satisfy all goals at the same time, and constant prioritization is necessary.

Element interconnectivity, dynamic effects and non-transparency make it extremely difficult for one single operator to understand the complete system. Instead multiple operators with unique insights in parts of the system have to collaborate closely to operate the system safely and effectively. The need for the collaboration brings about the final dimension of complexity which is called *social complexity*. A team handling the operation of a plant consists of several different team members, and each one occupies a specific role and function within the plant. Good coordination and communication between team members is vital to operate the plant effectively, avoid breakdowns, and reduce redundant work (Stachowski, Kaplan, & Waller, 2009). Shared mental models can improve the quality of teamwork by functioning as a common basis for clear and effective communication, which will be illustrated in the following chapters.

2.1.2 The Operators' Tasks

The challenges of working in a complex system are manifold, and engineers and interface designers have advanced numerous technical developments and reforms that render functional relationships more transparent and facilitate the operation of the plant. Yet, research into managing the demands of social complexity in the plant has developed more slowly. Demands of working in a complex system are most obvious when a fault has occurred (Stachowski et al., 2009; Zajac, Gregory, Bedwell, Kramer, & Salas, 2014). In the process industries teams consisting of a control room operator (CRO) and a field operator (FO) have to work together closely in order to diagnose and repair technical faults. As the team partners' working spaces are physically separated they are limited to verbal communication. Here, teamwork is characterized by the work in different environmental contexts which results in the availability of different information and possible control actions for team members. Consequently, teams that engage in collaborative fault diagnosis in a process plant face not only high technical complexity, but must also deal with high social complexity. It is therefore vital to take a closer look at the roles and working conditions of CROs and FOs.

Control room operator's task. The CRO supervises and controls the production process from a remote location, the control room. Hollnagel and Woods (2005) have described it as a room with a view of the past, present and future. They illustrate that the supervision of a plant has multiple components. To control the process it is necessary to have a good understanding of the current state of the plant and keep up to date with all changes. This information is provided by the distributed control system, which bundles information (e.g., current parameter values) from all parts of the plant (Bullemer, Cochran, Harp, & Miller, 1997). In order for the CRO to interpret the present state correctly it is vital to know about past actions and compare it to past states, which provide a frame of reference. This information can be gained from trend charts, logbooks and discussion during shift turnover (Mumaw, Roth, Vicente, & Burns, 2000). In sum, the operator forms an understanding of the system with the help of past and present states of the plant, which can then be used to form expectations of what will happen in the future (e.g., the consequences of a control action) (Hollnagel & Woods, 2005). It is likely that the affordances of the distributed control system lead the CRO to develop a rather broad and functional perspective on the plant, which includes all plant units and processes but neglects the physicality of the plant. The CRO's main challenge while monitoring the process during normal operation is to identify relevant changes and disturbances in the vast amount of constantly changing parameter values (Mumaw et al., 2000). In addition to monitoring to detect anomalies during normal operation, the CRO has to make decisions on how to react when anomalies occur, perform routine and non-routine actions (e.g., equipment tests and control actions to compensate deficits) and communicate (Kluge, 2014). When dealing with a fault, the CRO has to combine all of these tasks. As soon as a disturbance is detected during monitoring (e.g., the temperature of the reactor is too high), the CRO has to decide on how to react (e.g., he could either decide to closely monitor the temperature, regulate the heater responsible for the temperature in the reactor, or send an FO to the location of the reactor to investigate). If he or she decides to send an FO, the CRO must call the FO (usually over the phone) and instruct him or her on the current situation and the task, and keep in close contact until the problem is resolved.

The field operator's task. While CROs mostly occupy themselves with the routine task of monitoring the complete plant from afar, the FO closely monitors a specific section of the plant in the field (Bullemer et al., 1997). The FO is an important source of information on the plant for the CRO, as some parameters and devices cannot be checked from the control room (Mumaw et al., 2000). In addition to taking readings of instrumentation not available to the CRO, checking the status of devices on site, and verifying the information provided to the CRO by the instrumentation, the FO can perform manual operations (e.g., closing a manual valve), and function as a 'human sensor' (Bullemer et al., 1997). This means that he or she can use his or her senses to identify potential problems (e.g., visually checking the pipes for wet spots that could indicate leaks). The vast spatial dimension of a process plant likely leads the FO to develop a rather narrow physical perspective on the plant, which is limited to his or her assigned section, but can represent this section in great detail. When a disturbance has been detected, the FO plays a crucial role in assessing the situation. He or she can test equipment, supply additional information, diagnose the problem, and respond to emergencies (e.g., putting out a fire) (Bullemer et al., 1997).

2.1.3 Fault Diagnosis in Complex Systems

Due to the special characteristics of a complex system (i.e., interconnectivity, dynamic effects, non-transparency, multiple conflicting goals, and social complexity) fault diagnosis places high demands on operating crews, especially because the consequences of a fault that has been overlooked or misdiagnosed are potentially catastrophic, as can be seen in the example of the Texas Refinery Explosion. A fault diagnosis refers to the task in which the reason for a suboptimal state of a system needs to be detected. The process can be divided into the sub-processes of representing information, generating hypotheses, testing hypotheses, and evaluating hypotheses (Abele, 2017).

This subdivision describes that in order to diagnose a fault a person has to first gather and mentally represent all available information on the problem. During this phase, people often fail to consider all available information and instead focus on the information that seems most important to them (Rasmussen & Jensen, 1974). Thus, suboptimal performance in this process leads to an incomplete mental representation of the problem. For an effective fault diagnosis in process industries, both the broad functional perspective of the CRO and the narrow physical perspective of the FO are vital, as each perspective provides unique information on the plant. Information from both operators needs to be gathered and combined into a complete representation of the problem. If the CRO is alerted that the liquid level in a reactor is too low, he should not only consider his information about the liquid level of the reactor, but also inquire about the FO's information to reveal deviations of the sensor. Furthermore, information about the state of the surrounding vessels can supply important clues on the nature and the location of the problem

On the basis of the problem representation that is formed after the initial search for information, hypotheses about the possible causes of the observed problem can be generated and tested step by step. Previous research has shown that people tend to experience difficulties in generating a complete set of hypotheses (Mehle, 1982). If the CRO does not know that the parameter values of the FO deviate from the values in the control room, he will not consider a sensor problem as a possible cause for the alarm. Furthermore, during hypotheses testing people often rely heavily on the test procedures they consider the easiest (Bereiter & Miller, 1989; Konradt, 1995) or fixate on an incorrect hypothesis for a prolonged time (Bereiter & Miller, 1989). Taken together this can lead to biases in the search for information, as operators tend to sample information that can confirm their current hypothesis (Morrison & Upton, 1994). Thus, even when the CRO is aware of a sensor deviation, he can opt to test valves and pumps first, because he considers the test procedure for the malfunction of a sensor complicated.

During the last step of the fault diagnosis the gathered information needs to be discussed and evaluated to form a diagnosis. However, difficulties can arise when a hypothesis is believed to be true or false and therefore evidence to the contrary is reinterpreted (Arocha, Patel, & Patel, 1993). One belief commonly held by operators is that instrument readings are very valid (Patrick, Gregov, Halliday, Handley, & O'Reilly, 1999) and consequently a CRO confronted with a sensor deviation might conclude the observed difference is negligible and search for other causes of the current problem.

In addition to bringing their unique perspectives to the task, it is important that the information each team member possesses is not only available for the task, but is also communicated in the right moment and understood correctly. Only when communication and coordination between CRO and FO is effective can hypotheses be tested adequately and the test results can be evaluated correctly. One promising way to facilitate the communication between CRO and FO is the promotion of shared mental models. When CRO and FO share an adequate mental representation of the system, they should be able to communicate effectively in order to identify and resolve the problem successfully (Kluge, Nazir, & Manca, 2014).

2.2 Shared Mental Models during Fault Diagnosis

Adequate mental representation, what it encompasses, how it is acquired, and which effects it has on performance has long been discussed in the literature (e.g., Cannon-Bowers, Salas, & Converse, 1993; Funke, 1985; Klimoski & Mohammed, 1994; Rouse & Morris, 1986; Stout, Cannon-Bowers, Salas, & Milanovich, 1999). Depending on the research disciplines, several similar yet distinct theoretical concepts have evolved. While some (e.g., information sharing, transactive memory) focus on the information that is distributed across team members, others (e.g., cognitive consensus, group learning, shared mental models) put more emphasis on the amount of information that is shared between team members (see Mohammed and Dumville (2001) for a differential discussion of these concepts). In the process industries, where non-transparent and complex relations call for flexible adaption and the compatibility of information from different sources for consistent fault diagnosis, shared mental models have emerged as the most suitable construct (Cannon-Bowers et al., 1993).

A *mental model* of a system is defined as organized knowledge structures that allow individuals to describe, explain, and predict the impact of an event on their environment (Rouse & Morris, 1986). The construct of a mental model can be extended to the team level; the mental model will then encompass the shared assessment of knowledge about the main features of the environment and other team members. This construct is referred to as *shared mental model* (Klimoski & Mohammed, 1994). Most researchers agree on the view that a shared mental model does not refer to a single construct within a team, but rather to the overlap between multiple mental models that co-exist among team members (Klimoski & Mohammed, 1994). Consequently team members' mental models do not need to be identical, but should generate compatible output (i.e., expectations regarding system or human behavior) (Cannon-Bowers et al., 1993). This means that depending on the current task even mental models that seem rather different can function as shared mental models. If for example operator A mentally represents a pump as a 'black box that ensures the transportation of fluid' and operator B as a 'device with a rotor powered by electricity that ensures the transportation of fluid' this will not make any difference during normal operation, as both models generate the expectation that the pump will ensure the flow of liquid. Only when the pump does not function correctly the difference in mental models becomes obvious. Operator A cannot gain any clues regarding the cause of the problem from his or her 'black box' representation, whereas operator B can hypothesize that there may be a problem with either the rotor or the electricity. This example also shows that whether or not a mental model is adequate depends on the context of its use. Depending on their content Cannon-Bowers et al. (1993) distinguish four types of mental models. The *equipment model* contains information about the characteristics and function of equipment, the *task model* describes task procedures, possible problems, and environmental constraints, the *team* *interaction model* defines roles and responsibilities of team members, information sources that are available and interaction patterns, and the *team model* encompasses information about each team member's knowledge, skills, and preferences (Cannon-Bowers et al., 1993). Mathieu, Heffner, Goodwin, Salas, and Cannon-Bowers (2000) have argued that while the distinction of the four types is useful, the operationalization of all four models in one experiment is unwieldy. They therefore propose to condense the typology to two types of mental models. One includes all task-related features of the situation (i.e., equipment and task model), which will be called *task model* from now on and the other includes all team-related aspects of the situation (i.e., team interaction model and team model), which will be called *team model* (Mathieu et al., 2000). This categorization has been widely adopted in empirical research (e.g., Cooke et al., 2003; Lim & Klein, 2006; Mathieu, Rapp, Maynard, & Mangos, 2010)

2.2.1 Effects of Shared Mental Models

The collective consensus seems to be that shared mental models have a beneficial influence on team processes and performance. They can enhance the ability to coordinate within a team (Stout et al., 1999), facilitate the adaption to changing task and team requirements (Sander, van Doorn, van der Pal, & Zijlstra, 2015), improve decision making (Kellermanns, Floyd, Pearson, & Spencer, 2008), allow for the use of more effective communication strategies (Espevik, Johnsen, & Eid, 2011) and generally raise team performance (Turner, Chen, & Danks, 2014). Cannon-Bowers et al. (1993) propose that theses beneficial effects can occur because higher similarity in mental models leads to an improved understanding of the system and each person's role in task accomplishment. Going back to the previous example, one could assume that operator A, who views the pump as a 'black box', is an inexperienced field operator and works together with the CRO operator B, who views the pump as an 'electrical device with a rotor'. They have identified a pump in the plant as non-functional. The FO now wants to close all valves connected to the pump and exchange the whole device, while the CRO would like to investigate whether the pump is connected to the electric system and if there is a foreign object blocking the rotor. Each team member then makes their suggestion on how to proceed and after they have decided to continue testing, the CRO has to instruct the FO on which information is important for each test. If both operators had shared the same mental model of the pump, their deduction of the next necessary steps would have been identical and no further planning necessary. Thus, the improved understanding provided by a shared mental model can create compatible expectations about the necessary steps for task accomplishment, each other's information needs, and the behavior of team mates. This shared expectation of future states enables team members to act, adapt and coordinate according to these expectations without the process loss of overly long discussions and conflicts about strategies. This reasoning suggests that the impact of shared mental models is especially prominent during communication and coordination, and should be examined in more detail.

2.2.2 Markers of Effective Coordination and Communication

Shared mental models, especially a shared team model, enable teammates to anticipate the information requirements of other team members better (Mathieu et al., 2000). If the team members in the previous example had shared the same mental model, the FO would have been able to deduce which information was important for the CRO and communicate it to him (i.e., that there is an inspection window on the pump and no foreign object visible). In effective teamwork, communication strategies and content are varied depending on the current circumstances and workload (Oser, Prince, Morgan, & Simpson, 1991). *Implicit coordination* has been identified as an effective communication strategy for high workload situations. While during explicit coordination information and resources are transferred after a request, implicit coordination transfers information and resources without a prior request (Serfaty, Entin, & Volpe, 1993). As implicit coordination depends on the anticipation rather than the communication of team members' needs, it requires less resources than explicit coordination (Entin & Serfaty, 1999).

In critical situations people often tend to fall silent instead of speaking up (Hofinger, 2012). One communication strategy that can counteract this tendency is to 'fly by voice', a strategy which originated in the aviation domain. It refers to the concept of routinely announcing intended actions (Brindley & Reynolds, 2011) and could therefore be termed as *performance monitoring*. When team members speak up and give explicit statements about intended actions, this gives other team members an opportunity to double check the appropriateness of the proposed action and intervene before mistakes occur instead of correcting them after the event (Brindley & Reynolds, 2011). This means that if a CRO announces that he or she intends to shut down a pump regulating the outflow of a vessel, the FO can check the pressure of the vessel and warn the CRO before the shutdown would deteriorate already critical pressure conditions within the vessel. Furthermore, it can be hypothesized that this beneficial effect of performance monitoring is only possible if team members share a task model, as the check requires the FO to have sufficient information about the task and the system to allow a correct prediction of the consequences of an action.

The combined use of these strategies becomes obvious when another example is considered. A CRO discovers that the liquid level of a reactor is slowly but constantly sinking although it should remain constant. An FO is present on site and can assist with the diagnosis. The CRO contemplates the problem by himself and directs the course of action without telling the FO any details concerning the problem or the assumed cause. The team ends up laboriously checking the liquid levels of all adjourning vessels and testing the function of all connected devices. Finally, they come to the conclusion that the sensor must be malfunctioning, as every other device seems to function correctly. The FO exchanges the sensor. During the next shift the problem appears again and another team investigates. The CRO passes his information on the problem to the FO and announces that he intends to first check all adjourning vessels for anomalies and then test the function of all plant devices (i.e., engages in performance monitoring). The FO contemplates this and then announces that he would like to check the surrounding area first. During his check he notices a pipe pressure that is slightly lower than usual, but normally would not warrant further scrutiny. Knowing the nature of the problem in the reactor he nevertheless investigates it, and discovers a leak in the connection between the reactor and connecting pipe. He relays this information to the CRO (i.e., engages in implicit coordination), the connection is repaired and the problem does not appear again. Thus, the use of performance monitoring enabled the FO to actively participate in the diagnosis and supply the needed information without prior request. Through the use of effective communication strategies the team was able to diagnose the fault quickly and correctly, they performed better than the other team.

2.2.3 What makes Shared Mental Models Helpful?

Despite the prevalent consensus that shared mental models have a beneficial effect on performance, the question on the relative importance of shared team model and the shared task model remains controversial. In their study of army combat teams Lim and Klein (2006) found that team performance could be predicted similarly by both team and task model, which implies an equal importance of both constructs. In contrast to this Mathieu et al. (2000) studied dyadic teams completing missions in a low-fidelity flight combat simulation and could not identify direct relation between task model and team processes or performance. A shared team model, however, correlated with more effective team processes and better team performance. This implies a differential importance of the constructs which favors the effects of the team model. Mathieu et al. (2010) further investigated these findings and found that only shared task models exhibited a direct relationship with team effectiveness, while shared team models did not and thus favored the importance of the task model. Additionally, they concluded that the interaction of team and task models was positively related to team effectiveness, a finding that is echoed by Smith-Jentsch, Mathieu, and Kraiger (2005), who could only identify an influence of the interaction but no direct effects of both models on team performance. In sum these conflicting results could lead to the conclusion that while further research effort is necessary to determine which construct is more important, having a shared task model as well as a shared team model generates additional benefits.

Recently, various theoretical frameworks for the influence of shared mental models on team processes and performance have been proposed. Although the proposed explanatory mechanisms differ widely, most share an input-process-outcome (I-P-O) framework of team functioning (Mathieu et al., 2000; Uitdewilligen, Waller, & Zijlstra, 2010). The I-P-O framework postulates that the input variables to a team task (e.g., task difficulty) influence team processes (e.g., coordination) and that these processes in turn influence the quality of outputs (e.g., quality of fault diagnosis). Previous research has mostly looked at shared mental models as team processes (e.g., Espevik et al., 2011) or outputs of a team task (e.g., Stout et al., 1999). The team process view sees the construction of shared mental models as a product of the characteristics of the input, meaning for example that the familiarity of team members which each other is assumed to determine the similarity of mental models (Espevik et al., 2011; Espevik, Johnsen, Eid, & Thayer, 2006). The output view, however, sees high similarity in shared mental models as a product of team processes, for example by showing, that a higher quality of planning in teams leads to higher similarity of mental models (Stout et al., 1999). This entails that in most studies shared mental models and their changes were only assessed, but not manipulated directly.

One notable exception to this is a study by Sauer, Felsing, Franke, and Ruttinger (2006), who manipulated cognitive diversity within teams and observed the effects on communication and various team performance measures. Teams were trained for 4.5 hours in different levels of system understanding (i.e., procedure-oriented vs. knowledge-oriented) and specialization (i.e., general knowledge of all faults vs. specialized knowledge of half of the faults). This manipulation of cognitive diversity can be equated with the direct manipulation of similarity of shared mental models. The use of the term cognitive diversity represents their theoretical standpoint that more diverse mental models (i.e., cognitive diversity within a team) should be beneficial for team performance. During the experiment teams of two participants jointly worked in a simulated highly automated process control environment, where they operated a spacecraft's life support system. One of the teams' tasks was fault diagnosis and repair. The results suggest benefits of a deep system understanding (i.e., a more complex mental model), as teams in which at least one partner had been trained in a knowledge-oriented manner took less time to diagnose and repair faults and spent less time in an unsafe system state. Additionally, in generalist teams (i.e., teams with a shared model of fault scenarios) the system also spent less time in an unsafe state. Analyses of the communication process revealed that teams in which both partners had deep system knowledge as well as generalist teams communicated more.

Against the backdrop of current research the examination and direct manipulation of shared mental models seems a promising approach. Treating shared mental model as input variable to a task, their influence on the team processes of communication and diverse team outputs (i.e., error rate, time on task) can be investigated. The contemplation of this chain of factors stands in concordance with current research results that team processes may fully or partly mediate the influence of shared mental models on team outcomes (Marks, Sabella, Burke, & Zaccaro, 2002; Mathieu et al., 2000; Mathieu et al., 2010; Mathieu, Heffner, Goodwin, Cannon-Bowers, & Salas, 2005). This means that a higher similarity of shared mental models may have a direct effect on team performance or an indirect effect that is mediated by the quality of team communication. Furthermore this approach could shed new light on the as yet unanswered question which mental models need to be shared to achieve optimal performance.

2.2.4 Acquisition and Training of Shared Mental Models

Cannon-Bowers et al. (1993) suggest that training of team and task mental models is possible. While research on shared mental models has largely focused on the observation of antecedents and process variables that foster the development of shared mental models on the job or during the interaction of selected teams (e.g., Langan-Fox, Anglim, & Wilson, 2004; Rasker, Post, & Schraagen, 2000; Stout, 1995; van den Bossche, Gijselaers, Segers, Woltjer, & Kirschner, 2011), clues on how mental models can be instructed and thereby manipulated directly can be found in the literature on the development of individual mental models. Specific instruction can further the development of adequate mental models and reduce their variability between persons, thereby creating similar mental models in team members (i.e., shared mental models) (Cannon-Bowers et al., 1993; Edwards, Day, Arthur, & Bell, 2006). Rouse and Morris (1986) stress that the mere knowledge of theories and principles is not enough to construct an adequate mental model of a system, but that hierarchical organization of instruction content, guidance, and cuing on how and when to use this knowledge effectively is also necessary to improve performance. Cannon-Bowers et al. (1993) conclude that a presentation of explicit conceptual models during the instruction can also improve the acquisition of mental models. Additionally, increasing team members' understanding of not only their own but also the other team members' roles and responsibilities is an important aspect (Marks et al., 2002). Different levels of cross-training from the provision of information on all team roles to training all team members in all team positions have been found to further the development of similar mental models (Marks et al., 2002).

2.3 Present Study

This study aims to contribute new insight on the topic of the relative importance of shared task and team models by investigating shared mental models as an input to a collaborative fault diagnosis task between a CRO and an FO in a simulated process plant. As introduced before, the physical separation of team members results in an availability of different information and a reliance on verbal communication. A fault diagnosis task in this environment is therefore at high risk of mistakes and misunderstandings, making the attempt to improve cooperation highly relevant.

To evaluate the impact of shared mental models on the team process of communication and team performance a two-factorial between subjects design was used, which included the factor shared mental model group (i.e., SMM group) with the factor levels shared task model (Task SMM group), shared team model (Team SMM group), and shared task and team model (Task & Team SMM group) and the factor role with the levels CRO and FO. Shared team and task models were operationalized as shared knowledge on the team and task aspects of the experimental plant. The experiment consisted of two sessions. The first session served as manipulation of shared mental models. Participants were introduced to the system in groups during an instruction seminar. Each group learned the information about the team and task characteristics of the system relevant to their SMM group and role. As knowledge about the process industries was generally low and the plant simulation had not previously been used, participants had very little to no previous shared or conflicting knowledge. Thus shared knowledge and consequently shared mental models should depend solely on the manipulation. During the second session (i.e., the main experiment) a team of two participants, one acting as CRO and one as FO, jointly diagnosed five technical faults in a plant. To achieve this goal the CRO worked on a computer with a simulated plant, while the FO stood in front of a drawing that depicted the field view of the same plant. Both partners could access unique and shared information on the current state of the plant. In order to diagnose the fault correctly, verbal communication was necessary and numerous alternative reasons for the alarm had to be ruled out. The main experiment was supervised by two experimenters and audio and video as well as performance data was recorded.

As there is little precedent for the effects of the direct manipulation of shared mental models on communication and performance, a set of general hypotheses to drive exploratory evaluation of the data is proposed.

Different types and combinations of shared mental models have been shown to influence communication in a unique way. Therefore it is predicted, that the content and amount of communication exchanged between CRO and FO will differ between experimental SMM groups (*Hypothesis 1*).

In essence, the literature indicates that higher similarity of mental models enhances a team's ability to communicate and perform well. Thus, it is assumed that in the SMM group that shares more knowledge (i.e., Task & Team SMM group) more effective communication strategies will be used than the groups that share less knowledge (i.e., Team SMM group, Task SMM group) (*Hypothesis 2).* As a result the group that shares more knowledge will also perform better in the fault diagnosis task (*Hypothesis 3*).

As participants face a new and complex task, the task model is assumed to be the most important factor for goal-oriented task participation. This means that if the task model is shared, both partners can participate equally, and if the task model is not shared (i.e., in the Team SMM group) the CRO will have a more dominant role in the interaction compared to the other SMM groups, leading the CRO to be more directive in his or her communication and to take a higher proportional share in communication (*Hypothesis 4*).

If the team model is shared, the partner's information needs can be anticipated more accurately. This will result in more implicit coordination between team partners if the team model is shared (i.e., in the Team and the Task & Team condition) than when it is not shared (*Hypotheses 5*). If the team model is not shared (i.e., in the Task SMM group) only a very small amount of implicit coordination is possible, making more communication necessary compared to the other SMM groups (*Hypothesis 6*).

3 Method

3.1 Participants

Participants were recruited via the data base for participants of the Technical University of Dresden. In order to be eligible for the study, they were required to speak and understand German on a native level. The participants received 15€ or partial course credit as a compensation for their participation. It was intended to collect data on 10 teams per experimental SMM group resulting in an intended total of 60 (i.e., 3 x 10 x 2) participants. In order to accomplish this, a total of 72 people started the experiment. Five people failed to complete 75% of the knowledge test at the end of the instruction seminar correctly and were therefore excluded from the main experiment. Due to illness or other reasons five people were unable to participate in the second session. One team demonstrated a severe lack of system understanding during the main experiment and was consequently excluded from the data analysis and replaced by another team. The remaining 60 participants were mainly students (49 participants, 81.7 %) and female (39 participants, 65.0 %) the mean age was 25.88 years (*SD* = 5.45, min = 19, max = 46). There were three male teams, 12 female teams and 15 mixed teams. On a five-point Likert scale (1 = lowest, 5 = highest rating) participants rated themselves as having little to no prior knowledge about the process industries (*M* = 1.29, *SD* = 0.65) and estimated that they did not know their team partner well (*M* = 1.16, *SD* = 0.49).

3.2 Apparatus and Stimuli

3.2.1 Instruction Seminar

The manipulation was imparted during a mandatory instruction seminar, which consisted of three parts. First, there was an introductory section which detailed the experimental procedure and gave an overview of the process industries and the roles of CRO and FO. This section remained completely the same for all participants, while all further sections (plant equipment and function of the plant, fault diagnosis in the experimental plant) contained some general content and additional content belonging to either the team model or the task model. Therefore, these sections varied according to SMM group and role (see *Table 1* for a detailed overview). The combination of role and SMM group resulted in four different versions of the instruction seminar (i.e., task model & role of CRO, task model & role of FO, no task model & both roles, and task model & both roles).

Table 1
Overview of the content of the instruction seminar, divided according to sections of the instruction seminar, SMM group and role

	Role	Task SMM group	Team SMM group	Task & Team SMM group
Introduction	CRO & FO	• experimental procedure • introduction to process industries • working conditions of CRO and FO	• experimental procedure • introduction to process industries • working conditions of CRO and FO	• experimental procedure • introduction to process industries • working conditions of CRO and FO
Plant equipment and functional principles of the plant	CRO	• characteristics of plant equipment • depiction of plant equipment in the *control room* • information and operation available in the *control room* • functional principles of an intact plan	• characteristics of plant equipment • depiction of plant equipment in the *control room* and the *field* • information and operation available in the *control room* and in the *field* • functional principles of an intact plan	• characteristics of plant equipment • depiction of plant equipment in the *control room* and the *field* • information and operation available in the *control room* and in the *field* • functional principles of an intact plan

	FO	• characteristics of plant equipment • depiction of plant equipment in the *field* • information and operation available in the *field* • functional principles of an intact plan	• characteristics of plant equipment • depiction of plant equipment in the *control room* and the *field* • information and operation available in the *control room* and in the *field*	• characteristics of plant equipment • depiction of plant equipment in the *control room* and the *field* • information and operation available in the *control room* and in the *field* • functional principles of an intact plan
Fault diagnosis in the experimental plant	CRO	• general process of fault diagnosis • range of consequences of a fault on plant function • common faults and their consequences on plant function • work safety rules	• general process of a fault diagnosis • range of consequences of a fault on plant function • common faults and their consequences on plant function • work safety rules	• general process of a fault diagnosis • range of consequences of a fault on plant function • common faults and their consequences on plant function • work safety rules
	FO	• general process of fault diagnosis • range of consequences of a fault on plant function • common faults and their consequences on plant function • work safety rules	• general process of fault diagnosis	• general process of fault diagnosis • range of consequences of a fault on plant function • common faults and their consequences on plant function • work safety rules

The *team mental model* contained information about competencies of the team members and the characteristics of their work site. Competencies of a team member encompassed information a team member could collect in the system (e.g., a CRO could access liquid level and pressure of a tank, but not temperature), and which control actions a team member could execute (e.g., a CRO could change the state of pumps, mixers and heaters, while the FO could only change the state of valves) (see *Table 2* for an overview of available information and operations). In addition, participants received information about the way the plant equipment was depicted in their representation of the plant. If the team model was shared, CRO and FO received information about their own as well as their team partners' abilities and both representation of the system. If the team model was not shared the team members only received information about their own abilities and own representation of the system. Consequently, a shared team model enabled participants to gain a concrete impression of their partner's working conditions and gather an understanding of his or her partner's abilities to support the fault diagnosis in the experiment.

Table 1
Overview of available information and operations (in italics) in the control room and in the field

Plant equipment	Information and operation in the control room	Information and operation in the field
Reactor	• liquid level • temperature • pressure	• liquid level • temperature
Tank	• liquid level • pressure	• temperature
Pump	• state (on/off) • *change state*	• state (on/off)
Valve	• none	• state (open/closed) • *change state*
Heater	• state (on/off) • *change state*	• state (on/off)
Mixer	• state (on/off) • *change state*	• state (on/off)
Pipes	• none	• pipe pressure • leaks

The *task mental model* encompassed information about the function of the system, common faults, and work safety rules. If participants shared the task model, CRO and FO both received this information about the system, but if the task model was not shared only the CRO received this information, as at all times at least one person within the team had to have specific information on the task to allow them to diagnose a fault correctly. Participants learned how the experimental plant was intended to function. Three simple rules defined the flow of materials through every plant in the experiment. First, every inflow into one part of the plant (raw material or material from another part of the plant) amounts to 2l/s. Second, in a tank or a reactor there is the same amount of inflow as there is outflow (i.e., mass balance) and third, a valve splits the stream of material equally into all connected pipes. These three rules allowed participants to determine the flow of material through every element in the plant. If the flow in the plant differed from these rules, the reason was a fault. The participants were also taught common faults that could occur during the experiment (e.g., a pump could be damaged), and how these faults would change the function of the system (e.g., if a pump is damaged it has less throughput than it should when it is switched on, and will not be able to stop the flow of the liquid completely when it is switched off). Finally participants were introduced to two work safety rules that prohibited certain actions, when a specific problem had occurred within the plant (e.g., if there is an alarm that the liquid level of a tank is too low, actions that exacerbate this situation are not permitted. This means that pumps and valves that regulate the inflow should not be turned off and pumps and valves that regulate the outflow should not be turned on).

The instruction seminar was structured according to instructional principles to give participants very good conditions for learning the big amount of information in a short time period. Techniques used included test questions (Nungester & Duchastel, 1982), and animations of slides to reduce the risk of visual and cognitive overload (Mayer & Moreno, 2003). Furthermore, techniques that facilitated the development of accurate mental models were employed. Advance organizers (Mayer, 2008) and preview and summary slides were used to structure the material (Rouse & Morris, 1986), analogies were presented to encourage the development of an accurate task model (Cannon-Bowers et al., 1993) and information on the team partner's role, responsibilities and task were presented to encourage the development of an accurate team model (Marks et al., 2002). (Refer to the data CD for the complete instruction seminar and its versions).

3.2.2 Knowledge Test

The knowledge test tested participants on how much of the content of the instruction seminar they had managed to learn. It was possible to answer all questions correctly only through participating in the instruction seminar. No prior knowledge was necessary. There was only one version of the test, regardless of the experimental SMM group. This test consisted of 13 multiple choice questions and one question that asked participants to calculate the flow of material within the plant. All information relevant for the experiment was included in the test. A cutoff of 75% of correct answers was used for the admission to the main experiment to ensure that participants had understood and retained most of the seminar content. For participants that did not receive all possible information only questions on the information presented in the instruction seminar were evaluated. This means that FOs in the Team SMM group could reach 100% by answering all questions on the team model (i.e., all questions on the information and operations available in the control room and in the field) correctly.

3.2.3 Main Experiment

During the main experiment a pair of two participants, one functioning as a CRO and one as an FO, had to jointly diagnose a technical fault in the plant. The experiment consisted of a training scenario and five experimental scenarios. The scenarios concerned different plants and different faults. Participants were able to communicate freely with each other, but were divided by a screen that prevented them from seeing each other and their team partner's system representation (See *Figure 1* for a schematic overview of the experimental setup).

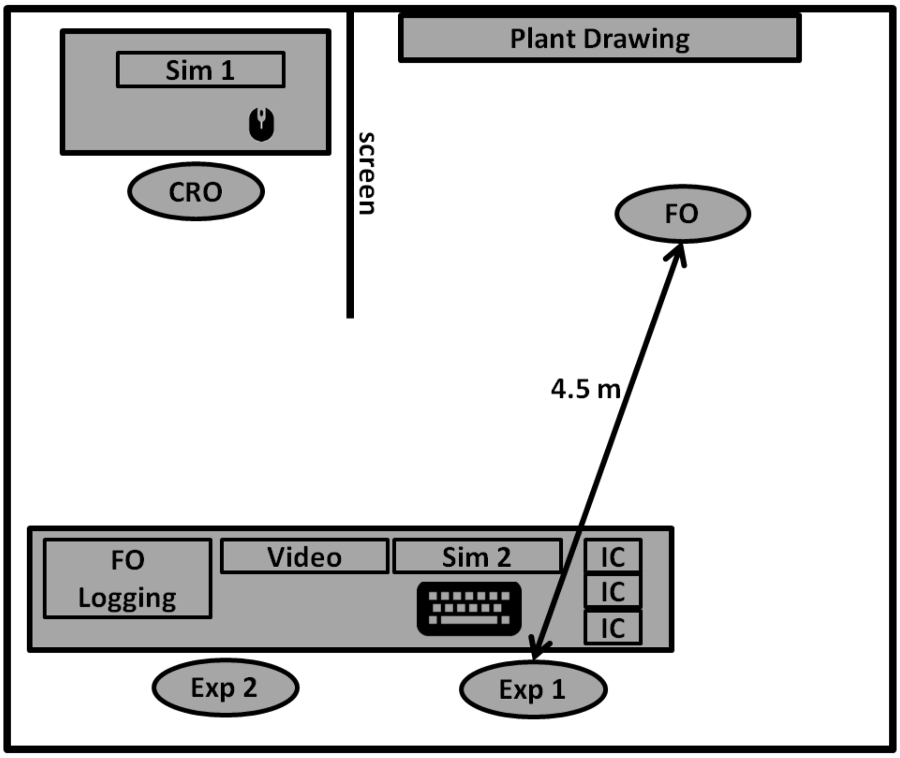

Figure 1. Schematic overview of the experimental setup. Exp 1/2: Position of experimenters 1/2; Sim 1/2: Position of simulation screen 1/2; IC: Interaction card.

Apparatus. The main experiment was recorded by two stationary cameras fixed to the ceiling and one mobile camera on a tripod *(Canon HD Legira HFR26)*. One camera recorded the computer screen of the CRO, while the actions of the FO were recorded from two different angles. Audio data was recorded by an ambience microphone. The recordings were controlled by *Mangold Lab Suite* (Version 2011, *Mangold International GmbH*, Arnstorf, Germany). The CRO worked with an interactive simulation of a plant created with *Experiment Builder* (*SR Research Ltd*., Osgoode, ON, Canada). The simulation ran on a computer with two 23 inch color monitors (Model P236H, 1920 x 1080 pixels, *Acer Inc.,* New Taipei, Taiwan). The plant was controlled by the CRO with the help of a wireless mouse (Model M705, *Logitech International S.A.,* Lausanne, Switzerland) and the main screen that depicted the plant simulation. The second screen and the keyboard were placed in front of the first experimenter. To include the influence of valves on the material flow in the system, valves were operated by the experimenter via the keyboard, whenever the FO gave the appropriate signal. The FO stood at an approximate distance of 4.5 meters to the experimenter. During the experiment all actions of the FO were logged on a separate computer by the second experimenter.

Scenarios. The main experiment consisted of five scenarios in randomized order. The identity of a scenario was coded by a combination of a letter and a number in the control room view and an animal in the field view. Each scenario consisted of a different plant in which an alarm had occurred due to a fault of the equipment. Ten different hypotheses were presented in randomized order for each scenario as possible causes for the alarm. Five to seven of these hypotheses were likely causes for the alarm that occurred (e.g., if the liquid level of a vessel is too low, the reason is likely to be a clogged inflow valve), while the others were highly unlikely (e.g., a damaged heater could not lead to a low liquid level). Only one of these ten hypotheses was correct, all others could be excluded by testing the function of the plant equipment (see *Table 3* for an overview and the *data CD* for a full listing of scenarios).

Table 3
Overview of scenario codes, alarms and correct hypotheses

Scenario Code (Control room/ field)	Alarm	Correct hypothesis
A 3 / Squirrel	'Liquid level in B002 is too high'	Wrong connection of pipes
D 3 / Salamander	'Liquid level in B003 is too low'	V005 is clogged
D 2 / Caterpillar	'Liquid level in B003 is too high'	B003 liquid level sensor CRO is defective
F 3/ Cat	'Liquid level in R001 is too high'	R001 liquid level sensor CRO is impaired
F 1/ Mouse	'Alarm only visible for the FO' Red light on in B003	H001 is defective

The control room view. The simulated system mimicked the function of a real plant, but was adapted and facilitated to improve system understanding in naive participants. The system was characterized by throughput of material, liquid level, temperature, pressure in vessels, and the current state of valves, pumps, heaters, and mixers. All parameters interacted with each other to simulate a complex technical system (e.g., the amount liquid entering into a vessel influenced its liquid level and pressure, the temperature of the liquid influenced temperature and pressure within the vessel and the amount of influence was determined by the current state of the vessel). To start testing the system one of ten presented hypothetical causes for the alarm needed to be selected. Only after a hypothesis had been selected the simulation became active and all parameters started to change in real time. The values visible to the CRO were refreshed every second depending on the current state of the plant. All current parameter values were also presented on a second screen in front of the first experimenter. They were refreshed every five seconds.

The starting point of the simulation was the hypotheses screen (see *Figure 2*). This screen showed the active alarm message at the top. On the left of the hypothesis screen there was a pipe and instrumentation diagram (P&I diagram) which showed a schematic overview of the whole plant. All plant equipment was labeled with a letter coding for the type of object and a number (e.g., P001 coded for pump number one). A combination of letter and number existed only once in a scenario. The P&I diagram allowed the CRO to identify all plant equipment within the plant and its connections. On the right side of the screen there was a list of ten hypotheses. On the hypotheses screen the CRO could not gather any more information than was depicted. It was necessary to select a hypothesis to continue with the system diagnosis. As soon as a hypothesis had been selected, the simulation switched to the plant screen.

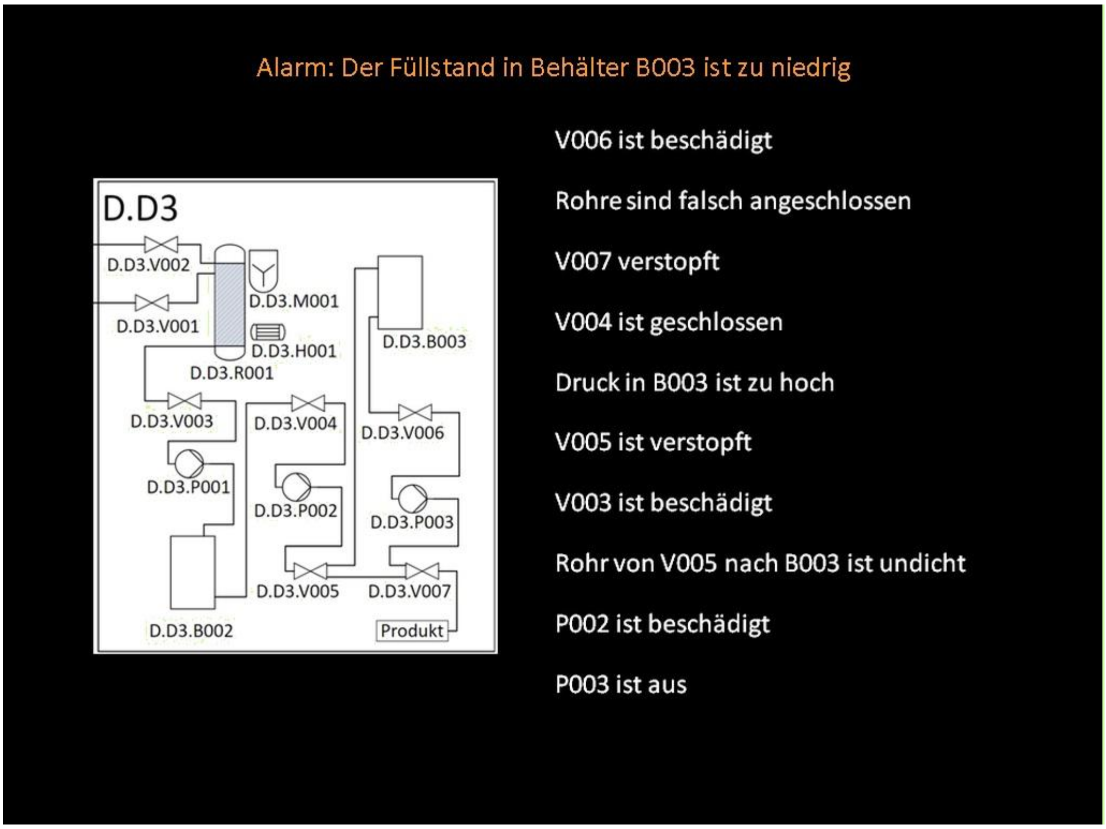

Figure 2. Hypotheses screen
Active alarm message at the top of the screen, P&I diagram on the left and list of ten hypotheses on the right.

On the plant screen (see *Figure 3*) the CRO could interact with the simulation. In this view the active alarm message and the currently selected hypothesis were displayed at the top of the screen. On the left side the P&I diagram was visible, and in this screen it was possible to select one element at a time in order to display information about the element. If a vessel was selected, a sector that displayed the available information appeared on the upper right side of the screen. The amount of information displayed varied according to the selected element. Whenever a vessel was selected the parameters temperature, liquid level, and pressure were shown, although the temperature of a tank would always be displayed as 0.0°, as the temperature of a tank was not available in the control room. When a pump, a mixer, or a heater was selected, information on the status (i.e., on or off) as well as a button that could change the status of the element appeared on the lower right side of the screen. If a valve was selected, the element was only marked within the layout of the plant, but no additional information became visible. The button 'Freigabe' (i.e., clearance) was always present. It served as a way to enable the FO to change the state of valves, which he or she could only do with clearance by the CRO. Apart from the sections that allowed the CRO to monitor and operate the plant, there were two more buttons in the bottom right corner of the screen. These allowed the CRO to reject or accept the current hypothesis.

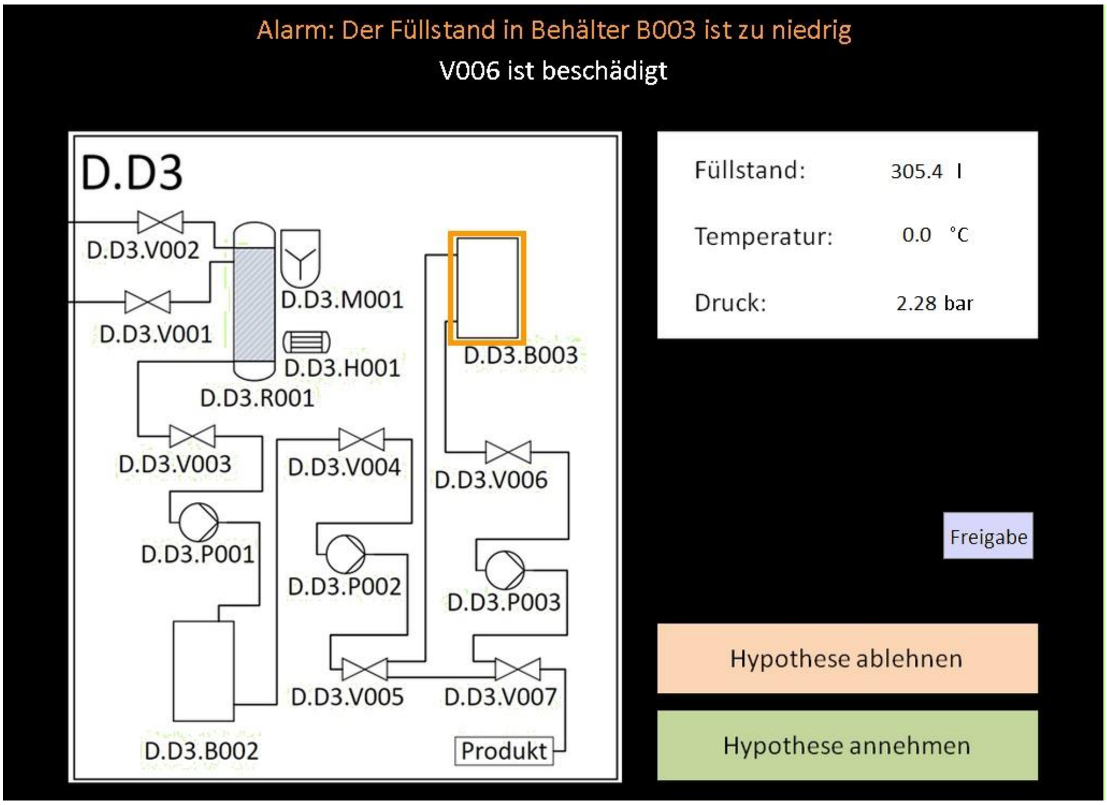

Figure 3. Plant screen

The active alarm message was displayed at the top and the currently selected hypothesis was displayed directly below. P&I diagram with selected plant equipment (marked by the orange rectangle) was on the left side of the screen. Current values of selected vessel were displayed on the upper right side. Buttons to accept and reject a hypothesis were in the bottom right corner.

The field view. While the CRO worked with a computer simulation of the plant, the FO worked with a very different representation of the system (see *Figure 4*). For this representation each P&I diagram of the plant was rendered into a 120 x 140 cm drawing that included all elements also visible on the diagram. An element was always depicted with the same characteristic marking (e.g., a tank was symbolized as a rectangle with a vertical violet stripe running through its height, and a control panel). The individual labels of the plant equipment mirrored those on the P&I diagram, thus the CRO and FO could ensure they were talking about the same object by naming it as it was labeled. The layout of the drawing differed from the orderly schematic view of the P&I diagram, but all plant equipment was included and had the same connections, except for scenario A3 where a wrong connection was the fault that had to be diagnosed. In each drawing a small animal which served as a code for the scenario was included. There could also be additional piping with valves that ran through the picture and served as distraction to make the field view more realistic. It was not relevant for the scenario.

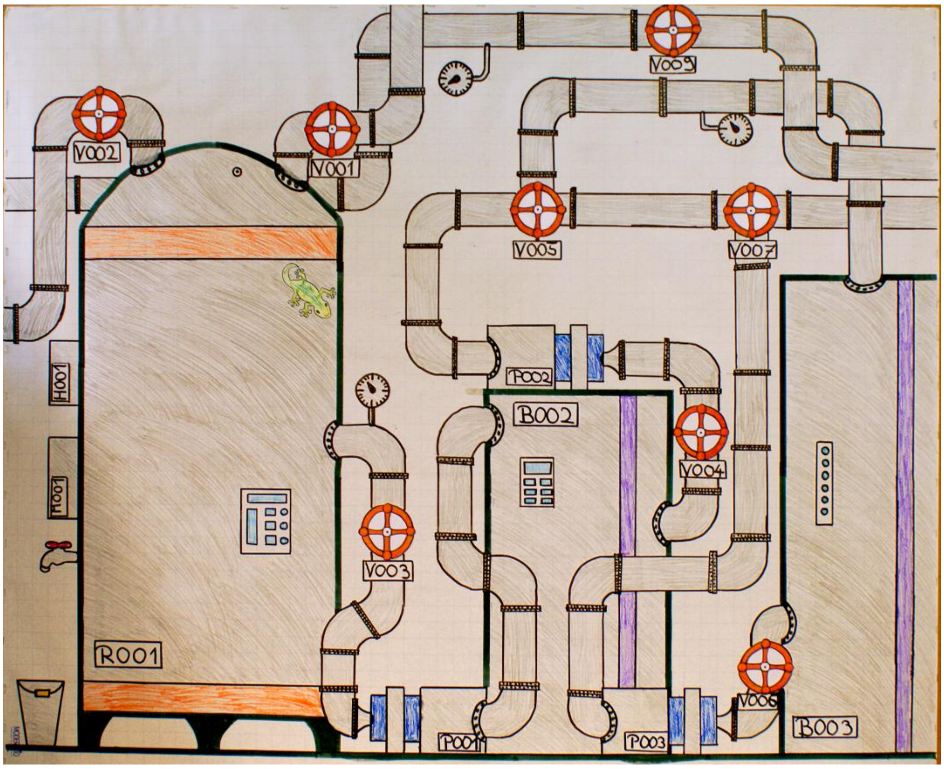

Figure 4. Field view

Each type of plant equipment was depicted with characteristic markings and labels that mirrored the labels on the P&I diagram of the CRO. A drawing of a small animal (salamander) was included as scenario code and additional piping (at the top, left to right) was included.

Additional materials. It was important that long term memory performance was not the main factor determining the amount of shared knowledge and participants could orient themselves within the system. Therefore, participants had access to a summary that listed the information of the instruction seminar that was most relevant for the experiment. It included information from the task model as well as information from the team model, depending on the SMM group of the participants. Additionally, both team members were presented with the same list of hypotheses for the current scenario as well as a list of all set points for the scenario.

Questionnaires. All questionnaires used were designed and administered with *LimeSurvey (*Version 2.67.3 *LimeSurvey GmbH,* Hamburg, Germany). The first questionnaire included questions on the participants' demographics (e.g., age, gender, education and occupation) and control variables (e.g., familiarity with team partner, previous knowledge on process industries, skill in working with computers, and general technical abilities). A second questionnaire collected information on the participants' experience during the experiment. This questionnaire was divided into two sections. The first asked for information about strategy during the selection and testing of hypotheses. It included questions on priorities during the experiment, the procedure of planning and executing the test of a hypothesis, and the subjective importance of the available information during the procedure. The second section was concerned with cooperation strategy. It asked for information about the relationship between CRO and FO, their hierarchy, who took the initiative in testing hypotheses and interpreting the plant behavior, and the perceived helpfulness of the partner. The questionnaire was available in two versions which differed only with regard to phrasing in order to reflect the perspective of CRO and FO (Refer to the *data CD* for the complete questionnaires).

3.3 Procedure

At the start of the first session participants were informed about the conditions of the experiments and gave their written consent. Then they took part in the instruction seminar as a group of maximum 15 participants and received one of the four versions of the instruction seminar. Three versions lasted approximately 60 minutes, while the forth which contained information on both roles but no information on the task lasted only about 30 minutes. As long as it was not predetermined by the content of the seminar, participants were not informed about which of the two roles (i.e., CRO, FO) they would perform during the main experiment. The seminar was interactive and structured to allow participants to ask questions and regularly check their understanding. Questions were answered, as long as the answer did not reveal knowledge that was not available in the experimental SMM group.

Directly after the instruction seminar participants were asked to complete the knowledge test. If a participant had given enough correct answers, he or she was allowed to take part in the main experiment. Participants who did not pass the test were also notified and given the opportunity to either repeat the seminar and test at the next available date for their SMM group or make arrangements with the experimenter to receive compensation for the time they had already spent on the experiment.

The second session (i.e., the main experiment) was supervised by two experimenters. When participants arrived for the second session, the first to arrive was assigned to role of FO, unless distribution of roles had been determined by the content of the instruction seminar. At the start of the session, participants answered the first questionnaire. Following this, CRO and FO were introduced to their workspace. The field view was covered until the CRO had taken his or her place on the other side of the screen to ensure the CRO did not receive any additional information about the characteristics of the field view before the experiment. The instruction included a brief instruction on how to interact with the experimental system and an explicit notification that the fault should not be repaired but only diagnosed, and that all communication was allowed. The instruction of CRO and FO was executed simultaneously by the two experimenters to minimize the participants' ability to overhear information about their team partner's role and experimental system. After the instruction was completed the participants started their evaluation of the plant. The teams were first confronted with a practice scenario, during which they received the opportunity to orient themselves within the plant and practice the interaction with each other and their plant representation. They were allowed to take as much time as they wanted to come up with a strategy of testing, but received hints on the interaction with the system if necessary. Only when participants had no idea on how to proceed they received general hints on the task, for example the hint that it was allowed to interact with the system by changing the status of plant equipment and that the liquid level of tanks and reactors and the way they changed could be interpreted. They did not receive any specific hints on test strategies or interpretation of certain system behaviors. As soon as the participants accepted a hypothesis in the practice scenario the main experiment started. The team did not receive any feedback on the correctness of their selected hypothesis.

During the main experiment hints by the experimenter were reduced to a minimum and only specific system behaviors (e.g., shutdown of pumps when a tank ran empty or that the faults could not be identified correctly if too much time had passed since the start of the scenario) were explained. The CRO interacted with the plant simulation by clicking on plant equipment and checking the available parameter values or status information as well as changing the status of mixers, heaters, and pumps. The FO stood in front of the field view and was also able to analyze the system and gather information. In order to interact with the system the FO pointed to the specific plant device he or she wanted to check and one experimenter would check the current value in the readout of the plant values and write it on an interaction card. This card was presented to the FO and thus he or she was able to check all parameters available to him or her, albeit with a short delay. At the same time the other experimenter would log the FO's action into the system. The FO could check the parameter values that were not available to the CRO as well as cross-check some information available to the CRO.

At the start of a scenario the CRO was always presented with the current alarm message and a list of ten hypotheses that could be possible causes for the alarm. To facilitate effective teamwork the CRO had to pass on the alarm to the FO. Barring guessing, it was not possible to diagnose a fault correctly without the help of the team partner. To end a scenario the CRO had to accept one of the hypotheses as correct. After the decision the plant simulation was paused in order to switch to the next scenario. This meant the participants received a new set of hypotheses, a new list of plant set points, and the field view was changed to the next plant. Then the simulation was reactivated and the team was informed that they could continue.

After the completion of the last experimental scenario the participants were asked to fill out the second questionnaire on their experience during the experiment. After they were finished, participants received compensation, feedback on their performance during the experiment, and were informed about the aim of the study.

4 Results

Data from the knowledge test, log files of the interaction of all teams with the plant simulation and all system states, audio and video data of the main experiment and questionnaire data were obtained. Data editing and analysis was conducted with *SPSS Version 21.0 (IBM Corp.,* 2012*).* The Shapiro-Wilk-test (Shapiro & Wilk, 1965) was employed to test all data for violations of normal distribution in small samples. Normal distribution was not given in all groups. Nevertheless, analyses of variance (ANOVAs) were used to analyze the data as these are regarded as very robust against the violation of normal distribution, especially when group sizes are equal (DeCarlo, 1997; Schmider, Ziegler, Danay, Beyer, & Bühner, 2010).

Due to technical problems the first few minutes of the audio and video data of two teams was missing. To extrapolate the number of missing statements from the recorded statements, they were multiplied by a factor reflecting the audio recording duration’s share in the total duration of the scenario.

4.1 Manipulation Check

The knowledge test, filled out by participants after completion of the instruction seminar, functioned as a manipulation check. To evaluate whether participants in the SMM groups differed in their knowledge of information about team and task model, the questions of the knowledge test were categorized as belonging to either the team or the task model. Each response option in the multiple choice test was treated as a decision and a correct response equaled one point. The team model encompassed four questions with a maximum score of 40 points, whereas the task model encompassed 10 questions with a maximum score of 87 points. To allow for comparisons of knowledge about the different models, all scores were analyzed as percentage scores and compared using a mixed-design ANOVA with the between subjects factors SMM-type (Task vs. Team vs. Task & Team) and role (CRO vs. FO) and the within factor knowledge type (task model vs. team model). Differences between individual groups were assessed using Bonferroni-corrected paired comparisons.

The analysis revealed a significant main effect of SMM group, $F(2,54) = 6.78$, $p = .002$, $\eta^2_p = .20$, indicating that knowledge in the Team SMM group was lower than in the Task & Team SMM group, $p = .002$ (80.9 vs. 87.9%). However, there was no significant main effect of role, $F < 1$, which indicated that CRO and FO did not generally differ in their knowledge. Additionally, there was a significant main effect of knowledge type, $F(1,54) = 35.78$, $p < .001$, $\eta^2_p = .40$, meaning that participants had more knowledge about the task than about the team model (87.8 vs. 80.4%). Furthermore, the significant interaction effect of SMM group and role, $F(2,54) = 3.88$, $p = .027$, $\eta^2_p = .17$, shows that CRO and FO differed significantly only in the Team SMM group, $p = .027$, with CROs reaching a higher score than FOs (83.9 vs. 77.8%). There was also a significant interaction effect of knowledge type and SMM group, $F(2,54) = 15.06$, $p < .001$, $\eta^2_p = .36$, indicating that participants in the Team SMM group had less task knowledge than in the other SMM groups (80.2 vs. 91.3 and 91.9%), $p < .001$, and participants in the Task SMM group had less team knowledge than in the Task & Team SMM group (76.0 vs. 83.8%), $p = .020$. There was a significant interaction effect of knowledge type and role, $F(1,54) = 22.70$, $p < .001$, $\eta^2_p = .30$, indicating that the CRO had more task knowledge than the FO (90.8 vs. 84.8%), $p = .001$, and less team knowledge than the FO (77.6 vs. 83.3%), $p = .015$. There was a significant interaction effect of knowledge type, SMM group, and role, $F(2,54) = 19.00$, $p < .001$, $\eta^2_p = .41$, indicating that knowledge type in the Team SMM group was influenced differentially by role, with the CRO having more task knowledge than the FO (91.6 vs. 68.9%), $p < .001$ and less team knowledge than the FO (76.3 vs. 86.8%), $p = .009$ (see *Figure 5*).

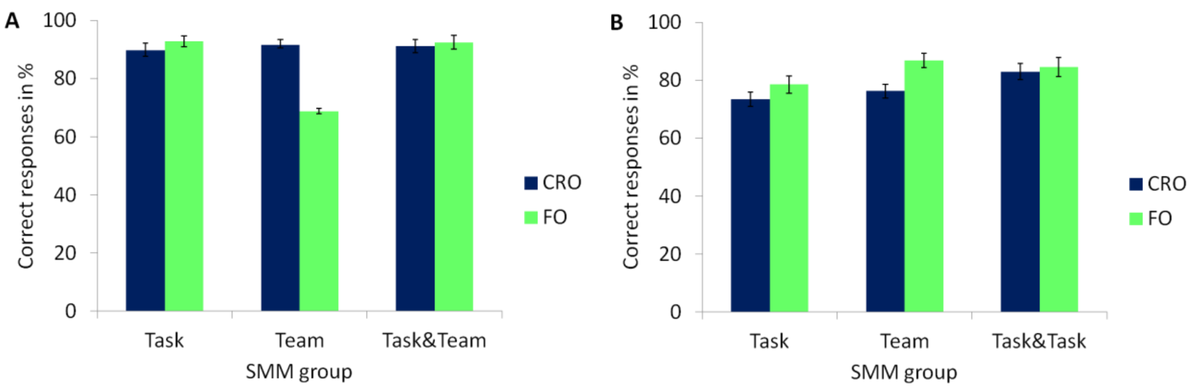

Figure 5. Percentage of correct responses on the knowledge about the task model (A) and the team model (B) separated by SMM group and team role. Error bars indicate the standard error of the mean.

4.2 Team Performance

To evaluate team performance, solution time, correctness of fault diagnosis, number and relevance of tested hypotheses and number of violations of work safety rules were analyzed (see Table 4). Performance measures were analyzed using a one-factorial ANOVA with the between subjects factor SMM group (Task vs. Team vs. Task & Team). Differences between individual groups were assessed using Bonferroni-corrected paired comparisons.

Solution time. There was no significant effect of SMM group on the solution time needed to reach a decision, $F < 1$.

Correctness score. The decision on a hypothesis was given a correctness score according to the following criteria: The correctness was scored 0 if the team chose one of the incorrect hypotheses, 1 if the team chose the correct hypothesis but did not collect sufficient evidence for the decision (i.e. they did not exclude all other hypothesis that could account for the system behavior they observed), and 2 if the team chose the correct hypothesis and did collect sufficient evidence for the decision. The correctness score was averaged across scenarios. There was no significant effect of SMM group on the correctness score, $F(2,29) = 0.65$, $p = .530$, $\eta^2_p = .05$.

Identity and number of hypotheses tested. As it was possible to test multiple hypotheses while only selecting one hypothesis in the plant simulation, the identity and number of hypotheses tested by participants were therefore reconstructed from the communication transcripts. The tested hypotheses were categorized according to their relevance for the alarm in a scenario and analyzed as percentage of relevant hypotheses tested. There was no significant effect of SMM group on either the number of hypotheses tested, $F(2,29) = 1.82$, $p = .181$, $\eta^2_p = .12$, or the percentage of relevant hypotheses tested, $F < 1$.

Violations of work safety rules. The log files of system states were analyzed for system states and combinations of system states that represented violations of work safety rules (i.e., the pressure rising over the threshold of 5.8 bar or combinations of states of pumps and valves that deteriorated system states). The mean number of violations of work safety rules per scenario showed a significant effect of SMM group, $F(2,29) = 5.35$, $p = .011$, $\eta^2_p = .28$. indicating that participants in the Team SMM group violated work safety rules more often than participants in the Task & Team SMM group (1.84 vs. 0.82), $p = .011$.

Table 4

Mean values and standard deviations (in parentheses) of performance data divided by SMM group and performance measure

	Task	Team	Task & Team
Solution time	9min 07s (4min 49s)	10min 21s (5min 30s)	9min 57s (6min 35s)
Correctness score	1.49 (0.80)	1.32 (0.87)	1.56 (0.70)
Sum of tested hypotheses	5.64 (2.64)	5.88 (2.76)	4.48 (2.45)
Percentage of relevant hypotheses	0.80 (0.14)	0.80 (0.19)	0.80 (0.19)
Violations of work safety rules	1.12 (0.67)	1.84 (0.92)	0.82 (0.49)

4.3 Verbal Communication

4.3.1 The Category System

The audio and video data of every team was transcribed and each statement was coded according to a category system. The category system was adapted from one that was devised to analyze communication in dyadic command and control teams (Rasker et al., 2000; Schraagen & Rasker, 2001). Command and control teams communicate to exchange information needed for task execution. As these teams commonly work in fast changing environments, they depend on a continuous flow of information. They need to monitor the environment, interpret and share information and insights with other team members, discuss and implement strategies and monitor the effects of their actions (Rasker et al., 2000). These preconditions and functions of communication shape the basic dimensions of the category system and are also relevant during cooperative fault diagnosis in a process plant. The dimensions proposed by Schraagen and Rasker (2001) are *information exchange, performance monitoring, determining strategies, evaluation, situation knowledge, team knowledge,* and *other*. The most important function of communication is *information exchange* about the environment and the situation. As one aim of the present study was to discern whether participants who have more similar mental models engage in more implicit coordination, the category was further divided into the subcategories *inquiry, reply,* and *voluntary*. The category *voluntary* describes the provision of information without explicit or implicit request, and a higher frequency of statements in this category can be understood to reflect a higher occurrence of implicit coordination. Furthermore, the category *alarm* was added to the system as the communication of the alarm in the current scenario was crucial for an effective fault diagnosis. In addition to exchanging information, participants can also *exchange knowledge* they have acquired during the instruction or the experiment. This category is subdivided into the exchange of *team knowledge* and *task knowledge.* The category *task knowledge* has replaced the originally proposed *situation knowledge* to reflect that learned information for the experiment belongs to either the team or task model. Statements describing what one has done, is currently doing, or is about to do are termed *performance monitoring* and a higher frequency of occurrence also signifies the use of effective communication strategies. These statements have to be distinguished from statements belonging into the category *determining strategies.* The important difference is that *determining strategies* statements are concerned with the suggestion and discussion of the next steps and actions in consultation with the team partner, while *performance monitoring* simply announces what is going to happen next without any affordances for discussion. As the complex system makes it extremely important to plan ahead and test equipment strategically, the category of *prediction* has been added, to reflect the amount of detailed planning participants engage in. Furthermore, statements in this category are marked as either true or false to reflect the accuracy of planning. The category *command action* describes statements that explicitly order the team mate to execute a certain action, and it was added to the system to reflect the direction and hierarchy of communication. *Evaluation* describes statements that interpret and judge the state of the system, the current implications, and the consequences of control actions. As participants could also draw false conclusions from the available information, these statements were also marked as either true or false. The category *acknowledgement* describes the extent to which participants signal that a statement was heard, understood, and accepted, and the category *clarification* reflects how often participants disambiguate or ask for disambiguation of unspecific information. In addition to the classification by content all statements were according to the speaker (CRO vs. FO). For a full description of the category system with definitions and examples of categories and subcategories see Table 5.

Table 5

Communication analysis, overview of categories, definitions, and examples

Category	Subcategory	Definition	Example
Information Exchange	Alarm	A statement that communicates the alarm of a scenario	CRO: "The alarm says that the liquid level of reactor R001 is too high."
	Inquiry	An inquiry for information about the scenario (i.e. liquid level of reactor, state of valve)	CRO: "Can you give me the state of valve 001, please?"
	Reply	A reply to an inquiry with information on the scenario	FO (in reply): "Valve 001 is open."
	Voluntary	A voluntary statements that provides information without explicit inquiry	FO: "The pipe is dripping."
Knowledge Exchange	Team	A statement that transfers information about the learned facts about the team, the team members' roles, and information available to them	CRO: "Only you can see the pressure in a pipe."
	Task	A statement that transfers information about the learned facts about the task, common procedures, common faults and task rules.	FO: "If the level of a container is low, we should not close off its inflow."
Determining Strategies		A statement that expresses intentions to adjust the way the team should engage in the task. Deliberations about alternatives, rationalizations of the strategy adopted so far.	CRO: "Should we test the hypothesis 'valve 004 is closed' first?"

Performance Monitoring	A statement about the tasks team members execute during the scenario. Telling each other what one is doing at the moment, giving advice what to do, and giving feedback about each other's performance.	CRO: "I'm checking the state of the pump now."
Prediction	A statement that predicts a system response to a specific control action under the condition of a test hypothesis. *Marked true or false*	FO: "If the valve is clogged and we close it, the liquid level in the container should still rise."
Command Action	A statement that tells the other team member which action should be executed.	CRO: "Please close valve 003."
Evaluation	A statement that evaluates or judges the current activities of the scenario. Analyses of why things went well or wrong. *Marked true or false*	FO: "I think this means the valve is clogged."
Acknow-ledgement	A statement representing that a message or order was received.	FO: "Okay, I will close valve 003."
Clarification	A statement that asks for or gives clarification.	CRO: "Do you mean P001?" FO: "Yes."
Other	Partial or unclear statements of unclear function or social exchange.	CRO: "I will take the blame if the plant explodes."

Note. All statements by the participants were additionally marked according to the speaker (i.e., CRO or FO).

4.3.2 Communication Frequency and Proportional Shares

To account for huge differences in the duration of scenarios between teams (*M* = 9.80 min, *SD* = 5.67 min, range = 1.28 - 27.71 min) the total number of statements per scenario was divided by the duration of the scenario to attain the communication frequency (i.e. number of statements per minute) Each team partner's communication frequency and their proportional share in the total amount of communication were aggregated across scenarios and compared using two-factorial ANOVAs with the between factors SMM group (Task vs. Team vs. Task & Team) and role (CRO vs. FO) (see *Table 6*). Differences between individual groups were assessed using Bonferroni-corrected paired comparisons.

Communication frequency per partner. The analysis revealed a significant main effect of SMM group, $F(2,59) = 9.00$, $p < .001$, $\eta^2_p = .25$, and role, $F(1,59) = 7.56$, $p = .008$, $\eta^2_p = .12$. This indicates that participants in the Task & Team SMM group had a lower communication frequency than participants in the Task SMM group, $p = .001$, and the Team SMM group, $p = .002$, (5.29 vs. 7.23 and 7.13 statements/min), and CROs had a higher communication frequency than FOs across all SMM groups (7.13 vs. 5.97 statements/min).The interaction effect SMM of group and role, however, was non-significant, F <1. The results of the analysis of the absolute number of statements per scenario did not differ from these results.

Proportional shares of communication. The main effect of SMM group in the analysis could not be reported as the sum of the shares of CROs and FOs within a SMM group always equaled 100 percent for each team. There was a significant main effect of role, $F(1,59) = 23.73$, p < 001, $\eta^2_p = .31$. The interaction effect of SMM group and role was non-significant, $F(2,59) = 1.64$, p =.24, $\eta^2_p = .05$. This indicates that across all SMM groups CROs had a higher proportional share in communication than FOs (54.6 vs. 45.4%).

Table 6

Mean values and standard deviations (in parentheses) of communication frequency and proportional communication shares divided by SMM group and role

	Task		Team		Task & Team	
	CRO	FO	CRO	FO	CRO	FO
Communication frequency	8.14 (1.94)	6.32 (1.03)	7.44 (1.40)	6.83 (1.51)	5.81 (1.77)	4.77 (1.94)
Communication share	55.9 (5.0)	44.1 (5.0)	52.3 (5.3)	47.7 (5.3)	55.6 (10.3)	44.4 (10.3)

4.3.3 Relative Proportions of Communication Categories

There were large differences in the number of statements per scenario and partner (M = 62.33, SD = 38.98, range = 6 - 190). To allow for comparisons between SMM groups and team roles, for each partner the relative proportions of the communication categories in the total amount of communication per scenario was calculated. (see *Figure 6*). These proportions were compared using two-factorial ANOVAs with the between subjects factors SMM group (Task vs. Team vs. Task & Team) and role (CRO vs. FO). Differences between individual groups were assessed using Bonferroni-corrected paired comparisons.

Information Exchange. There were significant main effects of SMM group, $F(2,59) = 6.14$, p =.004, $\eta^2_p = .19$, and role, $F(1,59) = 5.30$, p =.025, $\eta^2_p = .09$, but no interaction effect of SMM group and role, F < 1. This indicates that participants in the Task & Team SMM group used a higher proportion of their statements to exchange information than participants in the Team SMM group (37.0 vs. 28.9%), p = .003, and CROs used a lower proportion of statements for information exchange than FOs (30.9 vs. 35.3%). When the category was split into its subcategories *alarm*, *voluntary, inquiry,* and *reply*, no effect of SMM group or role on the communication of the alarm, $F < 1$, the proportion of voluntary information exchange, all Fs < 1.60, all ps >.212, or the proportion of inquiries for information, all Fs < 2.77, all ps >.071, could be identified. In the category reply there was no main effect of SMM group, $F(2,59) = 1.09$, $p = .342$, $\eta^2_p = .04$, but a significant main effect of role, $F(1,59) = 7.23$, $p = .010$, $\eta^2_p = .12$. This indicates that CROs exchanged less information on request than FOs (10.1 vs. 14.3%). No interaction effect of SMM group and role, $F(2,59) = 0.65$, $p = .524$, $\eta^2_p = .02$, could be identified.

Knowledge Exchange. No effects of SMM group or role in the proportion of knowledge exchange could be identified, all Fs < 2.42, all ps >.099. When the category was split into its subcategories *task knowledge* and *team knowledge,* in the category task knowledge there was no main effect SMM group, $F(2,59) = 3.00$, $p = .059$, $\eta^2_p = .10$, a significant main effect of role, $F(1,59) = 14.08$, $p < .001$, $\eta^2_p = .21$ and no interaction effect of SMM group and role, $F(2,59) = 3.15$, $p = .051$, $\eta^2_p = .10$. This indicates that CROs used a larger proportion of statements to exchange task knowledge than FOs (1.6 vs. 0.7%). No effects of SMM group or role in the proportion of team knowledge exchanged could be identified, all Fs < 3.41, all ps >.070.

Determining Strategies and Prediction. No effects of SMM group or role could be identified in the proportion of statements used for the discussion of strategy, all Fs < 0.53, all ps >.593, and the prediction of system responses, all Fs < 2.74, all ps >.074.

Performance Monitoring. There was no main effect of SMM group, $F(2,59) = 2.31$, $p = .109$, $\eta^2_p = .08$, there was a significant main effect of role, $F(1,59) = 18.29$, $p < .001$, $\eta^2_p = .25$, but no interaction effect of SMM group and role, $F < 1$. This indicates that CROs used a larger proportion of their statements to tell their partner about the tasks they had done or would be doing than FOs. (19.4 vs. 12.9%).

Command action. There was no main effect of SMM group, $F < 1$, there was a significant main effect of role, $F(1,59) = 7.29$, $p = .009$, $\eta^2_p = .12$, but no interaction effect of SMM group and role, $F < 1$. This indicates that CROs used a larger proportion of their statements to order FOs to execute a task than the other way around (1.9 vs. 0.5%).

Evaluation. There was no main effect of SMM group, $F < 1$, but a significant main effect of role, $F(1,59) = 6.99$, $p = .011$, $\eta^2_p = .12$, meaning that CROs used a higher proportion of statements for evaluation and judgment of information, actions, and consequences of activities in the scenario than FOs (13.5 vs. 10.1%). No interaction effect of SMM group and role, $F < 1$, could be identified.

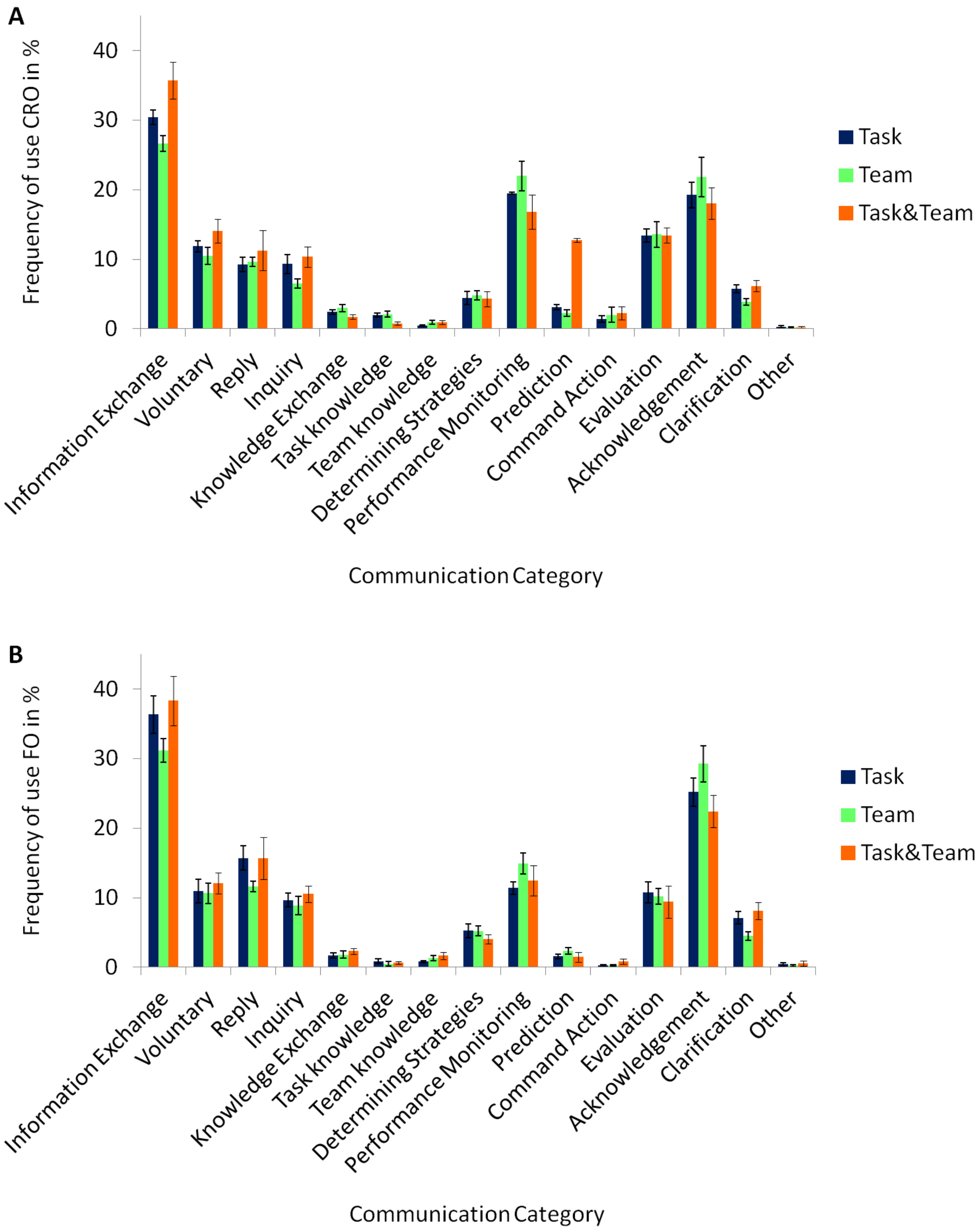

Figure 6.

Mean values of proportion of statements per communication category separated by SMM group for CRO (A) and FO (B). Error bars indicate the standard error of the mean.

Acknowledgement. There was no main effect of SMM group, $F(2,59) = 2.67$, $p = .079$, $\eta^2_p = .09$, but there was a significant main effect of role, $F(1,59) = 9.70$, $p = .003$, $\eta^2_p = .15$, This indicates that CROs used a lower proportion of statements that acknowledged that the contribution of their partner was heard, understood or accepted than FOs (19.7 vs. 25.6%). No interaction effect of SMM group and role, $F < 1$, was found.

Clarification. There was a significant main effect of SMM group, $F(2,59) = 7.25$, $p = .002$, $\eta^2_p = .21$, but no main effect of role, $F(1,59) = 3.86$, $p = .055$, $\eta^2_p = .07$, and no interaction effect of SMM group and role, $F < 1$. This indicates that participants in the Team SMM group used a lower proportion of statements to clarified or ask for clarification of ambiguous information than participants in the Task, $p = .021$, or the Task & Team SMM group, $p = .002$, (4.2 vs. 6.5 and 7.1%).

Other. No effect of SMM group or role in the proportion of statements that did not belong into any of the other categories all could be identified, Fs < 1.21, all ps >.277.

4.3.4 False Statements

All statements of the categories prediction and evaluation were marked as true or false. Statements were only marked as false when they could be identified as undoubtedly incorrect given the current state of the system and the information available to participants, while all other statements were marked as true. False statements within and across categories were analyzed as percentage scores and were aggregated across scenarios. Scores were compared using two-factorial ANOVAs with the between subjects factors SMM group (Task vs. Team vs. Task & Team) and role (CRO vs. FO) (see *Table 7*). Differences between individual groups were assessed using Bonferroni-corrected paired comparisons.

False statements per category. In the category prediction there was a significant main effect of SMM group, $F(2,59) = 4.69$, $p = .013$, $\eta^2_p = .15$, indicating that participants in the Task SMM group had a higher proportion of statements that predicted system reactions falsely than participants in the Task & Team SMM group (13.8 vs. 1.0%), $p = .010$. However, there was no main effect of role, $F(1,59) = 0.65$, $p = .423$, $\eta^2_p = .01$, and no interaction effect of SMM group and role, $F(2,59) = 1.26$, $p = .292$, $\eta^2_p = .05$. Furthermore, there were no significant effects of SMM group or role on the percentage of statements that evaluated system reactions falsely, all Fs < 1.69, all ps >.195.

False statements across categories. No significant effects of SMM group or role on the proportion of false statements across categories could be identified, all Fs < 2.22, all ps >.087.

Table 7

Mean values and standard deviations (in parentheses) of false statements by category and in total divided by SMM group and role

	Task		Team		Task & Team	
	CRO	FO	CRO	FO	CRO	FO
False predictions	14.5 (15.9)	13.0 (25.0)	2.7 (3.8)	13.0 (11.2)	1.3 (4.2)	0.8 (1.7)
False evaluation	11.8 (9.0)	12.5 (8.1)	11.8 (7.5)	11.3 (7.3)	7.5 (9.7)	8.1 (7.3)
False statements	2.0 (1.1)	1.8 (2.1)	2.1 (1.6)	1.7 (1.2)	1.0 (1.3)	1.0 (1.2)

4.4 Questionnaire Data

Questionnaire data about the participants approach to selection and testing of hypotheses and the quality of cooperation was obtained. The answers were analyzed using two-factorial ANOVAs with the between factors SMM group (Task vs. Team vs. Task & Team) and role (CRO vs. FO) (see *Table 8*). Differences between individual groups were assessed using Bonferroni-corrected paired comparisons.

Subjective hierarchy within the team. The analysis of the subjective hierarchy of team members was based on the question 'In which role did you see yourself in the cooperation with your partner?' (response options: employee, inexperienced colleague, equal colleague, experienced colleague, boss, scaled 1-5). There was no main effect of SMM group, $F(2,59) = 0.12$, $p = .892$, $\eta^2_p = .00$, but significant main effect of role, $F(1,59) = 5.62$, $p = .021$, $\eta^2_p = .09$, and no interaction effect of SMM group and role, $F(2,59) = 0.80$, $p = .453$, $\eta^2_p = .03$. This indicates that CROs felt they occupied a hierarchically higher position than FOs (2.30 vs. 1.83).

Utility of information. The analysis of the subjective utility of information for the testing of hypotheses revealed a pattern for the main parameters that characterized the plant (i.e., liquid level, temperature). It was based on the question 'how useful did you find the following information for the testing of hypotheses?' and rated on a scale from 1 (superfluous) to 5 (extremely useful). For the *utility of information about liquid level* there was no significant main effect of SMM group, $F(2,58) = 3.16$, $p = .051$, $\eta^2_p = .11$, or role, $F(1,58) = 0.28$, $p = .447$, $\eta^2_p = .01$, but a significant interaction effect of SMM group and role, $F(2,58) = 6.21$, $p = .004$, $\eta^2_p = .19$. The interaction indicates that CROs of the Team SMM group rated the utility of the information about liquid level lower than CROs of the Task, $p = 0.15$, and the Task & Team SMM group, p <001, (3.60 vs. 4.50 and 4.89) resulting in a significant difference in the subjective utility of the information about liquid level between CRO and FO in the Team SMM group (3.60 vs. 4.60), $p = .002$. For the *utility of information about temperature* a similar pattern with no main effect of SMM group, $F(2,58) = 0.21$, $p = .815$, $\eta^2_p = .01$, or role, $F(1,58) < 0.00$, $p > .99$, $\eta^2_p < .01$, but a significant interaction effect of SMM group and role, $F(2,58) = 4.95$, $p = .011$, $\eta^2_p = .16$ could be identified. This again indicates that CROs of the Team SMM group rated the utility of the information about temperature lower than CROs of the Task & Team SMM group, p = 037, (2.80 vs. 4.00), also resulting in a significant difference in the subjective utility of the information about temperature between CRO and FO in the Team SMM group (2.80 vs. 3.90), $p = .018$.

Table 8

Mean values and standard deviations (in parentheses) of responses in the questionnaire divided by SMM group and role

	Task		Team		Task & Team	
	CRO	FO	CRO	FO	CRO	FO
Hierarchy	2.20 (0.92)	2.00 (0.00)	2.40 (0.92)	1.60 (0.70)	2.30 (0.82)	1.90 (0.74)
Utility of information						
Liquid level	4.50 (0.53)	4.40 (0.70)	3.60 (0.97)	4.60 (0.70)	4.89 (0.33)	4.40 (0.70)
Temperature	3.50 (1.08)	3.30 (1.16)	2.80 (0.92)	3.90 (1.10)	4.00 (0.87)	3.10 (0.88)

5 Discussion

Fault diagnosis in a complex system requires operators to coordinate and communicate well. The present study investigated the influence of shared team and task mental models on remote communication and team performance during a fault diagnosis in a simulated process plant.

In line with *Hypothesis 1* the analyses showed different results for the three shared mental model groups (SMM groups). This suggests the mental models had some unique influences on the content and amount of communication exchanged between CRO and FO.

It was hypothesized that the higher similarity of mental models in the Task & Team SMM group would allow the teams to use more effective communication strategies (*Hypothesis 2*) and the increase of effectiveness would result in a benefit for performance (*Hypothesis 3*). A closer inspection of the differences between the SMM groups showed that the Task & Team SMM group did not use the communication strategies more often that had been deemed effective (i.e., voluntary provision of information and performance monitoring). However, communication in the Task & Team SMM group was different from communication in the other SMM groups, as participants in the Task & Team SMM group communicated less and used a higher percentage of statements for information exchange. As measures of team performance, the average solution time, the correctness score and the number and relevance of tested hypotheses did not differ between SMM groups. However, the Task & Team SMM group violated work safety rules less often than the Team SMM group. This suggests that the Task & Team SMM group might have experienced fewer difficulties during the diagnosis and thus had more capacity to comply with the rules. The reduction of communication in the Task & Team SMM group did not affect team performance negatively. In fact, the reverse seems possible, as the stricter adherence to the work safety rules even suggests a slight benefit. Consequently, the question arises whether the higher similarity of mental models may still have resulted in more effective communication, but the strategies responsible for the benefits differed from those identified by the literature (Brindley & Reynolds, 2011; Serfaty et al., 1993).

Hypothesis 4 assumed that a shared task model would enable both team partners to participate equally in the fault diagnosis and that the absence of a shared task model would lead to a more dominant role of the CRO. The results show a distinct hierarchy between participants with the CRO taking the hierarchically higher position in both, the subjective rating of participants as well as in the objective measures. CROs claimed a higher share of the total amount of communication and a higher proportion of statements in the more directive or active categories of communication (i.e., command action, task knowledge exchange, evaluation, and performance monitoring) while FOs used a higher proportion of passive statements (i.e., acknowledgement and information exchange in reply). Contrary to the hypothesis, the difference in status did not become more distinct when participants did not share the task model, but remained stable across all SMM groups instead.

A shared team model was assumed to allow for a better anticipation of the partner's information needs, and thus facilitate communication by enabling team partner to voluntarily share relevant information (*Hypothesis 5*). As this voluntary sharing would be more difficult when the team model was not shared, it was assumed that the Task SMM group would need to communicate more (*Hypothesis 6*). However, in the present study a shared team model did not result in a higher frequency of voluntarily provided information. And there was no difference between the Task and the Team SMM group in the amount of communication.

In short, the present study showed that sharing different mental models during fault diagnosis causes differences in communication and performance. These differences may even have been beneficial when both the task and the team mental model were shared. However, the presumed effects caused specifically by the presence of the task or the team mental model could not be identified here. Surprisingly, the analyses also revealed that the role of CRO was perceived as hierarchically higher than the role of FO across all SMM groups and that there was high interpersonal variance in all measures, especially solution time and amount of communication. These findings raise several questions.

5.1 Why was the Combination of Task and Team Model more Successful than the Individual Models?

The main finding of the present study is that only the combination of a shared team and task model caused a distinct effect on communication and a benefit for team performance when contrasted with the individual models, while the individual mental models did not lead to a distinct effects on communication and created no benefit for performance when contrasted with each other. These findings are echoed by previous research results that identified a beneficial effect of the interaction of a shared team and task model with slight to no distinct effects when only one of the models was shared (Mathieu et al., 2010; Smith-Jentsch et al., 2005). Yet, other studies were able to identify distinct effects of the task and the team model (e.g., Lim & Klein, 2006; Mathieu et al., 2005). Several explanations might account for this combination of effects.

First, it is necessary to consider that the manipulation of shared mental models might not have worked optimally. The results from the knowledge test indicate that the manipulation of the task model seems to have worked well, as the Team SMM group had lower task knowledge than the other groups, as was intended by the instruction seminar.

Moreover, only in the Team SMM group did CRO and FO differ in knowledge, with the CRO having more task knowledge than the FO. This difference mirrors the fact that in this group only the CRO received information on the task model during the instruction seminar. However, the manipulation of the team model did not succeed equally well. While the Task & Team SMM group had more knowledge about the team model than the Task SMM group, no such difference was found between the Task and the Team SMM group. Furthermore, participants generally had a higher amount of knowledge about the task model than about the team model, leading to the conclusion that information on the team model was not retained as well as information on the task model in this experiment. One reasons for this may be that the instruction about the team model did not seem as important to the participants as the instruction about the task model. The utility of knowledge about the distribution of information between team members was probably not immediately obvious to participants, as the information was presented rather early in the instruction seminar, when participants did not yet understand the functional principles of the simulated plant completely. Additionally, the early presentation of information about the team model may have made the information about the team model harder to learn than the information on the task model, which was presented later on, as the retention of information is better, when the new information can be linked to already existing knowledge structures (Reigeluth, Merrill, Wilson, & Spiller, 1980). In combination, these circumstances may have prevented participants from fully developing a shared team mental model and consequently its unique effects did not manifest in the experiment. However, this reasoning does not explain why the successful manipulation of the task model did not manifest either.

A point that also needs to be considered here is that while the manipulation in general may have worked, the present study only manipulated the accuracy of mental models and assumed that it would also reflect similarity of mental models. This is a common assumption in the literature, as one method of increasing the similarity of mental models is increasing their accuracy as compared to the mental model of an expert (Cannon-Bowers et al., 1993; Edwards et al., 2006). There are, however, diverging findings with regards to the effects of accuracy and similarity of mental models. While the accuracy of mental models generally seems to be a stronger predictor of team performance than their similarity (Edwards et al., 2006; Lim & Klein, 2006; Marks, Zaccaro, & Mathieu, 2000; Sander et al., 2015), some studies also conclude that mental model similarity creates additional benefits (Lim & Klein, 2006; Marks et al., 2000). These findings demonstrate that accuracy and similarity of mental models are closely related yet distinct constructs and should not be assumed to mirror each other perfectly. Although mental models may not have been as similar within groups as was intended, this assumption does not explain why there were changes in communication and benefits in performance when both mental models were shared.

Another reason why the individual models did not affect communication and performance is the influence of the specific task and the contextual constrains. As discussed before, when looking at previous research on shared mental models it is obvious that the importance of the different types of mental models differs between studies (e.g., Lim & Klein, 2006; Mathieu et al., 2000; Mathieu et al., 2010). The diverging findings could be explained if characteristics of the task and the context determined how the types of mental models affect the outcomes. All types of mental models can be useful, depending on the current problem (Rouse & Morris, 1986), and it has been suggested that the importance of mental model types varies between laboratory and field settings (Lim & Klein, 2006). In the present study the experiment was deliberately designed to make both the task and the team model essential for high task performance. This means that the task model enabled participants to successfully manipulate the system (e.g., turn off a valve in accordance with work safety rules) and interpret its reactions (e.g., interpret the amount of change in the liquid level). In addition, the team model enabled them to collect and exchange information efficiently (e.g., the CRO could ask the FO for the pipe pressure). Consequently, it seems only logical that both models needed to be shared in order to result in optimal performance. However, in any case all necessary information was available to the team and participants could simply make up for their lack of a shared mental model by exchanging information.

The unrestricted communication during the experiment is another explanation for the combined effect of team and task model. Several authors previously suggested that the commonly employed input-process-output framework is too simplistic, as its linear process does not reflect the diverse nature of the influence of shared mental models (Ilgen, Hollenbeck, Johnson, & Jundt, 2005; Marks, Mathieu, & Zaccaro, 2001; Uitdewilligen et al., 2010). Van den Bossche et al. (2011) found that team learning behaviors, like constructive conflict, influence the development of shared mental models and Stout et al. (1999) showed that effective planning increases the similarity of mental models. It is therefore reasonable to assume that the mental models are not set in stone once they have been developed, but are subject to continuous change and can evolve to become more similar over time. Rasker et al. (2000) concluded that specific communication types, like performance monitoring, are beneficial in the construction of similar mental models. Thus, teams sharing only one of the relevant mental models could have used the unrestricted communication during the experiment to actively construct more similar mental models. In this process, the manipulation of shared knowledge during the instruction seminar would have had the effect of a common knowledge basis for the construction of shared mental models, giving participants either a more or less similar mental model as a starting point. To work together effectively, participants needed to expand on this basis and develop more similar mental models. When they shared only one of the mental model types, they needed to invest more effort to reach this goal than when they already shared both types. Empirically, this effort manifested in the higher communication frequency found in the groups that only shared either the team or the task mental model. While this explanation seems logical and is in accordance with the empirical findings, in retrospect it is impossible to prove that the mental models of participants became more similar during the experiment. Further studies should investigate this question by adding an additional measurement of task and team knowledge at the end of the experiment, to assess whether the similarity of mental models changed due to the interaction.

It is not enough, however, to only focus on the finding that overall there seemed to be no distinct pattern of effects when the individual types of shared mental models were contrasted with each other, but one must also note that the results on the Task and the Team SMM group occasionally differed from each other when they were compared to the Task & Team SMM group. Effects which were only visible as a tendency when comparing the Task SMM group with the Task & Team SMM group were often significant effects when comparing the Task & Team SMM group with the Team SMM group. This suggests that although there were no significant differences between the Task and the Team SMM group, the models might still have different effects on communication and coordination in a team. Participants in the Team SMM group seemed to experience more problems during the fault diagnosis, resulting in the higher amount of violations of work safety rules and a lower percentage of statements concerned with information exchange when compared to the Task & Team SMM group, while the Task SMM group did not differ significantly from the Task & Team SMM group in these measures. Furthermore, the analysis of questionnaire data on the utility of parameter information revealed a curious pattern for this group, as CROs in the Team SMM Group rated the utility of information about liquid level and temperature lower than CROs in other groups and lower than FOs in the same group. This might be an indication that CROs who shared only the team model with their partners and were therefore the only team member with knowledge about the task, might have suffered from excessive cognitive demands. Additional strain was put on the CRO as he or she either needed to diagnose the fault mostly on his or her own while directing the FO's actions or tediously explain the functional principles of the plant. Being out of their depth and left to their own resources, CROs perceived the information available to them as less useful. In sum, it still seems reasonable to assume that a shared task model might be able to facilitate the collaboration of operators. However, the design of the study might not have been optimal to bring out the whole extent of the differences, as is indicated by a large amount of interpersonal variance.

5.2 What Caused the High Amount of Interpersonal Variance?

When looking at the results of the present study, a high amount of interpersonal variance could be observed. This was especially striking when looking at the range of the solution time and the number of statements per partner per scenario. Regrettably, the variance of obtained data is not commonly stated in the literature, but reference to the transformation of data in order to stabilize variance (Sauer et al., 2006) gives an indication that the observation might not be uncommon. As this variance may have obscured some influences of shared mental models, its sources need to be considered in order to reduce it in further studies.

As discussed before, the accuracy of mental models cannot be considered to reflect the similarity of mental models perfectly. The fact that a great deal of variance in the accuracy of mental models within the individual SMM groups was observed stands in line with this reasoning. It has been shown that people with higher general mental ability can construct more accurate mental models (Day, Arthur, & Gettman, 2001), and teams with higher ability can construct mental models that are more accurate as well as more similar (Edwards et al., 2006). Furthermore, it has been suggested that individual differences that are known to influence the learning process (e.g., like prior experience, learning styles) also influence the construction of mental models (Jih & Reeves, 1992; Kraiger & Wenzel, 1997). In line with these findings, participants showed varying degrees of accuracy in their understanding of the functional principles of the plant. A common misconception was the assumption that the stable liquid level of a vessel was an indication for a fault in the vicinity, when in fact it was an indication that the system was working correctly, as the rules of throughput clearly stated that all vessels should be subject to mass balance. Some participants also over-simplified the functional principles, as they assumed that the manipulation of the state of a functional pump or a valve should simply change the liquid level of the adjourning vessels. What they did not consider was the fact that the difference between functional and defective equipment lay in the amount of change caused by the manipulation. These misconceptions manifested in false statements during the experiment. Surprisingly, a higher percentage of false predictions could be observed in the Task in contrast to the Task & Team SMM group. Originally, it was assumed that the task model enabled participants to predict system responses better and not worse, as it included the functional principles of the plant. Thus, it seems reasonable to assume that some participants constructed less accurate task and team models due to individual differences. Consequently mental models might have differed a great deal in accuracy and therefore similarity between individual teams and these differences might have led to high variance in communication and performance.

Another source of variance is that not all participants took a systematic approach to the fault diagnosis task. In line with previous findings on fault diagnosis, participants did not use all available information on the problem (Rasmussen & Jensen, 1974). In many cases only the alarm message was considered and no additional information on the problem gather (i.e., parameter values and their changes in the concerned vessel), even when this information could have facilitated the fault diagnosis. Furthermore, many participants did not consider the relevance of the presented set of hypotheses, and preferred to test the list from the top down or select the ones that were easiest to test (Bereiter & Miller, 1989). This observation leads to the assumption that being confronted with a set of hypotheses participants may have simply skipped the phases of representing information and generating hypotheses and immediately started the phase of testing hypotheses (Abele, 2017). Additionally, participants often seemed to lack clear procedures for testing hypotheses and tended to abandon the test of a hypotheses midway without any conclusive test results. This lack of structure could have led participants to fixate on the investigation of irrelevant hypotheses (Bereiter & Miller, 1989) or repeat the test the same hypothesis multiple times (Patrick et al., 1999). Previous research has shown that the use of rules and procedures that specify the operation of a system can improve performance in a fault diagnosis task (Embrey, 2009; Morris & Rouse, 1985; Raaijmakers & Verduyn, 1996; Usher & Kaber, 2000). In order to reduce interpersonal variance caused by variance in the strategic approach, participants could receive a mandatory set of formal procedures for the fault diagnosis in general as well as individual hypotheses tests. Thus, participants could be instructed to actively consider all available information about the problem and select the hypotheses that need to be tested on the grounds logical arguments. By following the procedures for the individual tests, participants would always receive conclusive results. In combination, this would mean that participants experiencing difficulties during the selection and testing of hypotheses would have to execute fewer tests and be able to interpret the tests results better. In consequence, interpersonal variation in solution time and communication would be reduced.

A third source of variance is repeated misunderstandings between team members due to unclear communication. Communication problems, like misunderstandings, frequently contribute to accidents (Bullemer & Laberge, 2010; Hofinger, 2012) and were also a detrimental factor in the Texas Refinery Explosion (U.S. Chemical Safety and Hazard Investigation Board, 2007). It has been argued that communication occurs in a continuum between understanding und misunderstanding (Verdonik, 2010). This argument becomes clear when considering the example of a CRO who requests information about the state of a valve. A clear request by a CRO would be "Check valve 002. It is located between reactor 001 and Pump 001." This statement identifies the valve by its label and clearly states its location and thus enables the FO to quickly identify the correct valve. A moderately ambiguous request would be "Check the valve after the reactor". In this statement the valve is not labeled and therefore harder to identify. It would also be possible that there is more than one valve that regulates the outflow of the reactor, thus the FO would have to either clarify which valve should be checked, infer the valve that needs to be checked based on the context, or check all valves that fit the criteria. A very unclear request would be "Check the valve on the right side of the reactor." This statement demonstrates a clear lack of understanding of the FO's worksite. A valve located on the right side of the reactor in a schematic plant overview could be any of a number of different valves in the field and the request could prompt a prolonged search for the correct valve or even lead the FO to check an incorrect valve. In sum, this means that the team needs to communicate more when participants are prone to unclear communication. Accumulated over the duration of a scenario, the disambiguation of imprecise statements could produce large interpersonal variance. A possible method of reducing this variance would be the introduction of a communication protocol that regulates the manner of communication and thus, helps participants to exchange information in a way that reduces the risk of misunderstandings. Communication protocols have been employed across domains and are generally found to make communication clearer and more efficient (Howard, 2008; Kim, Park, Jung, Kim, & Kim, 2010; Wadhera et al., 2010). To ensure the information exchange is clear, participants could be required to follow a protocol which specifies the wording of statements. One example for such a rule would be that plant elements must always be named by their code (e.g., V001 for valve number one) to avoid ambiguous or unclear statements. By avoiding time loss due to misunderstandings interpersonal variance could be reduced.

In sum, the high amount of interpersonal variance seems to have been caused by interpersonal differences in ability, strategic approach to problem solving, and clarity of communication style. By providing participants with a structured procedure for fault diagnosis and a communication protocol to support clear communication, interpersonal variance could be reduced and the full extent of the effect of task and team mental model uncovered.

5.3 Which Communication Strategies are Helpful?

Another important question raised by the results is which communication strategies are really effective and relevant in this context. Contrary to the proposed hypotheses, better performance was not accompanied by a higher rate of the effective communication strategies identified in the literature (i.e., voluntarily provided information and performance monitoring) (Brindley & Reynolds, 2011; Entin & Serfaty, 1999; Hofinger, 2012; Mathieu et al., 2000; Oser et al., 1991; Serfaty et al., 1993). However, the group with the most similar mental models communicated less. Previous research has found that the similarity of mental models can either decrease the amount of communication within a team (Espevik et al., 2011) or increase it (Sauer et al., 2006). It is therefore important to consider not only the amount but also the quality of communication between team members. In the present study the Task & Team SMM group communicated less and at the same time performed slightly better than the other groups. This could indicate that communication was more effective when mental models were more similar. The fact that participants did not use the seemingly effective communication strategies more frequently when they had more similar mental models can be explain in two different ways.

Implicit coordination trough the voluntary provision of information has often been shown to have beneficial effects on performance and it can be argued that it is indeed an efficient communication strategy (Entin & Serfaty, 1999; Espevik et al., 2011; Mathieu et al., 2000). However, the way the category 'voluntary provision of information' was defined in the present study may not be an optimal indicator for the occurrence of implicit coordination. Relevance and timing of the presented information are vitally important to the concept of implicit coordination, as they decide whether or not a statement is helpful (Espevik et al., 2011). Providing the team partner with irrelevant or misleading information resulted in a distraction from the relevant information, while providing relevant information with a delay could not prevent long fruitless discussions or the extensive consideration of a false search path. During the categorization process, all information that was provided without an implicit or explicit request was categorized as 'voluntary' and therefore presumed to reflect the use of implicit coordination. Consequently, the result that shared team model did not increase the use of implicit coordination may partly be due to the fact that the category 'voluntary provision of information' does not reflect whether or not the information provided supported or hindered an effective fault diagnosis and thus cannot be assumed to signify the occurrence of implicit coordination.

Conclusions about the effectiveness of communication in the present study have generally been limited by the category system. It was based on the basic function of communication in dyadic command and control teams (Rasker et al., 2000; Schraagen

& Rasker, 2001). Consequently, statements were categorized by function and not by content. However, this causes a loss of some qualities that are necessary for the complete appraisal of the effectiveness of communication. When for example the CRO inquires for information about the state of a valve and the FO tells him that it *should* be open, this exchange will be labeled as 'CRO information exchange inquiry' and 'FO information exchange reply' but the categorization does not reflect that the information the CRO received was not the state information requested, but instead the set point of the valve. Moreover, distinct patterns of communication are also not reflected by the category system, but instead each statement is treated as a specific occurrence of a category and is analyzed separately from the statements immediately before and after. It has been argued that data from the analysis of frequencies is often ambiguous (Bowers, Jentsch, Salas, & Braun, 1998). When communication is analyzed in frequencies, there is no way of identifying instances when important statements were ignored by the partner or when an inquiry for information went unanswered. Thus, analyzing communication patterns can reveal additional insights (Bowers et al., 1998). Another restriction of the category system used is that it relies on surface criteria of speech to distinguish the different types of statements. This means that during the categorization process the wording of a statement was sometimes more important than its meaning in the context of the conversation. For example, a statement like 'next we will look at the pipe pressure between V002 und R001' will be categorized 'performance monitoring' as it states the next action that is going to happen without giving room for debate. However, when the statement is made by the CRO, who cannot read pipe pressure, it carries an implicit command for the FO to read the pipe pressure at the required location and relate the information to the CRO. Only the categorization obvious in the wording of the statement will be transferred by through the categorization process, while the implicit content will be lost. Further research effort needs to be invested to refine and adapt the category system to reflect more qualities of communication in order to assess its effectiveness. Special emphasis should be put on features that enable a rater to clearly identify misunderstandings and detrimental conversation patterns, as these frequently contribute to accidents (Bullemer & Laberge, 2010; Hofinger, 2012).

However, it should also be considered that the finding that effective communication strategies were not used more often when participants had more similar mental models could be due to the fact that implicit coordination and performance monitoring are not the most relevant strategies for effective communication in complex systems. The slight benefit for performance in the Task & Team SMM group was accompanied by a lower communication frequency and a higher percentage of statements used for information exchange. This suggests that 'relying on information exchange' may be an effective communication strategy. Decision quality in groups has been found to increase as a function of group size in small teams (i.e., 2-6 team members) (e.g., Bray, Kerr, & Atkin, 1978; Laughlin & Johnson, 1966). Thus, the quality of a common interpretation of information should be higher than the quality of individual interpretations. The transfer of the information itself enables the teammate to form an individual interpretation, and then decide on the correct interpretation in collaboration with the partner. The use of the strategy 'relying on information exchange' becomes clear when considering an example. A CRO announces that the liquid level of a reactor is too high. Following this declaration, the team will then discuss possible reasons and implications of these conditions. However, this discussion is superfluous if the CRO simply confused the lines on the list of set points and read off the limits for temperature instead of liquid level and the liquid level is in fact within the allotted range. The loss of time and energy could have been prevented if instead the CRO had exchanged all relevant information with the FO and announced the precise liquid level of the reactor. In this case, both team members could have compared the level to the list of set points and the FO would have identified the CRO's mistake. The exchange of the information instead or in addition to the evaluation avoids misunderstandings and misevaluation of the state of the system and therefore 'relying on information exchange' seems to be an effective communication strategy during fault diagnosis. Further research on effective communication strategies should incorporate this notion into experimental design in order to evaluate the use of the strategy.

5.4 Can the Results be Generalized?

Finally, it has to be considered whether the results from the present study are representative for fault diagnosis in process industries or perhaps even fault diagnosis in general. Although the plant simulation used in the present study simplified the functional principles of a process plant, the simplification was necessary to make it understandable for naive participants. However, the simulated plant still displayed all characteristics of a complex system (Brehmer & Dörner, 1993; Funke, 2010; Sterman, 1994). Participants were confronted with a system in which all parameters interacted with each other, (i.e., interactivity of elements), plant parameters changed in real time (i.e., dynamic effects), the full extent of changes in the plant was not immediately obvious to the individual operators (i.e., non-transparency), the fault diagnosis was complicated by the need to follow work safety rules (i.e., multiple goals), and two operators needed to coordinate their actions and information in order to diagnose the fault correctly (i.e., social complexity). Furthermore, a distinct hierarchy between operators in subjective ratings as well as objective measures could be observed across all types and combinations of mental models. Hierarchy between operators was not directly discussed or manipulated at any point during the experiment and thus, seems to be an inherent characteristic of the task itself. The observed hierarchy mirrors the conditions commonly found in the process industries, where FOs may be seen as an additional information source for CROs (Mumaw et al., 2000) and the coordination of FOs' actions is one of the CRO's responsibilities (Bullemer et al., 1997). In sum, this suggests that the plant simulation of the present study is a reasonable representation of a process plant. Thus, it is possible to generalize the insights of the present study to collaborative fault diagnosis in process industries.

In contrast, the application of results to fault diagnosis in other domains seems limited, as the present study focused on the collaborative fault diagnosis between team partners of hierarchically different status, who had access to different information. However, participants seemed to employ strategies and experience difficulties that mirror the general findings on fault diagnosis. They often failed to consider additional information about the problem (Rasmussen & Jensen, 1974), preferred to test the easiest hypotheses first (Bereiter & Miller, 1989; Konradt, 1995), fixated on certain hypotheses (Bereiter & Miller, 1989), reproduced the hypotheses tests multiple times (Patrick et al., 1999), and preferred to confirm their hypotheses instead of disconfirming them (Morrison & Upton, 1994). In combination, these similarities suggest that the findings of the present study might also be applicable to fault diagnosis in other domains, but further research should validate this assumption.

5.5 Applications in Research and Industry

For further research, the results of the present study indicates that while more similar mental models can be beneficial for performance and communication, more effort needs to be spent to investigate the specific nature and prerequisites that make these benefits possible. A special emphasis should be put on the development of a new categorization system for communication which allows researchers to better understand characteristics that make communication and coordination more effective. Additionally, it would be useful to consider the question which characteristics of the tasks determine the importance of shared mental models for good team performance and whether operators may be able to compensate for the lack of shared mental models by relying more heavily on communication strategies that facilitate the construction of shared mental models (Rasker et al., 2000; Stout et al., 1999; van den Bossche et al., 2011).

For the process industries, the most important contribution of the present study is the indication that training all operators on the complete team task instead of focusing training on their own responsibilities is worthwhile. The consequences of accidents in process industries are high and thus, work safety is an immensely important topic. A reduction of violations of work safety rules could ultimately be the factor that makes the difference between a catastrophic accident and a near miss. Consequently, the prospect of reducing violations of work safety rules by supplying operators with training that gives them enough knowledge to develop shared mental seems to be worth the effort. Furthermore, companies should be aware of the fact that hierarchical differences within a team can bias communication. Therefore, companies should emphasize an open working atmosphere which encourages all employees to speak up regardless of rank.

6 References

Abele, S. (2017). Diagnostic Problem-Solving Process in Professional Contexts: Theory and empirical investigation in the context of car mechatronics using computer-generated log-files. *Vocations and Learning*, *28*(4), 167. https://doi.org/10.1007/s12186-017-9183-x

Arocha, J. F., Patel, V. L., & Patel, Y. C. (1993). Hypothesis generation and the coordination of theory and evidence in novice diagnostic reasoning. *Medical decision making : an international journal of the Society for Medical Decision Making*, *13*(3), 198–211. https://doi.org/10.1177/0272989X9301300305

Bereiter, S. R., & Miller, S. M. (1989). A field-based study of troubleshooting in computer-controlled manufacturing systems. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(2), 205–219. https://doi.org/10.1109/21.31027

Bowers, C. A., Jentsch, F., Salas, E., & Braun, C. C. (1998). Analyzing communication sequences for team training needs assessment. *Human Factors*, *40*(4), 672–679. https://doi.org/10.1518/001872098779649265

Bray, R. M., Kerr, N. L., & Atkin, R. S. (1978). Effects of group size, problem difficulty, and sex on group performance and member reactions. *Journal of Personality and Social Psychology*, *36*(11), 1224–1240. https://doi.org/10.1037/0022-3514.36.11.1224

Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, *9*(2-3), 171–184. https://doi.org/10.1016/0747-5632(93)90005-D

Brindley, P. G., & Reynolds, S. F. (2011). Improving verbal communication in critical care medicine. *Journal of critical care*, *26*(2), 155–159. https://doi.org/10.1016/j.jcrc.2011.03.004

Bullemer, P., & Laberge, J. C. (2010). Common operations failure modes in the process industries. *Journal of Loss Prevention in the Process Industries*, *23*(6), 928–935. https://doi.org/10.1016/j.jlp.2010.05.008

Bullemer, P. T., Cochran, T., Harp, S., & Miller, C. (1997). Managing abnormal situations II: Collaborative decision support for operations personnel. *ASM Consortium*.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), *Individual and group decision making: Current issues* (pp. 221–246). Hillsdale, NJ: Erlbaum.

Cooke, N. J., Kiekel, P. A., Salas, E., Stout, R., Bowers, C., & Cannon-Bowers, J. (2003). Measuring team knowledge: A window to the cognitive underpinnings of team performance. *Group Dynamics: Theory, Research, and Practice*, *7*(3), 179–199. https://doi.org/10.1037/1089-2699.7.3.179

Day, E. A., Arthur, W., & Gettman, D. (2001). Knowledge structures and the acquisition of a complex skill. *Journal of Applied Psychology*, *86*(5), 1022–1033. https://doi.org/10.1037/0021-9010.86.5.1022

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, *2*(3), 292–307. https://doi.org/10.1037/1082-989X.2.3.292

Edwards, B. D., Day, E. A., Arthur, W., & Bell, S. T. (2006). Relationships among team ability composition, team mental models, and team performance. *Journal of Applied Psychology*, *91*(3), 727–736. https://doi.org/10.1037/0021-9010.91.3.727

Embrey, D. (Ed.) 2009. *A human factors approach to managing compentency in handling process control disturbances*. *Symposium Series: Vol. 155*.

Entin, E. E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors*, *41*(2), 312–325. https://doi.org/10.1518/001872099779591196

Espevik, R., Johnsen, B. H., & Eid, J. (2011). Communication and performance in co-located and distributed teams: An issue of shared mental models of team members? *Military Psychology*, *23*(6), 616–638. https://doi.org/10.1080/08995605.2011.616792

Espevik, R., Johnsen, B. H., Eid, J., & Thayer, J. F. (2006). Shared mental models and operational effectiveness: Effects on performance and team processes in submarine attack teams. *Military Psychology*, *18*, S23-S36. https://doi.org/10.1207/s15327876mp1803s_3

Funke, J. (1985). Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle. *Zeitschrift für Psychologie*. (193), 443–465. https://doi.org/10.11588/heidok.00008140

Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive processing*, *11*(2), 133–142. https://doi.org/10.1007/s10339-009-0345-0

Hofinger, G. (2012). Kommunikation. In P. Badke-Schaub, G. Hofinger, & K. Lauche (Eds.), *Human factors* (pp. 141–162). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-19886-1_8

Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. Boca Raton, FL: Taylor & Francis. Retrieved from http://ebookcentral.proquest.com/lib/subhh/detail.action?docID=263746

Howard, J. W. (2008). Tower, am I cleared to land? Problematic Communication in Aviation Discourse. *Human Communication Research*, *34*(3), 370–391. https://doi.org/10.1111/j.1468-2958.2008.00325.x

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual review of psychology*, *56*, 517–543. https://doi.org/10.1146/annurev.psych.56.091103.070250

Jih, H. J., & Reeves, T. C. (1992). Mental models: A research focus for interactive learning systems. *Educational Technology Research and Development*, *40*(3), 39–53. https://doi.org/10.1007/BF02296841

Kellermanns, F. W., Floyd, S. W., Pearson, A. W., & Spencer, B. (2008). The contingent effect of constructive confrontation on the relationship between shared mental models and decision quality. *Journal of Organizational Behavior*, *29*(1), 119–137. https://doi.org/10.1002/job.497

Kim, M. C., Park, J., Jung, W., Kim, H., & Kim, Y. J. (2010). Development of a standard communication protocol for an emergency situation management in nuclear power plants. *Annals of Nuclear Energy*, *37*(6), 888–893. https://doi.org/10.1016/j.anucene.2010.01.003

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of Management*, *20*(2), 403–437. https://doi.org/10.1177/014920639402000206

Kluge, A. (2014). Controlling complex technical systems: The control room operator's tasks in process industries. In A. Kluge (Ed.), *The acquisition of knowledge and skills for taskwork and teamwork to control complex technical systems* (pp. 11–47). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-5049-4_2

Kluge, A., Nazir, S., & Manca, D. (2014). Advanced applications in process control and training needs of field and control room operators. *IIE Transactions on Occupational*

Ergonomics and Human Factors, *2*(3-4), 121–136. https://doi.org/10.1080/21577323.2014.920437

Konradt, U. (1995). Strategies of failure diagnosis in computer-controlled manufacturing systems: Empirical analysis and implications for the design of adaptive decision support systems. *International Journal of Human-Computer Studies*, *43*(4), 503–521. https://doi.org/10.1006/ijhc.1995.1057

Kraiger, K., & Wenzel, L. H. (1997). Conceptual development and empirical evaluation of measures of shared mental models als indicators of team effectiveness. In M. Brannick, E. Salas, & C. T.Prince (Eds.), *Series in applied psychology. Team performance assessment and measurement: Theory, methods, and applications* (pp. 63–84). New York, Mahwah, N.J: Lawrence Erlbaum Associates.

Langan-Fox, J., Anglim, J., & Wilson, J. R. (2004). Mental models, team mental models, and performance: Process, development, and future directions. *Human Factors and Ergonomics in Manufacturing*, *14*(4), 333–352. https://doi.org/10.1002/hfm.20004

Lau, N., Jamieson, G. A., & Skraaning, G., Jr. (2012). Situation awareness in process control: A fresh look. In American Nuclear Society (Ed.), *Proceedings of the 8th American Nuclear Society international topical meeting on nuclear plant instrumentation & control and human-machine interface Technologies (NPIC & HMIT).*

Laughlin, P. R., & Johnson, H. H. (1966). Group and individual performance on a complementary task as a function of initial ability level. *Journal of Experimental Social Psychology*, *2*(4), 407–414. https://doi.org/10.1016/0022-1031(66)90032-1

Lim, B.-C., & Klein, K. J. (2006). Team mental models and team performance: A field study of the effects of team mental model similarity and accuracy. *Journal of Organizational Behavior*, *27*(4), 403–418. https://doi.org/10.1002/job.387

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, *26*(3), 356–376. https://doi.org/10.5465/AMR.2001.4845785

Marks, M. A., Sabella, M. J., Burke, C. S., & Zaccaro, S. J. (2002). The impact of cross-training on team effectiveness. *Journal of Applied Psychology*, *87*(1), 3–13. https://doi.org/10.1037//0021-9010.87.1.3

Marks, M. A., Zaccaro, S. J., & Mathieu, J. E. (2000). Performance implications of leader briefings and team-interaction training for team adaptation to novel environments. *Journal of Applied Psychology*, *85*(6), 971–986. https://doi.org/10.1037//0021-9010.85.6.971

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Cannon-Bowers, J. A., & Salas, E. (2005). Scaling the quality of teammates' mental models: Equifinality and normative comparisons. *Journal of Organizational Behavior*, *26*(1), 37–56. https://doi.org/10.1002/job.296

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology*, *85*(2), 273–283. https://doi.org/10.1037//0021-9010.85.2.273

Mathieu, J. E., Rapp, T. L., Maynard, M. T., & Mangos, P. M. (2010). Interactive effects of team and task shared mental models as related to air traffic controllers' collective efficacy and effectiveness. *Human Performance*, *23*(1), 22–40. https://doi.org/10.1080/08959280903400150

Mayer, R. E. (2008). *Learning and instruction* (2. ed.). Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

Mehle, T. (1982). Hypothesis generation in an automobile malfunction inference task. *Acta Psychologica*, *52*(1-2), 87–106. https://doi.org/10.1016/0001-6918(82)90028-2

Mohammed, S., & Dumville, B. C. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior*, *22*(2), 89–106. https://doi.org/10.1002/job.86

Morris, N. M., & Rouse, W. B. (1985). The effects of type of knowledge upon human problem solving in a process control task. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-15*(6), 698–707. https://doi.org/10.1109/TSMC.1985.6313453

Morrison, D. L., & Upton, D. M. (1994). Fault diagnosis and computer integrated manufacturing systems. *IEEE Transactions on Engineering Management*, *41*(1), 69–83. https://doi.org/10.1109/17.286326

Mumaw, R. J., Roth, E. M., Vicente, K. J., & Burns, C. M. (2000). There is more to monitoring a nuclear power plant than meets the eye. *Human Factors*, *42*(1), 36–55. https://doi.org/10.1518/001872000779656651

Nungester, R. J., & Duchastel, P. C. (1982). Testing versus review: Effects on retention. *Journal of Educational Psychology*, *74*(1), 18–22. https://doi.org/10.1037/0022-0663.74.1.18

Oser, R. L., Prince, C., Morgan, B. B., JR., & Simpson, S. S. (1991). *An analysis of aircrew communication patterns and content* (No. NTSC-TR-90-09): Naval Training Systems Center Orlando Fl.

Patrick, J., Gregov, A., Halliday, P., Handley, J., & O'Reilly, S. (1999). Analysing operators' diagnostic reasoning during multiple events. *Ergonomics*, *42*(3), 493–515. https://doi.org/10.1080/001401399185603

Perrow, C. (1999). *Normal accidents: Living with high-risk technologies ; with a new afterword and a postscript on the Y2K problem*. Princeton, NJ: Princeton Univ. Press. Retrieved from http://www.loc.gov/catdir/description/prin021/99032990.html

Raaijmakers, J. G. W., & Verduyn, W. W. (1996). Individual differences and the effects of an information aid in performance of a fault diagnosis task. *Ergonomics*, *39*(7), 966–979. https://doi.org/10.1080/00140139608964517

Rasker, P. C., Post, W. M., & Schraagen, J. M. (2000). Effects of two types of intra-team feedback on developing a shared mental model in command & control teams. *Ergonomics*, *43*(8), 1167–1189. https://doi.org/10.1080/00140130050084932

Rasmussen, J., & Jensen, A. (1974). Mental procedures in real-life tasks: A case study of electronic trouble shooting. *Ergonomics*, *17*(3), 293–307. https://doi.org/10.1080/00140137408931355

Reigeluth, C., Merrill, M.D., Wilson, B., & Spiller, R. (1980). The elaboration theory of instruction: A model for sequencing and synthesizing instruction. *Instructional Science*, *9*(3). https://doi.org/10.1007/bf00177327

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, *100*(3), 349–363. https://doi.org/10.1037/0033-2909.100.3.349

Sander, P. C., van Doorn, R. R. A., van der Pal, J., & Zijlstra, F. R. H. (2015). Team adaptation to an unforeseen system failure: Limits of the potential aids of shared knowledge and standardized communication. *European Journal of Work and Organizational Psychology*, *24*(5), 796–811. https://doi.org/10.1080/1359432X.2015.1006199

Sauer, J., Felsing, T., Franke, H., & Ruttinger, B. (2006). Cognitive diversity and team performance in a complex multiple task environment. *Ergonomics*, *49*(10), 934–954. https://doi.org/10.1080/00140130600577502

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? *Methodology*, *6*(4), 147–151. https://doi.org/10.1027/1614-2241/a000016

Schraagen, J. M., & Rasker, P. C. (2001). Communication in command and control teams. In US Naval Academy (Ed.), *Proceedings 6th international command and control research and technology symposium.*

Serfaty, D., Entin, E. E., & Volpe, C. (1993). Adaptation to stress in team decision-making and coordination. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *37*(18), 1228–1232. https://doi.org/10.1177/154193129303701806

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (Complete samples). *Biometrika*, *52*(3/4), 591. https://doi.org/10.2307/2333709

Smith-Jentsch, K. A., Mathieu, J. E., & Kraiger, K. (2005). Investigating linear and interactive effects of shared mental models on safety and efficiency in a field setting. *Journal of Applied Psychology*, *90*(3), 523–535. https://doi.org/10.1037/0021-9010.90.3.523

Stachowski, A. A., Kaplan, S. A., & Waller, M. J. (2009). The benefits of flexible team interaction during crises. *Journal of Applied Psychology*, *94*(6), 1536–1543. https://doi.org/10.1037/a0016903

Sterman, J. D. (1994). Learning in and about complex systems. *System Dynamics Review*, *10*(2-3), 291–330. https://doi.org/10.1002/sdr.4260100214

Stout, R. J. (1995). Planning effects on communication strategies: A Shared Mental Model Perspective. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *39*(20), 1278–1282. https://doi.org/10.1177/154193129503902008

Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, shared mental models, and coordinated performance: An Empirical Link Is Established. *Human Factors*, *41*(1), 61–71. https://doi.org/10.1518/001872099779577273

Turner, J. R., Chen, Q., & Danks, S. (2014). Team shared cognitive constructs: A meta-analysis exploring the effects of shared cognitive constructs on team performance. *Performance Improvement Quarterly*, *27*(1), 83–117. https://doi.org/10.1002/piq.21163

U.S. Chemical Safety and Hazard Investigation Board. (2007). *Investigation report: Refinery explosion and fire*. Retrieved from http://www.csb.gov/bp-america-refinery-explosion/

Uitdewilligen, S., Waller, M. J., & Zijlstra, F. R. H. (2010). Team cognition and adaptability in dynamic settings: A review of pertinent work. In G. P. Hodgkinson & J. K. Ford (Eds.), *International review of industrial and organizational psychology: Vol. 25, 2010.* Chichester, West Sussex: Wiley-Blackwell.

Usher, J. M., & Kaber, D. B. (2000). Establishing information requirements for supervisory controllers in a flexible manufacturing system using GTA. *Human Factors and Ergonomics in Manufacturing*, *10*(4), 431–452. https://doi.org/10.1002/1520-6564(200023)10:4<431::AID-HFM5>3.0.CO;2-C

van den Bossche, P., Gijselaers, W., Segers, M., Woltjer, G., & Kirschner, P. (2011). Team learning: Building shared mental models. *Instructional Science*, *39*(3), 283–301. https://doi.org/10.1007/s11251-010-9128-3

Verdonik, D. (2010). Between understanding and misunderstanding. *Journal of Pragmatics*, *42*(5), 1364–1379. https://doi.org/10.1016/j.pragma.2009.09.007

Wadhera, R. K., Parker, S. H., Burkhart, H. M., Greason, K. L., Neal, J. R., Levenick, K. M.,. . . Sundt, T. M. (2010). Is the "sterile cockpit" concept applicable to cardiovascular surgery critical intervals or critical events? The impact of protocol-driven communication during cardiopulmonary bypass. *The Journal of thoracic and cardiovascular surgery*, *139*(2), 312–319. https://doi.org/10.1016/j.jtcvs.2009.10.048

Zajac, S., Gregory, M. E., Bedwell, W. L., Kramer, W. S., & Salas, E. (2014). The cognitive underpinnings of adaptive team performance in ill-defined task situations. *Organizational Psychology Review*, *4*(1), 49–73. https://doi.org/10.1177/2041386613492787

7 Appendix

A Content of the Data CD

- Instruction Seminars
- Knowledge Test
- Overview of Scenarios
- Information Overview for Participants
- Questionnaires

Declaration of Authorship

I, Rica Bönsel, hereby declare that the submitted master thesis, written under the supervision of Dr. rer. nat. Romy Müller, is my own unaided work. I used no other aids than indicated, and all direct or indirect sources are acknowledged as references.

Dresden, January 30th 2018

signature