Cognitive diversity and team performance in a complex multiple task environment

JÜRGEN SAUER*†, TOBIAS FELSING‡, HOLGER FRANKE‡ and BRUNO RÜTTINGER‡

†Department of Psychology, University of Fribourg, Rue de Faucigny 2, 1700 Fribourg, Switzerland
‡Institute of Psychology, Darmstadt University of Technology, Alexanderstr.1, 64283 Darmstadt, Germany

This article examines the multiple effects of cognitive diversity in teams operating complex human-machine-systems. The study employed a PC-based multiple-task environment, called the Cabin Air Management System, which models a process control task in the operational context of a spacecraft's life support system. Two types of cognitive diversity were examined: system understanding and team specialization. System understanding referred to the depth of understanding team members were given during training (low-level procedure-oriented vs. high level knowledge-oriented training). Team specialization referred to the degree to which knowledge about system fault scenarios was distributed between team members (specialized vs. non-specialized). A total of 72 participants took part in the study. After having received 4.5 h of training on an individual basis, participants completed a 1-h experimental session, in which they worked in two-person teams on a series of fault scenarios of varying difficulty. Measures were taken of primary and secondary task performance, system intervention and information sampling strategies, system knowledge, subjective operator state, communication patterns and conflict. The results provided evidence for the benefits of cognitive diversity with regard to system understanding. This manifested itself in better primary task performance and more efficient manual system control. No advantages were found for cognitive diversity with regard to specialization. There was no effect of cognitive diversity on intra-team conflict, with conflict levels generally being very low. The article concludes with a discussion of the implications of the findings for the engineering of cognitive diversity in teams operating complex human-machine-systems.

Keywords: Team performance; Team diversity; Training; Conflict; Team heterogeneity

1. Introduction

1.1. *Teamwork in complex technical work environments*

Teamwork is a common feature in a number of technical work environments, such as aircraft cockpits, nuclear power plants and ship's bridges. In most of these teams, performance is a critical issue since it is associated with system safety and productivity. There is an extensive body of literature on the performance of work teams (e.g. Cordery 2002). This literature has looked at a range of factors that are associated with team performance, such as job design, interdependence, team processes, team context and team composition. One aspect of team composition refers to the differences between team members (TMs) or, put differently, the degree of diversity within a team. The present article is concerned with the question of how the performance of two-person teams working with complex technical systems is affected by within-team diversity.

Within the literature on team diversity, two main research strands can be identified. One is concerned with the importance of shared mental models between TMs (i.e. reducing within-team diversity) managing highly complex technical systems (e.g. Cannon-Bowers *et al.* 1993). The other is concerned with increasing team diversity as a means to improve overall team performance (e.g. Webber and Donahue 2001). The two research strands are characterized by somewhat opposing basic assumptions, with the former emphasizing the importance of similarity of TMs with regard to their mental model while the latter stresses the benefits of having differences between TMs to improve overall team performance. The strands also differ with regard to the dominant task environment in which the research is carried out. The literature on shared mental model largely focuses on teamwork during the use of technical systems while the diversity literature is more concerned with decision-making in management and project teams. However, both research strands are united in pursuing the critical question about the optimal degree of team diversity under given situational circumstances. For the present study, relevant work from both research strands is examined.

1.2. *Cognitive diversity in teams*

Differences between members of a team are often referred to as 'diversity' (Jackson *et al.* 1995). Since TMs can differ from one another in many attributes, diversity is considered a multifaceted concept (Stumpf and Thomas 1999). In a model by McGrath *et al.* (1995) attributes of diversity in teams are assigned into five clusters: (i) demographic attributes (e.g. age, gender, functional background); (ii) task-related knowledge, skills and abilities; (iii) values, beliefs and attitudes; (iv) personality and cognitive and behavioural styles; (v) status in the work group's embedding organization (e.g. organizational rank). The clusters are not independent of each other. For example, diversity in demographic factors such as functional background is linked to differences in task-related attributes, such as knowledge, skills and abilities. The multifaceted nature of diversity is also reflected in a model by Jackson (1996), which classifies diversity in work groups into two bipolar dimensions: (a) readily detected attributes vs. underlying attributes; and (b) task-related attributes vs. relationship-oriented attributes. This results in a two-by-two table, into which TM attributes can be assigned. An example for an underlying and task-related attribute would be knowledge and expertise of the work team, whereas gender would be a relationship-oriented and readily detected attribute.

The present paper is concerned with diversity in underlying and task-related attributes, such as abilities, knowledge, expertise and problem-solving strategies. In theories of

information processing and decision-making, these attributes are often described as particularly relevant for team performance (Milliken and Martins 1996, Williams and O'Reilly 1998). Because of their cognitive nature, the term 'cognitive diversity' will be used to refer to within-team differences in these attributes.

Evidence from the research literature suggests that the relationship between cognitive diversity and team performance is of considerable complexity, as it may be affected by a number of moderating variables (Milliken and Martins 1996, Shaw and Barrett-Power 1998, Williams and O'Reilly 1998). McGrath *et al.* (1995) have identified three factors that may moderate that relationship: work group characteristics; technology; and task.

- Work group characteristics: Work teams differ with regard to their underlying function (e.g. operating a human-machine-system, working on a time-limited project). McGrath *et al.* (1995) have argued that teams operating human-machine-systems (called crews in their terminology) benefit less from team diversity than other forms of work groups, such as task forces and management teams. This is largely due to the standardized way of completing task activities often found in teams operating technical systems. However, there are also situations with lower levels of standardization (e.g. emergencies), in which team diversity would be more advantageous.
- Technology: McGrath *et al.* (1995) refer to communication technology as an influencing factor. While communication technology is of importance, there are much broader technological issues that need to be looked at, too. Technical systems may differ with regard to a number of factors, such as application area (e.g. transportation, process control), system complexity (e.g. time lags) and task interdependence. The underlying logic of the application area determines the level of team diversity required. For example, while considerable overlap between the knowledge structures of TMs is typically found in civilian aircraft (i.e. low diversity), space crews are much more diverse because of differences in crew member qualification (e.g. physician, engineer).
- Task: The importance of the task type as a moderator of the relationship between team diversity and performance is stressed by several authors (Shaw and Barrett-Power 1998, Williams and O'Reilly 1998). For example, tasks that involve the solving of complex problems benefit more from task-related diversity than less complex tasks (Mohammed and Ringseis 2001). In the present study, tasks of different complexity levels were examined to identify any differential effects of cognitive diversity.

Although cognitive diversity may provide benefits for team performance, it also involves the risk of incurring process losses that may partly offset these advantages. This has been coined the 'double-edged sword'' of team diversity (Milliken and Martins 1996).

The potential for conflict in highly diverse groups has been considered such a risk factor for process losses (Williams and O'Reilly 1998). Empirical work has found that certain aspects of diversity were connected to certain types of conflict (Pelled *et al.* 1999). For example, diversity in task-related aspects (e.g. professional background) was associated with task-related conflict but not with emotional conflict. Conversely, diversity in relationship-oriented attributes (e.g. age) was associated with emotional conflict but not with task-related conflict. Overall, task-related conflict is not necessarily undesirable since it may result in better decisions (Priem *et al.* 1995).

While most of the work on team diversity has been carried out in a non-technical work context, a smaller number of studies has also been conducted in the context of complex technical work environments. That work may be summarized under the heading of 'shared mental models', as it examines to what extent a strong overlap of TMs' mental models is of benefit for team performance. The studies have been conducted in a range of settings, such as simulated military applications (e.g. Stout *et al.* 1999, Cooke *et al.* 2003), air traffic control (e.g. Tschan *et al.* 2000) but also in generic laboratory settings (e.g. Banks and Millward 2000). The findings from these studies generally suggest that a strongly shared mental model is advantageous for team performance.

However, not all aspects of the mental model need to be shared to the same extent. Cannon-Bowers *et al.* (1993) have argued that it is very important that the following aspects of a mental model are shared: task (e.g. managing system faults); team interaction (e.g. communication); and team resources (e.g. knowledge about team mate's abilities and skills). However, it is less critical that there is a strong overlap for the mental model of the equipment (i.e. knowledge about system). Another distinction was made by Rouse and Morris (1986) between rapidly evolving dimensions of the mental model (e.g. current state of the system) and temporally more stable dimensions (e.g. system function, form and purpose). The current state of the system shows considerable variation over time since this aspect of the mental model evolves with the changing situation. Whilst the reviewed work from the literature has been primarily concerned with the mental model of the current system state, the present study focuses on the temporally more stable dimensions of the mental model. While a high degree of mental model overlap appears to be beneficial with regard to the current system state, it remains to be empirically tested whether this would also apply to temporally more stable aspects of the mental model. In particular, it needs to be examined how team performance is affected if there are differences in the performance potential of individual TMs. A classification system by Steiner (1972) proposes several principles, according to which group task performance is determined by the performance potential of each TM. First, the disjunctive task principle may apply, which proposes that the best TM determines overall team effectiveness. Second, the compensatory task principle may apply, suggesting that the performance of single TMs is averaged. Third, the conjunctive task principle may apply, which proposes that team performance equals the performance of the worst TM. Cognitive diversity in teams would only be beneficial if the disjunctive task principle applied.

1.3. *The present study*

In the present study, a complex work environment was modelled by means of a computer-based simulation, which is suitable for performance tests for individuals as well as teams. The good data-gathering facilities of the simulation environment permitted the collection of performance measures on a range of task components. This is in contrast to most previous studies examining team diversity, which did not collect any performance measures or needed to rely on self-assessed performance alone. In addition to performance measures, important supplementary measures were collected to assess the multiple effects of cognitive diversity, such as information sampling and control behaviour, subjective operator state, system knowledge, team perception, conflict and communication patterns.

The current study used training as a means to create cognitive diversity within teams. Of the many facets of cognitive diversity, this study examined two aspects: system understanding and specialization.

The main independent variable examined was the type of system understanding, which refers to the level of understanding operators had of the system functions (e.g. components, cause–effect relationships between parameters, constraints of operations). This was an important aspect because creating an effective mental model of the operator has long been considered a principal factor for optimizing the human-machine-system. In the first experimental group, both TMs were provided with a deep understanding of the system (system-oriented training; SOT). In the second group (procedure-oriented training; POT), both TMs were given precise procedures to follow when a system disturbance occurred. These two homogeneous groups were compared to a third group (mixed; MIX), in which a system-trained and a procedure-trained TM were paired up, hence representing a diverse group with regard to their understanding of the system. The design of the training regimes to achieve different levels of system understanding was based on a distinction by Morris and Rouse (1985), who identified two forms of training (guidance in the use of system knowledge, guidance in the use of rules or algorithms) as more effective than alternative training regimes. Based on that theoretical distinction, two training regimes have been developed in an earlier study (Sauer *et al.* 2000a), upon which the training regimes used in the present work were based.

A further independent variable in the present study referred to the degree of TM specialization. Since training time is at a premium in many work environments, an effective allocation of resources is required. Against this background, this study addressed the question of whether an operator should only be trained on the management of a selected set of possible fault scenarios (e.g. those with a higher frequency) or should the operator be given a very broad introduction to all possible fault scenarios? If there is a focus on a selected set, the operator might benefit from overtraining (Patrick 1992), although this may be at the expense of knowledge transfer problems to unfamiliar fault states. In technical systems managed by a single operator, one may be inclined to opt for a broader introduction to fault scenarios to ensure that the operator is not confronted with unfamiliar fault states, which he/she may not be able to manage. However, in the context of teamwork one can also see the benefits of stronger specialization since TMs are able to complement each other in their fields of expertise. Against this background, the present study examined whether a low or high level of specialization would be more advantageous for team performance. In one condition, TMs would have a broad but less profound knowledge of fault scenarios (i.e. low specialization of TMs because their knowledge strongly overlapped). In the other condition, each TM's knowledge would be limited to a small set of possible fault scenarios, but for these selected faults scenarios knowledge would be much more profound than in the other condition (i.e. high specialization of TMs because their knowledge of fault scenarios did not overlap).

On the basis of the research literature reviewed, a number of predictions were made. First, a complex pattern of effect was expected for diversity in system understanding. It was predicted that the homogeneous teams (SOT and POT) were particularly effective under specific operational conditions. SOT would show better primary task performance than POT at the most difficult fault scenarios. This was because of their deeper understanding of the task environment, which enabled SOT-teams to transfer their better knowledge to unfamiliar situations more easily. Conversely, it was hypothesized that POT-teams showed better primary task performance than SOT-teams on familiar faults because the former were more familiar with the application of effective procedures. Whilst directed expectations were put forward for the homogeneous POT and

SOT-teams, no such clear predictions were made for diverse teams. It may be that the best TM in this MIX team would determine team performance (i.e. the disjunctive task principle of Steiner's classification system applies) or, alternatively, any of the two other principles (conjunctive or compensatory) may apply. Therefore, an important goal of this piece of research was to determine which principle would be most likely to apply. Second, with regard to team specialization, it was predicted that specialized teams would perform better than non-specialized teams at familiar fault scenarios because the specialized TM had more experience of dealing with the system fault than a TM who received a broader introduction to system faults. Third, the effects of cognitive diversity were expected to be task dependent in that more complex tasks (e.g. fault diagnosis) would benefit more from diversity than less complex tasks (e.g. alarm acknowledgement).

2. Method

2.1. *Participants*

A total of 72 participants (8.3% female) took part in the study. All of them were students from science and engineering departments of Darmstadt University of Technology. This was to ensure that they had a satisfactory understanding of technical systems, similar to operators of real technical systems. Their ages ranged from 20 to 40 years (mean 24.5 years). Participants received a payment of €50 for their involvement.

2.2. *Design*

A $3 \times 2 \times 4$ mixed design was employed in the experiment, with two between-participants variables (system understanding, specialization) and one within-participants variable (difficulty of fault scenario). System understanding was manipulated at three levels (POT, SOT and MIX, i.e. POT for one TM and SOT for the other) whilst specialization was varied at two levels (specialized teams vs. non-specialized teams). Within each 1-h testing period, four levels of the factor 'difficulty of fault scenario' were included: fault-free (F-free); practised faults (PracF); novel faults (NovF); control panel failures (CPF). Since the unit of analysis was a two-person team (i.e. 72 participants formed 36 teams), each of the six cells contained six cases.

The variable 'scenario difficulty' was varied at four levels: (i) In the F-free condition, the automatic system was perfectly reliable; (ii) PracF referred to fault states that participants had already extensively practised during the training sessions; (iii) NovF were fault states that participants had not encountered before but they were of the same quality as the practised faults; (iv) CPF referred to a group of faults that were much more difficult to manage than those practised in training. This was because the most obvious strategy of dealing with them was unavailable (i.e. the control panel that was necessary to manage the fault was out of order). Furthermore, it was not possible to repair the fault since the maintenance facility was also unavailable.

2.3. *Simulation task*

In this study, a task environment was employed that modelled a highly automated process control environment in the operational context of a spacecraft's life support system. The task environment, known by the name of Cabin Air Management System

(CAMS), consists of five automatic controllers that maintain their corresponding system parameters (O_2, CO_2, cabin pressure, temperature and humidity) within a predefined target range. The main interface of CAMS may be found in figure 1. This screenshot displays the topographical layout of the system with its principal features, comprising history display, parameter control panels, maintenance facility, alarm system and flow meter readings.

The operator is required to complete four tasks, which are divided into primary and secondary tasks according to the priorities attached to them. The two primary tasks are system stabilization and fault diagnosis. Maintaining a stable system state is achieved by monitoring the safe functioning of the automatic controllers and, if required in the event of a system fault, by adopting manual control. Fault diagnosis refers to the identification and repair of system faults by making use of the CAMS maintenance facility. The two secondary tasks are responding to system alarms and carrying out status checks at regular intervals. The acknowledgement of system alarms was a reaction time task that measured the response latency of operators. The status of the O_2 tanks had to be checked regularly (i.e. at 3-min intervals), which represented essentially a prospective memory task.

In addition to these performance measures, CAMS also allowed the collection of data on different aspects of the subjective state of the operator, including mental effort, anxiety and fatigue. This was done at 30-min intervals by means of visual analogue scales (100 mm lines), which were embedded in the CAMS environment. All system and operator responses are automatically written into a results file, allowing the reconstruction of an experimental session for further analysis. A detailed description of CAMS may be found in Sauer *et al.* (2000b).
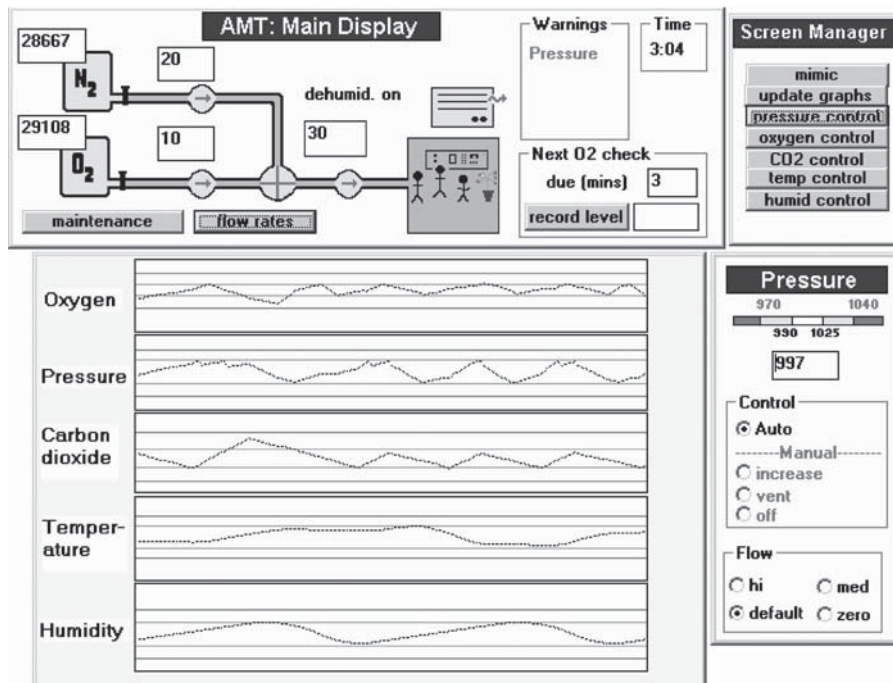

Figure 1. Main interface of Cabin Air Management System.

2.4. *Training*

Participants were randomly assigned to one of four training conditions, which consisted of two dimensions with two levels each: specialization (specialized vs. non-specialized); system understanding (procedure- vs. system-oriented). As the two dimensions were completely crossed, four different training conditions emerged, such as high specialization combined with system-orientation or low-specialization combined with procedure-orientation (see table 1). The precise content of each training approach is described below.

Since the CAMS task was of considerable complexity, the participants received extensive training in small groups of three to four. Total training time was approximately 4.5 h, given in two separate sessions. In the first session (2.5 h), participants were taught about the basic operation of CAMS and the set of tasks they had to complete during the experimental session. After having worked with CAMS during normal system operation, participants received extensive practice on a set of system faults. In the second training session (2 h), more intensive practice on these system faults was given. A fault finding guide was provided to all participants, which they could consult during training and the experimental testing session. The training approach was highly interactive, with the participants completing especially designed training scenarios on individual computers. At the end of the training programme, a 7-min performance test was given to determine whether participants had received acceptable performance levels. All participants reached satisfactory standards and none had to be excluded.

2.4.1. Specialization. In specialized teams, each TM was intensively trained on three fault states, with each fault scenario being practised six times. A TM would either be trained on set A of fault states (i.e. leak of O_2 valve, block of N_2 valve, vent permanently open) or on set B (i.e. CO_2 scrubber ineffective, set point failure of N_2 controller, O_2 valve permanently open). For the experiment, a TM trained on set A would be paired up with a TM trained on set B so that they would complement each other with regard to their familiarity with a complete set of six system faults. In non-specialized teams, each TM was trained on the complete set of all six faults (i.e. both TMs received exactly the same training) before they were paired up for the experimental session. Because non-specialized TMs were introduced to more fault states than specialized TMs, non-specialized TMs had less intensive practice on each of them, only facing each fault state three times during training (i.e. half as often as the specialized TMs).

Table 1. Design of training conditions.

	Type of system understanding		
	Procedure-oriented training		System-oriented training
Specialization			
Specialists	◖ or ◗	n = 18	◖ or ◗ n = 18
Generalists	○	n = 18	● n = 18

Filled symbol = system-oriented training; open symbol = procedure-oriented training; full circle = generalists; two half circles = specialists.

2.4.2. System understanding. A total of 36 participants were given POT and 36 participants received SOT.

The POT approach focused on the importance of procedures for system management. The instructor explained these procedures to the trainee, who had to follow them strictly during task completion. If the trainee adopted a different system management strategy, the instructor would intervene and stress the importance of the correct procedure being followed. No explanations were given to the trainee about how subsystems interact with each other or how system parameters are affected by fault states. If questions were asked to gain a deeper system understanding, it was emphasized that the most efficient way of managing CAMS was to follow the procedures precisely as taught during training and as outlined in the fault finding guide.

The SOT approach focused on the relationship and interaction between different system components. Trainees should make their interventions on the basis of their understanding of the cause of the problem rather than simply following the step-by-step procedures from the fault finding guide. Participants were permanently encouraged to make a priori predictions about system behaviour and to give a posteriori explanations after having observed actual system behaviour. The instructor would directly challenge explanations given by trainees if they deviated from the actual processes of the system.

2.5. *Experimental procedure*

For the experimental testing session, participants were assigned to two-person teams according to the experimental plan (see table 2). The experimenter ensured that teams were not formed with participants who had met prior to taking part in the experiment. During the 1-h experimental session, the teams had to deal with eight fault scenarios (four PracF, two NovF and two CPF), following a schedule that was known only to the experimenter.

TMs were told that they should work on the task scenarios together as equal partners. One of the TMs was put in charge of the mouse and keyboard for 30 min and then they changed so that the other TM was in charge of the controls for the remaining 30 min. This functional role was strictly defined and changed halfway through the experiment to achieve an equal contribution as possible from the TMs over the session.

The session was videotaped for subsequent analysis of communication patterns. A content analysis of intra-team communication (see Mayring 1988) was carried out,

Table 2. Experimental conditions.

	Type of system understanding					
	POT-team (two POT-TMs)		SOT-team (two SOT-TMs)		Mixed team (one POT-TM, one POT-TM)	
Team specialization	Generalist	Specialist	Generalist	Specialist	Generalist	Specialist
Number of teams	6	6	6	6	6	6
Team composition	○○	◐◐	●●	◑	○●	◑

POT = procedure-oriented training; SOT = system-oriented training; TM = team member; filled symbol = system-oriented training; open symbol = procedure-oriented training; full circle = generalists; two half circles = specialists.

with communication units being assigned to one of the following categories: (i) fault diagnosis (causes of a disturbed system state); (ii) intervention measures (control actions to stabilize the system); (iii) CAMS task (technical environment and its components); (iv) secondary task; (v) training (kind of training received); (vi) allocation of tasks (how should tasks be allocated within the team); (vii) personal matters; (viii) other.

A communicative act was defined as a statement made that referred to one of the topic categories (e.g. 'This must be an oxygen leak!'). However, short utterances, such as 'good' or 'yeah' were not considered as communicative acts since their meaning has often remained unclear. A sample of each tape was analysed, including three different fault scenarios (one PracF and two CPF). This made it possible to make a comparison between different levels of fault difficulty and to examine the effects of the functional role of the TM. The videotape analysis was carried out by a rater who was unaware of the study's research questions but was familiar with the task environment CAMS to be able to make sense of participant statements. To ensure the objectivity of the content analysis, inter-rater agreement was determined by employing a second rater on sub-samples of the videotape material. The inter-rater reliability coefficient (Cohen's Kappa) calculated provided a satisfactory result (K = 0.66).

After the completion of the working session, a set of three tests and questionnaires was given to participants in the following order:

1. To measure the TM's explicit (i.e. verbalizable) system knowledge, a test of user knowledge about CAMS was applied. The test had already been used in previous studies employing the CAMS task to assess the system knowledge of the participants (Sauer *et al.* 2000a). It comprised 12 three-alternative, multiple-choice items (e.g. What happens to the cabin temperature when nitrogen is pumped into the cabin? (a) increase, (b) decrease or (c) no effect) and three open questions about technical subsystems (a question each about O_2, pressure and temperature control, such as: Please explain which components or processes have an impact on cabin oxygen levels and describe the direction of that relationship).
2. A team evaluation questionnaire was especially developed for evaluating various aspects of teamwork. Comprising 17 items, the questionnaire had four sub-scales: (i) task engagement of participant and of team mate (six items); (ii) competence of participant and of team mate (two items); (iii) overall team behaviour (three items); (iv) perceived level of team diversity (three items). A 5-point Likert scale was used for all items. An example of an item was: My approach to system management was different from my team mate's (strongly agree/agree/neither agree nor disagree/ disagree/strongly disagree).
3. A German-language questionnaire (FAKT) was used to measure social conflict within the team. The FAKT-questionnaire measures different facets of social conflict (Windel *et al.* 1999, Adolph 2000). The number of items was reduced from 54 to 26 to exclude those items that were not applicable to teams in experimental settings.

3. Results

3.1. *Primary task performance*

3.1.1. System control failures.   This measure refers to the percentage of time any of the five key parameters had been in an unsafe state. The data underwent a logarithmic transformation to stabilize variances. The data (untransformed values) are presented in

table 3. The F-free condition was not included in the analysis since virtually no errors were committed under that condition. The data showed a main effect of system understanding (F = 4.81; df = 2, 30; $p < 0.05$). Post-hoc least significant difference (LSD)-tests indicated that performance for MIX and system-trained teams was significantly better than for procedure trained teams ($p < 0.05$). Surprisingly, the results showed better performance for non-specialized teams (i.e. generalists) than for specialized teams (F = 4.34; df = 1, 30; $p < 0.05$). No interaction between the two types of cognitive diversity was observed (F < 1). There was an effect of fault scenario difficulty in the predicted direction, with control errors increasing with more difficult fault scenarios (PracF 3.4%; NovF 6.0%; CPF 10.3%; F = 91.9; df = 2, 60; $p < 0.001$). No interaction between scenario difficulty and any of the other independent variables was observed.

3.1.2. Fault diagnosis.   The efficiency of fault diagnosis was evaluated by analysing the time needed for a correct diagnosis and repair. CPF were not included in this analysis because the maintenance facility was not available for that fault type. The data are presented in table 3. As for the system control failures, the analysis revealed a main effect of system understanding (F = 4.33; df = 2, 30; $p < 0.05$). MIX- and SOT-teams repaired system faults more quickly than POT-teams (LSD-test: $p < 0.05$). No effect of team specialization was found (F < 1). A main effect of scenario difficulty was found (F = 20.3; df = 1, 30; $p < 0.001$), with NovF requiring significantly longer to be identified than PracF (137.1 s vs. 90.8 s). No interaction was found.

3.2. *Secondary task performance*

3.2.1. Prospective memory (tank level recording).   This measure indicates the percentage of prospective memory failures (i.e. responses omitted or more than 20 s after scheduled time). The analysis revealed a significant main effect of scenario difficulty (F-free 28.5%; PracF 48.2%; NovF 41.0%; CPF 39.6%; F = 3.61; df = 3, 90; $p < 0.05$). Post-hoc LSD-tests indicated that only the difference between F-free and the other three fault conditions was significant ($p < 0.05$). No significant effect of system understanding was observed (F = 1.25; df = 2, 30; $p > 0.05$) and there was no significant difference between specialized and non-specialized groups (F = 1.80; df = 1, 30; $p > 0.05$).

3.2.2. Reaction time (annunciator acknowledgement).   This measure refers to the time needed to respond to the appearance of an annunciator. There was no significant effect of system understanding (F < 1) and no significant difference between generalists and

Table 3. Performance as a function of team specialization and system understanding.

	System-oriented training	Procedure-oriented training	Mixed team	Overall
System control failures (%)	6.48	7.26	5.99	
Team specialization	7.11	7.53	6.15	6.93
No team specialization	5.85	6.99	5.84	6.23
Diagnosis time (s)	100.8	137.5	104.7	
Team specialization	96.3	136.4	106.3	113.0
No team specialization	105.3	138.7	101.1	115.0

specialists (F < 1). However, there were indications of reaction time slowing down when scenario difficulty increased (F-free 1.60 s; PracF 1.91 s; NovF 1.86 s; CPF 1.95s; F = 5.28; df = 3, 90; $p < 0.005$). LSD-tests indicated that only the differences between F-free and the other fault scenarios were significant ($p < 0.01$).

3.3. *Control actions and information sampling behaviour*

3.3.1. Control actions. This measure indicates the frequency of system interventions carried out by the team (no./min). The analysis revealed that SOT-teams (1.60/min) intervened more frequently than POT-teams (1.11/min) or MIX-teams (1.09/min). This difference between SOT and the two other conditions was significant (F = 7.85; df = 2, 30; $p < 0.005$; LSD-test: $p < 0.001$). In particular, during the presence of CPF, the SOT-team carried out more interventions than the other teams, resulting in a significant interaction between system understanding and scenario difficulty (F = 5.67; df = 4, 60; $p < 0.001$), as the data in figure 2 demonstrate. A strong main effect of scenario difficulty was also observed, with teams carrying out more control actions during the presence of CPF than NovF or PracF (F = 26.8; df = 2, 30; $p < 0.001$; LSD-tests: $p < 0.001$), although the interaction was clearly more interesting. No effect of specialization was recorded (F < 1).

3.3.2. Flow meter sampling. An analysis of the sampling frequency of flow meters (they remained visible on screen for a period of 10 s when evoked) revealed no significant effects of cognitive diversity, neither for specialization (F < 1) nor for system
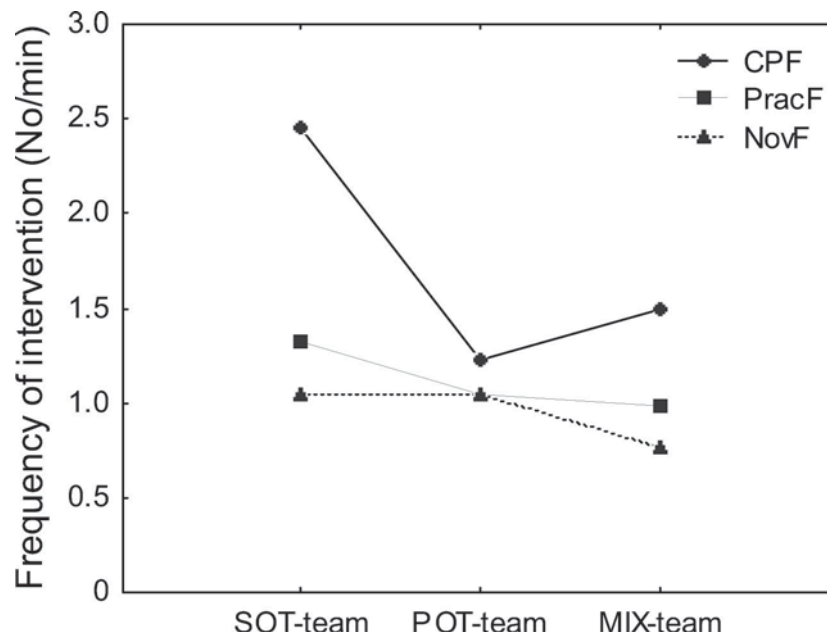

Figure 2. Control action as a function of system understanding and fault scenario. CPF = control panel failures; PracF = practised faults; NovF = novel faults; SOT = system-oriented training; POT = procedure-oriented training; MIX = mixed teams.

understanding (F = 1.13; df = 2, 30; $p > 0.05$). However, a strong effect of scenario difficulty was found (F = 50.6; df = 3, 90; $p < 0.001$), with sampling frequencies decreasing with increasing scenario difficulty (F-free 2.04/min; PracF 1.66/min; NovF 1.54/min; CPF 1.02/min). No interactions were found.

3.3.3. History display sampling. This measure indicated the percentage of time during which the history display was shown on screen (when evoked, it remained visible for a period of 30 s). Overall, it was on-screen for 75.7% of the time, although there was little difference between experimental diversity conditions (all F < 1). A strong main effect of fault scenario was found (F = 23.6; df = 3, 90; $p < 0.001$), with a higher sampling rate during fault-free phases (80.9%) than the other three (PracF 74.8%; NovF 73.3%; CPF 73.8%). This was confirmed by LSD-tests ($p < 0.001$).

3.4. *Subjective operator state and team evaluation*

3.4.1. Subjective operator state. As an embedded feature of CAMS, the subjective state of the participants was measured at 30-min intervals. Responses were made on 100-mm visual analogue scales for three variables: effort, anxiety and fatigue. No effect was found for any of the two types of cognitive diversity on mental effort expenditure, anxiety and fatigue (for all F < 1). However, ratings increased for all three measures, as the experimental session progressed. The data are presented in figure 3. Mean team ratings of mental effort increased from T_{30} (+30 min) to T_{60} (+60 min) significantly (F = 20.4; df = 1, 30; $p < 0.001$). A similar pattern was found for anxiety, with team ratings also monotonously increasing over the three points of measurement, with the first being taken prior to the beginning of the session. This effect was highly significant (F = 25.1; df = 2, 60; $p < 0.001$), with LSD-tests confirming that all three levels were different from each other ($p < 0.01$). Fatigue showed largely the same pattern, with TMs becoming increasingly tired, as the session progressed (F = 4.31; df = 2, 60; $p < 0.05$). Here, only the difference between T_0 and T_{60} was significant (LSD-test: $p < 0.05$).

3.4.2. Within-team analysis. An additional within-team analysis was carried out to determine whether the functional role within the team affected the operational state of the participant. This was achieved by entering functional role as an additional independent variable into the ANOVA. The results showed that the role of the controller (i.e. participant using the mouse and keyboard) was associated with higher subjective state ratings on all three measures than the role of the co-controller (i.e. participant not using the mouse and keyboard). When the participants swapped their roles halfway through the experiment, ratings changed accordingly. The effects of functional role were highly significant for mental effort (F = 41.5; df = 1, 30; $p < 0.001$) and anxiety (F = 32.5; df = 1, 30; $p < 0.001$), and moderately significant for fatigue (F = 5.08; df = 1, 30; $p < 0.05$). Interactions between functional role and any other independent variable were not observed. The data for all three subjective measures are presented in figure 3.

3.4.3. Team evaluation questionnaire. ANOVAs were carried out on four sub-scales of the questionnaire to identify any effect of cognitive diversity on team evaluation. The analysis revealed little awareness of cognitive diversity among TMs (neither for system understanding nor for team specialization), with all tests of main effects and interactions showing non-significant results. For the subscale level of diversity, even a single item
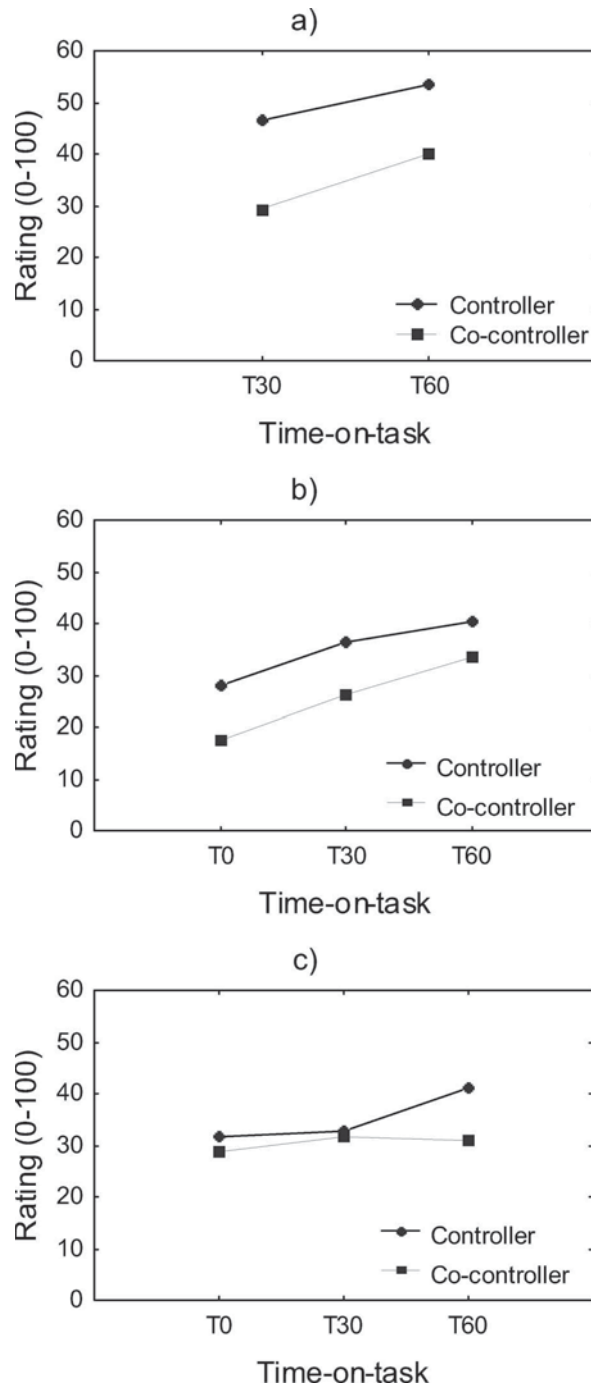
Figure 3. Ratings of subjective state measures (0–100) as a function of time on task and functional role for (a) mental effort, (b) anxiety, and (c) fatigue. T_0 = prior to task session; T_{30} = 30 min into task; T_{60} = 60 min into task.

analysis (within-team differences in task management strategies, in system knowledge and in overall performance) showed no significant effects (all F < 1).

3.5. *Communication patterns and social conflict*

3.5.1. Videotape analysis. Overall, the videotape analysis showed that within-team communication was largely task-related. The main topic of within-team communication was intervention measures (5.69 communicative acts/min), followed by fault diagnosis (1.34/min) whilst other issues featured much less prominently in within-team conversation. This includes communication topics such as CAMS task (0.77/min), secondary tasks (0.71/min) and personal matters (0.40/min). Interestingly, there was very little explicit communication about the allocation of tasks between TMs (0.06/min) and little reference was made to the content of the training received (0.16/min).

ANOVA was carried out on the aggregated communication frequency score (see table 4), revealing an effect of system understanding that showed more communication in SOT-teams than in the other two (F = 3.94; df = 2, 21; $p < 0.05$; LSD-test: $p < 0.05$). The analysis also provided evidence for more within-team communication for non-specialized teams than for specialized ones (F = 4.84; df = 1, 21; $p > 0.05$). With regard to fault type, there was no difference between the three fault types (PracF 9.24/min; CPF1 9.44/min; CPF2 10.14/min; F = 2.75; df = 2, 42; $p > 0.05$). No interaction was recorded. Separate ANOVAs were carried out on each of the eight communication topics. Two of them (intervention and secondary task) showed a similar pattern to the aggregated communication frequency score, whereas for the other categories no effects of the independent variables were found.

3.5.2. Social conflict questionnaires. Levels of conflict within the team were generally very low, as the ratings of the FAKT questionnaire have indicated. The mean rating across all scales for the whole sample was 1.66 on a scale ranging from 1 to 5. Only one of the sub-scales (conflict due to interdependence) showed a higher rating (mean 2.78). ANOVA showed no effects of cognitive diversity, neither for system understanding (F < 1), nor for specialization (F < 1).

3.6. *System knowledge*

The data collected by the knowledge questionnaire were analysed for individual TMs (i.e. n = 72). Each proposition made by a TM about CAMS was scored with regard to its correctness (maximum score was 33). The analysis revealed that SOT-TMs made a higher number of correct propositions (mean 16.0) than POT-TMs (mean 12.9). This difference in the quality of explicit system knowledge was significant

Table 4. Communication frequency (communicative acts/min) as a function of system understanding and team specialization.

	System-oriented training	Procedure-oriented training	Mixed team	Overall
Overall	11.67	8.82	8.43	
Team specialization	10.01	7.86	7.99	8.62
No team specialization	13.75	9.77	9.17	10.95

(F = 7.42; df = 1, 65; $p < 0.005$). There was no difference as a function of specialization (generalists: mean 14.4; specialists: mean 14.3; F < 1) and no interaction was observed (F < 1). Separate analyses of the two sections of the knowledge questionnaire (multiple choice and open questions) provided the same pattern of results; that is, an effect of system understanding but none of specialization.

4. Discussion

The study aimed to investigate the effects of two types of cognitive diversity on team performance in a technical work environment: system understanding and team specialization. The data analysis revealed a number of interesting results, which may be summarized into four major points. First, there were differences in performance between the two types of cognitive diversity. Diversity in system understanding showed some advantages over homogeneous teams, whereas no such effect was observed for diversity in team specialization. Second, the results indicated that the effects of cognitive diversity were moderated by task complexity, with only complex tasks benefiting from cognitive diversity whilst more simple tasks were unaffected. Third, typical problems associated with diversity (e.g. increased conflict) were not observed, with levels of conflict being generally rather low. Fourth, diverse teams were, surprisingly, not aware of their diversity. These main findings are now discussed in more detail.

The results for diversity in system understanding partly supported the assumption that diverse teams were superior to homogeneous ones. As expected, the MIX-teams showed a significantly higher performance than the POT-teams. In contrast to expectations, there were no differences between MIX-teams and SOT-teams. These results suggest that Steiner's (1972) disjunctive task principle has been applicable to teamwork in a task environment such as CAMS. As the disjunctive task principle proposes that the best TM determines overall team effectiveness, this means in the present study that within MIX-teams the SOT-TM determined team performance and, therefore, the performance of the MIX-teams was similar to the SOT-teams rather than the POT-teams.

However, at this point one needs to explore why MIX-teams did not benefit as much from their diversity as expected. Because of the overall low level of conflict observed, the 'process loss'-explanation (Milliken and Martins 1996) can be discounted, which suggests that diversity is associated with certain costs (e.g. conflict or communication problems) that offset some of its advantages. Instead, the pattern of results might be explained as follows. More benefits of diversity in system understanding would have been found if a broader range of working conditions had been simulated, allowing POT-TMs to make more use of their performance potential. The working conditions simulated in the present study may not have been sufficiently favourable for POT to benefit. For example, if environmental stressors such as noise are present, POT may have had advantages over SOT, as has been demonstrated in a previous study (Hockey *et al.* 2005). This was because POT has led to cognitively less demanding control strategies, requiring fewer cognitive resources for task completion and, hence, task performance was less vulnerable to environmental stressors than performance of operators having received SOT. As no such adverse working conditions were present in the current study, it was not possible to demonstrate the differential benefits of POT for performance in diverse teams.

Despite the absence of these differential benefits for POT-teams, some evidence for the utility of POT was provided by the data on the efficiency of system management behaviour. This manifested itself in more efficient system control strategies for

POT-teams and MIX-teams (due to their procedure-trained TM) compared to SOT-teams, with intervention efficiency being measured by the number of system control actions to achieve system stability. Whilst intervention efficiency is not a primary task goal, it may be considered a secondary aspect of successful system management since it allows the conservation of operator resources and the reduction of machine wear. More frequent system intervention shown by system-trained operators compared to procedure-trained operators has also been found in a comparative study (Sauer *et al.* 2000a). This difference between training regimes may be due to two factors. First, SOT-teams were not given a set of effective procedures to deal with fault scenarios so that they had to develop their own, which they derived from their overall system understanding. These were generally not as efficient (in the sense of minimizing the number of interventions) as the ones given to the POT-teams. Second, SOT encouraged operators to explore the system's properties, which has led to interventions that were not absolutely necessary but may have improved their general understanding of the system. The procedure-trained TM in the MIX-team supported the system-trained TM in finding more efficient intervention strategies that allowed the MIX-team to achieve the same performance levels as the SOT-team while accomplishing that goal more efficiently.

In contrast to diversity in system understanding, there were no benefits for diversity in team specialization, with the system stabilization measure even showing advantages for the homogeneous, that is, non-specialized teams. The generalist teams may have benefited from a broader knowledge of CAMS because they were trained on a larger number of system faults. In contrast, team specialization did not provide any advantages because TMs did not communicate as much as needed to ensure that they made best use of their team resources. Furthermore, it is conceivable that training may have led to the unintended side effect that specialists were more similar to POT-teams with regard to fault management strategies (i.e. strong proceduralization because they practised few fault states very intensively) and generalists were more similar to SOT-teams (i.e. broader system understanding because they were trained on a larger number of system faults but practised these less thoroughly).

All effects of cognitive diversity were observed in the more complex primary tasks while no such effects were found for the less complex secondary tasks (complexity may be defined here as the training time required for the operator to achieve satisfactory performance levels). The fault diagnosis task appears to be the most complex task in the simulation environment because it involves the storage and retrieval of complex information (e.g. fault symptoms and step-by-step procedures for fault recovery), requiring a substantial amount of training. Being perhaps of only slightly lesser complexity, the system control task represents a tracking task comprising five interdependent tracking loops, with particularly high information-processing demands arising during the management of CPF. By contrast, both secondary tasks are not inherently complex as training requirements are minimal for both of them (i.e. operators can successfully complete the annunciator acknowledgement task as well as the logging task after a few minutes of practice). This is not to say that these tasks are undemanding since they are to be completed as part of a multiple-task environment requiring careful allocation of attentional resources. In particular, the prospective memory task is characterized by heavy working memory demands. However, due to the low complexity of the secondary tasks, the team is unlikely to benefit from cognitive diversity during their completion. These findings correspond to other work, which suggests that complex tasks benefit more from cognitive diversity than simple tasks (e.g. Mohammed and Ringseis 2001).

The differences between primary and secondary tasks may also be demonstrated by referring to the classification system of Straus and McGrath (1994). The two primary tasks embedded in CAMS may be considered intellective tasks; that is, they require the identification of the correct solution to a problem (e.g. which sequence of control actions needs to be carried out to stabilize the CPF). By contrast, the two secondary tasks have a strong focus on the implementation of routine action sequences ('execute' in the model of Straus and McGrath). The model may also help make a prediction for the effects of diversity in system understanding when even more complex tasks are to be completed. While there have been effects of cognitive diversity for intellective tasks, one may speculate that for even more complex tasks, such as judgements tasks, the benefits of cognitive diversity would be even greater. This is because judgement tasks do not have a correct answer to a problem (e.g. whether a nuclear power plant be shut down after an accident). Therefore, beliefs, attitudes and values play a much stronger role in the decision-making process, with the discussions gaining from more diverse views within the team.

In contrast to some research that found increasing social conflict to be one of the negative side-effects of diversity (Jehn *et al.* 1997, Pelled *et al.* 1999), cognitive diversity was no cause for conflict in this study, as shown by the data from the conflict questionnaire FAKT and the videotape. The absence of conflict may be partly due to the kind of tasks embedded in the CAMS environment. The factors typically observed as causes for conflict in teams (e.g. different views on the best task management approach; Jehn and Mannix 2001) are more likely to have an effect on judgement tasks than on intellective or execute tasks, which were the two types of tasks embedded in CAMS. The type of tasks used in the present study may therefore have contributed to the very low prevalence of conflict. A second possible cause for the low prevalence of conflict referred to system management strategies used by teams. Any idea generated by a TM to deal with an unknown fault state was usually put to the test without delay to check whether it was effective, suggesting the prevalence of an action-oriented system management approach. If the tested plan of action turned out to be ineffective, a test of an alternative plan of action could immediately follow. In a more reflective system management approach a number of ideas would have been collected first, followed by an evaluation of their effectiveness before the most promising one would have been implemented. The reflective approach is expected to create more conflict within teams. However, this approach was hardly used because there was a need for rather rapid decision-making processes, which was due to the system state of CAMS constantly changing (even in the absence of operator intervention).

While there was no effect of diversity on conflict, some differences in communication patterns were found between groups. Members of SOT-teams communicated more frequently with each other than members of MIX- and POT-teams. This may be because members of SOT-teams were encouraged during their training to think about the system and its components. This more thorough thinking process, if done aloud in a team, may have increased the communication rate but, with communication being a two-way process, only when both TMs had received system-based training. This may explain why the communication rate dropped during the presence of two or even one POT-TM.

Interestingly, within diverse teams there was a rather low awareness of their diversity, as the data collected by the team evaluation questionnaire had demonstrated. This was supported by the findings from the video analysis, which showed that within the diverse teams, the different training methods were very rarely the subject of team

communication. Both data sources, questionnaire and video analysis, provided evidence for a poor transactive memory within these teams, which refers to the level of knowledge of TMs about their within-team resources (see Wegner 1986). This poor transactive memory has probably contributed to specialized teams not making effective use of their expertise, as they may not have assigned the task to the most effective TM. A prerequisite for effective team performance is that the profile of strengths and weaknesses for each TM are known within the group (e.g. Cannon-Bowers *et al.* 1993). A further prerequisite is that the task is carried out by the more competent TM to make the best use of team skills (i.e. the less competent TM must not dominate the team).

The quality of the transactive memory is influenced by the duration of shared task involvement of the team. Because of TMs having worked together for only 1 h in the present study, there has only been very limited time for gaining a good knowledge of overall group resources. This may not be a problem in some technical work environments, where roles are very clearly defined, so that such adaptation processes are not really needed (McGrath *et al.* 1995). However, in those technical work environments in which a substantial transactive memory is required, one needs to be aware that it takes generally more time to develop a shared mental model in teams operating technical systems than in other work groups, such as task forces and management teams (McGrath *et al.* 1995).

The study also provided some interesting findings concerning the effects of teamwork on subjective operator state. A surprising result was that the functional role assigned to the operator had a great impact on their subjective state. The purpose of assigning these functional roles was to balance out the influence of TMs during task completion while it was emphasized that participants should work as equal TMs. While team effectiveness models acknowledge the important part that TM roles play for overall team performance (Tannenbaum *et al.* 1992), this result shows that even temporary minor roles (e.g. having strict responsibilities assigned in task completion to avoid diffusion of responsibility) may substantially increase operator strain.

In addition to subjective operator state being affected by functional roles, ratings of all three state variables (fatigue, effort and anxiety) were found to increase as the working session progressed. Rising levels of fatigue as a result of increasing time-on-task were not unexpected and have been found in comparative studies (e.g. Sauer *et al.* 2003). Somewhat rarer were increases in mental effort expenditure to respond to augmenting work demands since the strategy is typically only applied to cope with short-term peak demands, as an extensive use of that strategy may risk a complete depletion of the operator's resource pool (Hockey 1997). In the present case, demands associated with the coordination of teamwork may have contributed to the observed increase in effort expenditure. Most surprising was, however, that anxiety levels increased in the present study, although in comparative studies examining single operator performance no such increase in anxiety was measured (e.g. Sauer *et al.* 2002). The increasing anxiety may be attributed to the presence of a second operator and may be interpreted as evaluation apprehension within the context of social facilitation theory (Cottrell *et al.* 1968). As performance levels could only be maintained by increasing effort expenditure, TMs may have felt more anxious about successful task completion because their performance was monitored by their team mate.

The article closes with a discussion of the issue of cognitive diversity in a broader context. The present study has extended the focus of previous work by including aspects of cognitive diversity that refer to permanent features of a team's mental model rather

than temporary features. This is an important point since it may dissolve the seeming contradiction between the research literature that advocates that mental models need to be strongly shared and the findings of the present study arguing for the benefits of diversity in system understanding. On the basis of the benefits found for diversity, it could be argued that the potential of cognitive diversity may be even larger under more favourable circumstances. There may be three factors that help create more favourable conditions for cognitively diverse teams. Two of them (a good transactive memory and sufficiently complex tasks) have already been discussed but a third, of no lesser importance, refers to the degree of interdependence between TMs for achieving task goals. TM interdependence was low in this study because each TM would have been able to carry out all given tasks alone (albeit at the cost of excessive workload). If there had been a higher level of interdependency between TMs (such as in distributed systems), one would expect to find an increase in the advantages of cognitive diversity. Overall, cognitive diversity appeared to be a promising concept to convey the idea of team diversification to achieve performance improvements.

Acknowledgements

References

ADOLPH, L., 2000, *Soziale Konflikte bei verschiedenen Formen industrieller Gruppenarbeit und ihre Auswirkungen* (Frankfurt am Main: Lang).

BANKS, A.P. and MILLWARD, L.J., 2000, Running shared mental models as a distributed cognitive process. *British Journal of Psychology*, **91,** 513–531.

CANNON-BOWERS, J.A., SALAS, E. and CONVERSE, S., 1993, Shared mental models in expert team decision making. In *Individual and Group Decision Making: Current Issues*, N.J. Castellan (Ed.), pp. 221–246 (Hillsdale, NJ: Lawrence).

COOKE, N.J., KIEKEL, P.A., SALAS, E., STOUT, R., BOWERS, C. and CANNON-BOWERS, J., 2003, Measuring team knowledge: A window to the cognitive underpinnings of team performance. *Group Dynamics: Theory, Research, and Practice*, **7,** 179–199.

CORDERY, J., 2002, Team working. In *Psychology at Work*, 5th ed., P. Warr (Ed.), pp. 326–350 (London: Penguin Books).

COTTRELL, N.B., WACK, D.L., SEKERAK, G.J. and RITTLE, R.M., 1968, Social facilitation of dominant responses by the presence of an audience and the mere presence of others. *Journal of Personality and Social Psychology*, **9,** 245–250.

HOCKEY, G.R.J., 1997, Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology*, **45,** 73–93.

HOCKEY, G.R.J., SAUER, J. and WASTELL, D.G. (in press), Adaptability of training in simulated process control: comparison of knowledge- and rule-based guidance under task changes and environmental stress. *Human Factors*.

JACKSON, S.E., 1996, The consequences of diversity in multidisciplinary work teams. In *Handbook of Work Group Psychology*, M.A. West (Ed.), pp. 53–75 (Chichester: Wiley and Sons).

JACKSON, S.E., MAY, K.E. and WHITNEY, K., 1995, Understanding the dynamics of diversity in decision-making teams. In *Team Effectiveness and Decision Making in Organizations*, R.A. Guzzo and E. Salas (Eds.), pp. 204–261 (San Francisco: Jossey-Bass).

JEHN, K.A., CHADWICK, C. and THATCHER, S.M.B., 1997, To agree or not to agree: The effects of value congruence, individual demographic dissimilarity, and conflict on work group outcomes. *The International Journal of Conflict Management*, **8,** 287–305.

JEHN, K.A. and MANNIX, E.A., 2001, The dynamic nature of conflict: A longitudinal study of intragroup conflict and group performance. *Academy of Management Journal*, **44,** 238–251.

MCGRATH, J.E., BERDAHL, J.L. and ARROW, H., 1995, Traits, expectations, culture, and clout: The dynamics of diversity in work groups. In *Diversity in Work Teams, Research Paradigms for a Changing Workplace*, S.E. Jackson and M.N. Ruderman (Eds.), pp. 17–47 (Washington: American Psychological Association).

MAYRING, P., 1988, *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (Weinheim: Deutscher Studien Verlag).

MILLIKEN, F.J. and MARTINS, L.L., 1996, Searching for common threads: understanding the multiple effects of diversity in organizational groups. *Academy of Management Review*, **21,** 402–433.

MOHAMMED, S. and RINGSEIS, E., 2001, Cognitive diversity and consensus in group decision making: The role of inputs, processes, and outcomes. *Organizational Behavior and Human Decision Processes*, **85,** 310–335.

MORRIS, N.M. and ROUSE, W.B., 1985, Review and evaluation of empirical research in troubleshooting. *Human Factors*, **27,** 503–530.

PATRICK, J., 1992, *Training: Research and Practice* (London: Academic Press).

PELLED, L.H., EISENHARDT, K.M. and XIN, K.R., 1999, Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly*, **44,** 1–28.

PRIEM, R.L., HARRISON, D.A. and MUIR, N.K., 1995, Structured conflict and consensus outcomes in group decision making. *Journal of Management*, **21,** 691–710.

ROUSE, W.B. and MORRIS, N.M., 1986, On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, **100,** 349–363.

SAUER, J., HOCKEY, G.R.J. and WASTELL, D.G., 2000a, Effects of training on short- and long-term skill retention in a complex multiple-task environment. *Ergonomics*, **43,** 2043–2064.

SAUER, J., WASTELL, D. and HOCKEY, G.R.J., 2000b, Using micro-worlds to simulate highly-automated work environments: the case of the Cabin Air Management System. *Computers in Human Behavior*, **16,** 45–58.

SAUER, J., WASTELL, D., HOCKEY, G.R.J., CRAWSHAW, M., ISHAK, M. and DOWNING, J., 2002, Effects of display design on performance in a simulated ship navigation environment. *Ergonomics*, **45,** 329–347.

SAUER, J., WASTELL, D., HOCKEY, G.R.J. and EARLE, F., 2003, Performance in a complex multiple-task environment during a lab-based simulation of occasional night work. *Human Factors*, **45,** 657–669.

SHAW, J.B. and BARRETT-POWER, E., 1998, The effects of diversity on small work group processes and performance. *Human Relations*, **51,** 1307–1325.

STEINER, I.D., 1972, *Group Processes and Productivity* (New York: Academic Press).

STOUT, R.J., CANNON-BOWERS, J.A., SALAS, E. and MILANOVICH, D.M., 1999, Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors*, **41,** 61–71.

STRAUS, S.G. and MCGRATH, J.E., 1994, Does the medium matter? The interaction of task type and technology on group performance and member reactions. *Journal of Applied Psychology*, **79,** 87–97.

STUMPF, S. and THOMAS, A., 1999, How to cope with the diversity of diversity research on group effectiveness: some outlines for future empirical research, theory building and practical applications. *Psychologische Beiträge*, **41,** 403–421.

TANNENBAUM, S.I., BEARD, R.L. and SALAS, E., 1992, Team building and its influence on team effectiveness: an examination of conceptual and empirical developments. In *Issues, Theory, and Research in Industrial/ Organizational Psychology*, K. Kelley (Ed.), pp. 117–153 (New York: Elsevier Science).

TSCHAN, F., SEMMER, N.K., NÄGELE, C. and GURTNER, A., 2000, Task adaptive behavior and performance in groups. *Group Processes & Intergroup Relations*, **3,** 367–386.

WEBBER, S.S. and DONAHUE, L.M., 2001, Impact of highly and less job-related diversity on work group cohesion and performance: a meta-analysis. *Journal of Management*, **27,** 141–162.

WEGNER, D.M., 1986, Transactive memory: A contemporary analysis of the group mind. In *Theories of Group Behaviour*, B. Mullen and G.R. Goethals (Eds.), pp. 185–208 (New York: Springer).

WILLIAMS, K.Y. and O'REILLY, C.A., 1998, Demography and diversity in organizations: A review of 40 years of research. *Organizational Behavior*, **20,** 77–140.

WINDEL, A., KRONZ, E., ADOLPH, L. and ZIMOLONG, B., 1999, Der Fragebogen zu arbeitsbezogenen Konflikten (FAKT): Entwicklung, statistische Kennwerte und Anwendungshinweise. Unpublished questionnaire. Arbeits- und Organisationspsychologie, Universität Bochum.