

An agent-based approach for structured modeling, analysis and improvement of safety culture

Alexei Sharpanskykh · Sybert H. Stroeve

Published online: 12 February 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract Safety culture is broadly recognized as important for operational safety in various fields, including air traffic management, power plant control and health care. Previous studies addressed characterization and assessment of safety culture extensively. Nevertheless, relations between safety culture and formal and informal organizational structures and processes are yet not well understood. To address this gap, a new, formal, agent-based approach is proposed. This paper shows the application of the approach to an air navigation service provider, including structured modeling, analysis and identification of improvement strategies for the organizational safety culture. The model results have been validated using safety culture data that had been achieved by an independent safety culture survey study.

Keywords Safety culture · Agent-based organizational modeling · Social simulation · Model analysis · Model validation

This contribution is based on a published conference paper (Sharpanskykh and Stroeve 2010), which was extended and revised significantly. There exists a companion paper (Stroeve et al. 2011) targeted to a safety science audience, which focuses on coupling of the obtained model results with results of a survey-based study and on validation.

Electronic supplementary material The online version of this article (doi:10.1007/s10588-011-9083-9) contains supplementary material, which is available to authorized users.

A. Sharpanskykh (✉)

Department of Artificial Intelligence, Vrije Universiteit Amsterdam, De Boelelaan 1081a,
1081 HV Amsterdam, The Netherlands
e-mail: sharp@cs.vu.nl

S.H. Stroeve

National Aerospace Laboratory NLR, Air Transport Safety Institute, Amsterdam, The Netherlands
e-mail: stroeve@nlr.nl

1 Introduction

Organizational and safety culture are broadly recognized as important for operational safety in various fields, including air traffic management, power plant control and health care. For example, as indicated in Leveson et al. (2005), investigations into many NASA mission failures pointed to cultural problems and the need for cultural and organizational improvements. Currently, as a prelude to systemic changes in air traffic management via innovation programmes SESAR in Europe and NextGen in the USA, more and more Air Navigation Service Providers (ANSP's) go through safety culture measurement and improvement processes.

The main aspects of organizational culture are reflected in a definition by Uttal (1983): 'Shared values (what is important) and beliefs (how things work) that interact with an organization's structures and control systems to produce behavioral norms (the way we do things around here).' There exists a variety of definitions of safety culture. Most of them encompass five components: (1) informed culture that collates data from accidents and incidents and combines them with information from proactive measures (e.g., safety audits); (2) reporting culture in which employees feel free to report safety-related occurrences; (3) just culture characterized by an atmosphere of trust; (4) flexible culture that successfully manages safety during organizational changes; (5) learning culture needed to draw conclusions from the information collected along with the will to implement necessary changes. The term *safety culture* is used in this paper as those aspects of organizational culture that may have an effect on safety, in line with reasoning of Hopkins (2006).

Various studies focused on characterization of safety culture and on assessment of safety culture of various organizations, including ANSPs (cf. Ek et al. 2007). However, the links of safety culture with organizational structures and processes are yet not well understood and this affects the determination of ways to improve safety culture. Traditional approaches to safety analysis (Bedford and Cooke 2001) focus on failure events and human errors and put events into chains or trees, which are used for sequential cause-effect reasoning for accident causation. However, such trees do not account for complex, non-linear dependencies and dynamics inherent in ANSPs and, therefore, are not well suited for safety culture analysis. More promising approaches for modeling and analysis of safety culture using system dynamics were proposed in Leveson et al. (2005) and in Mohaghegh and Mosleh (2009). These approaches abstract from single events and actors and take an aggregate view on the organizational dynamics. Consequently, in safety culture models aggregated values are assigned to such highly individual variables as 'fear of reporting', 'employee participation', 'perceived risk'. By taking such an aggregated view, important effects of individual differences on local interaction and global organizational dynamics may be overlooked.

A number of agent-based approaches for control, coordination, optimization and safety assessment of operations in air traffic have been proposed in the area of multi-agent systems. In particular in Tumer and Agogino (2007) a distributed agent-based air traffic flow management system is described. Through reinforcement learning this system ensures observance of separation safety norms by aircraft in a simulated airspace. Another agent-based system described in Wolfe (2007) is used to support

air traffic controllers in their decision making in air traffic control tasks. In Lee et al. (2005) a cognitive agent model for an air traffic controller is proposed, used for analysis of different working modes through simulation. In Blom et al. (2006) stochastic dynamic agent-based models are used to describe the performance of people and systems in air traffic scenarios, addressing their performance in both nominal and off-nominal situations. Monte Carlo simulations of these models provide rare event probability estimates of safety-critical scenarios, including the accident risk of air traffic operations.

However, above agent-based approaches do not address the organizational layer and interaction between a formal organization and autonomous behavior of organizational agents explicitly. Furthermore, these systems do not focus on organizational and safety culture modeling and analysis.

Influences of national culture and some aspects of the organizational culture on project team performance were investigated in Horii et al. (2004). The authors also provide an extensive model of the organizational context based on the Virtual Design Team framework Kunz et al. (1998) in this study. Safety aspects are mentioned only briefly in this study and influence of aspects of national culture on safety culture is not considered.

An agent-based approach proposed in this paper focuses on modeling and analysis of safety culture in particular and addresses an existing gap between safety culture and organizational structures and processes. It provides a formal basis for understanding the causal relations between organizational processes that influence safety culture, such that robust and flexible policies may be identified to improve and maintain a sufficient level of safety culture in an organization.

This paper demonstrates the application of the approach to an air navigation service provider, including structured modeling, analysis and identification of improvement strategies for the organizational safety culture. The development of the organizational model has been focused on safety occurrence reporting at a real ANSP (called ANSP3 throughout the paper) using an existing formal organization modeling framework (Sharpanykh 2008) and data provided by ANSP3. The reporting of safety-related occurrences is obligatory in Air Traffic Management (ATM) and it is an important constituent of the Safety Management System (SMS) of ANSPs. It is recognized that there is a strong reciprocal relation between the organizational safety culture and the reporting behavior of air traffic controllers (Ek et al. 2007). In the model proposed this relation is elaborated formally, in detail.

As a basis for the validation, results of a safety culture survey questionnaire for ANSP3 and of the related workshops administered by EUROCONTROL were used. This safety culture survey approach has been applied at a large number of ANSPs in Europe. The survey questionnaire data obtained in these applications were analyzed statistically to ensure that the questions appropriately address the key aspects of safety culture (EUROCONTROL/FAA AP15 2008).

Thus, the main contributions of the paper are:

- (1) A formal, agent-based organizational model for ANSP3 developed by a structured methodology; the model addresses both prescriptive aspects of the formal organization and autonomous behavior of agents.

- (2) Identification of explicit, formal relations between safety culture indicators and the agent-based organizational model.
- (3) Analysis methods for structured identification of safety culture issues and safety culture improvement options.
- (4) Validation of the obtained model results.

The paper is organized as follows. In Sect. 2 a case study is introduced. An overview of the modeling approach is provided in Sect. 3. Modeling of the formal organization from the case study is described in Sect. 4. Modeling of agent-specific aspects for the case study is described in Sect. 5. Furthermore, it is described in Sect. 5 how safety culture indicators identified in the paper can be related to the model. The developed method for safety culture analysis is presented in Sect. 6. Also model validation is considered in Sect. 6. Finally, Sect. 7 concludes the paper.

2 Case study

In operations performed by air traffic controllers hazardous situations may occur, e.g., separation minima infringement, runway incursion, aircraft deviation from ATC clearance. According to air traffic safety regulations, controllers are obliged to report safety occurrences of a large number of predefined types. In particular, information about safety occurrences is useful for safety analysis (e.g., identification of safety trends). Some ANSPs reprimand controllers for occurrences. Sometimes also rewards are provided for reporting occurrences. Although reporting is obligatory, it is believed that not all identified occurrences are properly reported by controllers. To understand the reasons for such behavior, the occurrence reporting cycle including the controller's decision making whether to report an occurrence is modeled in this study in the context of an existing ANSP.

Controllers work in shifts, within each shift a pair of controllers is allocated to each air traffic control sector. A controller supervisor manages work in a shift. After a controller decides to report an observed occurrence, s/he creates a notification report. The draft report is reviewed and maybe corrected by the supervisor and it is provided to the Safety Investigation Unit of the ANSP. Depending on the occurrence severity and the collected information about (similar) occurrences, the Safety Investigation Unit makes the decision whether to initiate a detailed investigation. During the investigation accumulated organizational knowledge and data about safety related issues (in particular, learned from notification reports) is used. As the investigation result, a final occurrence assessment report is produced, which should be provided to the controller-reporter as a feedback. Furthermore, often final reports contain recommendations for safety improvement, which are required to be implemented by ANSP (e.g., provision of training, improvement of formal procedures, extension of staff).

3 Modeling approach: overview

In this section an overview of the agent-based modeling approach is provided. First, modeling requirements are discussed in Sect. 3.1. Then, the organization modeling

framework used is considered in Sect. 3.2. After that, the formal modeling language used for the model specification is described briefly in Sect. 3.3.

3.1 Modeling requirements

To identify safety culture aspects relevant for the safety occurrence reporting, safety culture survey results of two ANSPs (different from ANSP3) and safety culture data from the literature were analyzed and interviews were conducted with experts at EUROCONTROL Head Quarters and at ANSP3. As result of this analysis, a categorized set of safety culture issues that impact safety occurrence reporting was determined; some examples are given in Table 1 and a full list is specified in (supplementary material).

For each issue from the set, required organization modeling aspects have been identified. For example, for S1.9 the following modeling aspects were identified: ‘Cognitive and psychological characteristics’, ‘Generation of beliefs on safety-related information’, ‘Influence of psychological, cognitive and other individual characteristics on (the attitude towards) reporting’; for S2.3: ‘Trust to a peer’, ‘Team composition’, ‘Collaboration in a team and its influence on safety’. Based on the three criteria—importance for modeling of safety occurrence reporting, availability of data, maturity level of modeling techniques—the selection of the most relevant modeling aspects has been performed for further inclusion in the model. More details on identification of safety culture issues and modeling aspects are given in (supplementary material).

Furthermore, based on the identified set of safety culture issues, eight safety culture indicators were identified. These indicators were used in our study for the evaluation of the quality of the ANSP’s safety culture in relation to occurrence reporting.

Table 1 Examples of the identified safety culture issues

Group 1: Individual aspects

- S1.1: Occurrence reporting may lead to ‘naming and blaming’ and therefore it may not be in the personal interest of an actor
- S1.9: The confidentiality of reporting is not trusted
- S1.2: Actors are not motivated to report their safety concerns because of the lack of feedback and interest experienced in the organization

Group 2: Team aspects

- S2.3: Willingness of actors to cooperate with an actor may decrease after s/he has been involved in a (serious) incident
- S2.6: Problems are not raised as actors do not want to be seen as trouble-makers
- S2.9: Supervisors may not effectively reinforce safety culture

Group 3: Intra-organizational aspects

- S3.1: Importance of safety-related goals may be threatened by performance-related goals
 - S3.17: Feedback/lessons learned from incidents comes too late or not at all
 - S3.13: Information about changes in procedures and in the system is not provided on time to (some) actors that require this information
-

The values of these indicators are determined by simulation based on the ANSP3's model described in this paper. These values are used as input for the validation of the approach proposed.

SCI1: Average reporting quality of controllers. It refers to the ratio of reported to observed occurrences.

SCI2: Average quality of the processed notification reports. It refers to the correctness and completeness of information about the reported occurrences.

SCI3: Average quality of the final safety occurrence assessment reports. It refers to the completeness of the occurrence report with respect to the causes of the occurrence.

SCI4: Average quality of the monthly safety overview reports received by controllers. It refers to the completeness of the report with respect to the safety trends.

SCI5.1: Average commitment to safety of controllers.

SCI5.2: Average commitment to safety of a team as perceived by controllers.

SCI6: Average commitment to safety of a supervisor as perceived by controllers.

SCI7: Average commitment to safety of management as perceived by controllers.

The range of each indicator is $[0, 1]$.

3.2 Organization modeling framework

For modeling safety occurrence reporting in ANSP3 the organization modeling framework from Sharpanskykh (2008) was used. In this framework organizations are considered from different perspectives (views). *Process-oriented view* describes the workflow as well as static structures of tasks and resources. *Performance-oriented view* is characterized by a goal structure, a performance indicators structure, and relations between them as well as relations between goals and tasks, performance indicators and processes, goals and roles or agents. *Organization-oriented view* defines the organizational roles, each associated with a set of tasks and characterized by authority and responsibility relations on tasks, resources and information. Commitment, obligation and power relations and sets of competences required for agent allocation to roles are also defined. *Agent-oriented view* identifies different types of agents with their capabilities and behavior.

Furthermore, this framework describes a sequence of organization design steps, an overview of which is provided in Table 2. This framework includes all the modeling aspects related to the formal organization of ANSP3, which were identified as required to be included in the model. Modeling of the formal organization from the case study is described in Sect. 4. The framework does not prescribe any specific architecture or approach for modeling of agents. Modeling of agent-specific aspects for the case study considered in the paper is described in Sect. 5.

3.3 Formal modeling language

To specify organizational dynamic properties, the order-sorted predicate logic-based language called LEADSTO is used (Bosse et al. 2007). Dynamics in LEADSTO is represented as evolution of states over time. A state is characterised by a set of properties that do or do not hold at a certain point in time. To specify state properties for

Table 2 Overview of steps in organizational modeling and their relation with the views considered

Design step	View			
	Organization	Performance	Process	Agent
Identification of organizational roles	x			
Identification of interactions between roles and with the environment	x			
Identification of requirements for roles	x			x
Identification of organizational performance indicators and goals		x		
Identification of resources			x	
Identification of organizational tasks and relations between tasks, resources and goals		x	x	
Identification of authority relations	x		x	
Identification of workflows			x	
Identification of characteristics (skills, psychological and cognitive characteristics) of agents		x		x
Identification of goals and needs of agents				x
Identification of commitments, obligations and responsibilities of agents				x
Identification of attitudes and beliefs of agents				x
Identification of relations between agents (e.g., interaction, trust and informal power relations) and informal structures of agents	x			x
Identification of shared beliefs, attitudes, norms and values of (groups of) agents				x
Identification of performance variability in formal and informal flows of control			x	x
Allocation principles of agents to organizational roles				x
Identification of organizational constraints	x	x	x	x
Specification of the environmental dynamics				x

components of an organization, ontologies are used which are defined by a number of sorts, sorted constants, variables, functions and predicates (i.e., a signature or a vocabulary). For every component A a number of ontologies can be distinguished: the ontologies $IntOnt(A)$, $InOnt(A)$, $OutOnt(A)$, and $ExtOnt(A)$ are used to express respectively internal, input, output and external state properties of the component A. Input ontologies contain elements for describing perceptions of a component (e.g., organizational actor) from the external world.

In particular, to specify observation requests of components in the environment, function $to_be_observed: STATE_PROPERTY \rightarrow STATPROP$ is used. Observation results are provided to the component’s input using the function $observation_result: STATE_PROPERTY \rightarrow STATPROP$, which indicates whether the state property holds in the environment. For example, $observation_result(occurrence(o, ot))$ specifies the observation of occurrence o of type ot.

Output ontologies describe actions and communications of components. To specify communications the following function is included in the output ontologies:

communicated_from_to: COMPONENT \times COMPONENT \times MSG_TYPE \times CONTENT \rightarrow STATPROP

Here the first argument denotes the component-source of information, the second—the component-recipient of information, the third argument denotes the types of the communication (which may be one of the following {observe, inform, request, decision, feedback}) and the fourth—the content of the communication. The sort COMPONENT is a composite sort that comprises all subsorts of the components of particular types (e.g., CONTROLLER). The sort CONTENT is also the composite sort that comprises all names of terms that are used as the communication content. Such terms are constructed from sorted constants, variables and functions in the standard predicate logic way.

For example, communicated_from_to(ag_controller, ag_supervisor, inform, notification_report_for_B)) specifies communication of notification_report_for_B from ag_controller to ag_supervisor.

Actions of a component generated at its input interface are formalized by function performed: ACTION \rightarrow STATPROP. For example, performed(action(initiate_reporting(o, ot))) specifies the action to initiate reporting about occurrence o of type ot.

Internal ontologies describe internal states of components. The modelling language allows specifying a large variety of cognitive architectures (Bosse et al. 2007). For example, to specify beliefs, which form a part of many existing cognitive models, function belief: STATPROP \times TIME \rightarrow STATPROP may be included in the internal ontology of a component. Sort TIME contains a set of linearly ordered time points. For example, belief(occurrence_A, 10) specifies the belief that occurrence_A was observed at time point 10.

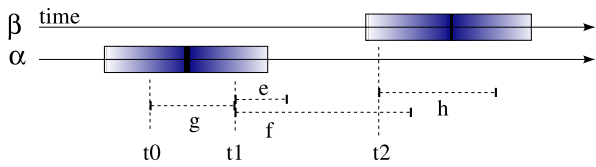
External ontologies are used to describe states of the environment. For example, to specify safety occurrences in the environment function occurrence: OCCUR \times OCCUR_TYPE can be used, where sort OCCUR contains all names of the occurrences, OCCUR_TYPE contains all names of the occurrence types.

For a given ontology Ont, the propositional language signature consisting of all state ground atoms based on Ont is denoted by APROP(Ont). State properties are specified based on such ontology by propositions that can be made (using conjunction, negation, disjunction, implication) from the ground atoms. Then, a state S is an indication of which atomic state properties are true and which are false: S: APROP(Ont) \rightarrow {true, false}.

LEADSTO enables modeling of direct temporal dependencies between two state properties in successive states, also called *dynamic properties*. A specification of dynamic properties in LEADSTO is executable and can be depicted graphically. The format is defined as follows. Let α and β be state properties of the form ‘conjunction of atoms or negations of atoms’, and e, f, g, h non-negative real numbers. In the LEADSTO language the notation $\alpha \rightarrow_{e,f,g,h} \beta$ means: if state property α holds for a certain time interval with duration g, then after some delay (between e and f) state property β will hold for a certain time interval of length h (see Fig. 1).

When e = f = 0 and g = h = 1, called standard time parameters, we write $\alpha \rightarrow \beta$. LEADSTO specifications can be translated automatically to a general state transition system format as described in Sharpanskykh and Treur (2010). Such translated

Fig. 1 The timing relationships for a LEADSTO expression



specifications can be executed by a large variety of tools, including Matlab, MetateM (Fisher 1996), and automata-based simulation environments.

Consider an example dynamic property for a controller agent:

$$\text{observation_result}(\text{occurrence}(o, ot)) \rightarrow \text{performed}(\text{action}(\text{initiate_reporting}(o, ot)))$$

Informally, this example expresses that if the agent observes occurrence o of type ot , then after that the agent will initiate reporting of the occurrence during 1 time unit.

In addition, LEADSTO allows expressing mathematical operations. To this end functional symbols $-, +, /, \bullet : \text{VALUE} \times \text{VALUE} \rightarrow \text{VALUE}$ are introduced, here sort VALUE is an ordered set of numbers. In the model description in Sects. 4 and 5 for better readability mathematical expressions will be used instead of logical ones, e.g. the rule $\text{has_value}(x, v) \rightarrow \text{has_value}(x, v * 0.25)$ is specified by the mathematical expression $x_t = x_{t-1} * 0.25$.

4 Modeling formal organization

A model for the formal organization from the case study was built along the design steps described in Sharpanskykh (2008).

4.1 The identification of the organizational roles

A *role* is a (sub-)set of functionalities of an organization, which are abstracted from specific agents who fulfill them. Each role can be composed by several other roles, until the necessary detailed level of aggregation is achieved. A role composed of (interacting) subroles is called a composite role. A formal or informal organizational group is modeled as a composite role. Grouping of roles in a composite role may be done by knowledge and skill, by work process and function, by time, or by place. At a higher abstraction level, sets of simple roles clustered in a composite role act as a single entity interacting with other (composite) roles. Each role has an input and an output interface, which facilitate interaction with other roles. Interfaces are formalized using ontologies, defined by a number of sorts, sorted constants, variables, functions and predicates (i.e., signatures). An input ontology specifies which types of information can be provided to the input interface of a role, and an output ontology determines which information types can be generated by a role. Furthermore, using input/output interfaces information abstraction (or transformation) can be performed. For example, a composite role may represent a department meeting at which the progress of particular students is discussed. An output of this role to the composite role representing Managerial Board may be the total number of students not making satisfactory progress.

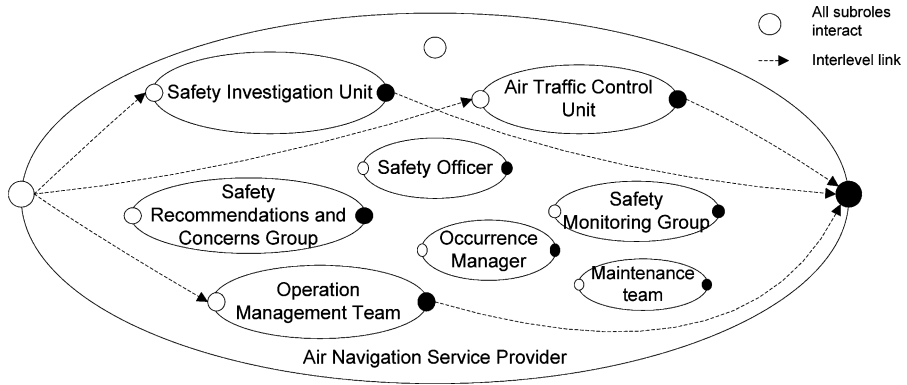


Fig. 2 The interaction relations between the subroles of the role Air Navigation Service Provider at the aggregation level 2

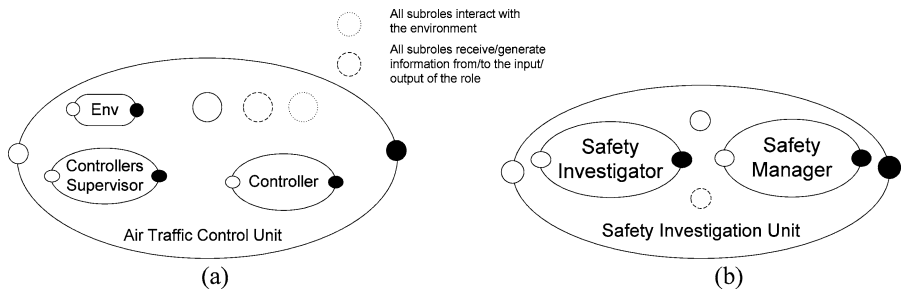


Fig. 3 The interaction relations between the subroles of the role Air Traffic Control Unit (a) and of the role Safety Investigation Unit (b) at the aggregation level 3

The environment represents a special component of a model, which also has input and output interfaces. In the case study roles are identified at three aggregation levels. At the aggregation level 1 the Air Navigation Service Provider is considered as one composite role. The subroles of the Air Navigation Service Provider are described at the aggregation level 2 (see Fig. 2). Two ANSP’s roles are refined further, as shown at aggregation level 3 in Fig. 3.

4.2 The specification of the interactions between the roles

Relations between roles are represented by interaction and interlevel links. An *interaction link* is an information channel between two roles at the same aggregation level. An *interlevel link* connects a composite role with one of its subroles. The interaction relations at the ANSP’s aggregation levels 2 and 3 are shown in Fig. 2 and Fig. 3.

4.3 The identification of the requirements for the roles

In this step the requirements on knowledge, skills and personal traits of an agent implementing a role are identified. A prerequisite for the allocation of an agent to a role

is the existence of a mapping between the capabilities and traits of the agent and the role requirements. For example, among the requirements for the controller role are: passed a rigid medical examination; thorough knowledge of the flight regulations; air traffic control training; excellent listening and communication skills; quick decision-making skills. Note that the measure of the level of development may be associated with some capabilities. For example, controllers may differ in the amount of hours that they spent for air traffic control training.

4.4 The identification of the organizational performance indicators and goals

A performance indicator (PI) is a quantitative or qualitative indicator that reflects the state/progress of the company, unit or individual. PIs can be hard (e.g., occurrence investigation time) or soft, i.e., not directly measurable, qualitative (e.g., level of collaboration between controllers). *Goals* are objectives that describe a desired state or development and are defined formally as expressions over PIs. The characteristics of a goal include, among others: *priority*; *horizon*—for which time point/interval the goal should be satisfied; *hardness*—hard or soft. A goal can be refined into subgoals forming a hierarchy. For example, goal G18 ‘It is required to maintain timeliness and a high quality of occurrence investigation’ is based on two PIs ‘timeliness of occurrence investigation’ and ‘quality of occurrence investigation’. This goal is refined in several subgoals: G18.1 ‘It is required to maintain a high proficiency level of incident investigators’, G18.2 ‘It is required to maintain a sufficient level of details of notification reports’, G18.3 ‘It is required to maintain the timely investigation of an occurrence’ and G18.4 ‘It is required to maintain a high level of thoroughness of occurrence investigation’. Formally, $is_refined_to(G18, L) \ \& \ is_in_goal_list(G18.1, L) \ \& \ is_in_goal_list(G18.2, L) \ \& \ is_in_goal_list(G18.3, L) \ \& \ is_in_goal_list(G18.4, L)$. To ensure the satisfaction of G18, the (sufficient degree of) satisfaction of its subroles is required. Goals are related to roles. E.g., G18 is attributed to Safety Investigation Unit role of the ANSP. Formally, $is_committed_to(Safety \ Investigation \ Unit, G18)$.

4.5 The specification of the resources

Resource types are characterized by: *name*, *category*: discrete or continuous, *measurement unit*, *expiration duration*: the time interval during which a resource type can be used; *location*; *sharing*: some processes may share resources. Examples of resource types are: aircraft, incident classification database, a notification report.

4.6 The identification of the tasks and relations between the tasks, the resources and the goals

A task represents a function performed in the organization and is characterized by *name*, *maximal* and *minimal duration*. Tasks can be decomposed into more specific ones using AND- and OR-relations. Each task performed in an organization should contribute to the satisfaction of one or more organizational goals. For example, task T4 ‘Safety occurrence reporting and the report handling’ is refined into more specific tasks, among which T4.1 ‘Create a notification report’, T4.3 ‘Making decision about the investigation necessity’, T4.4 ‘Investigation of an occurrence’.

Formally, $\text{is_decomposed_to}(T4, L) \ \& \ \text{is_in_task_list}(T4.1, L) \ \& \ \text{is_in_task_list}(T4.2, L) \ \& \ \text{is_in_task_list}(T4.3, L) \ \& \ \text{is_in_task_list}(T4.4, L)$. Task T4.4 is related to resources as follows: it uses a notification report and produces a final occurrence assessment report. Formally, $\text{task_uses}(T4.4, \text{notification_report}, 1)$ and $\text{task_produces}(T4.4, \text{final_occurrence_assessment_report}, 1)$. Furthermore, T4.1 contributes to goal G18.2, and T4.4 contributes to goals G18.3 and G18.4. Formally, $\text{is_realisable_by}(G18.2, L) \ \& \ \text{is_in_list}(T4.1, L) \ \& \ \text{is_realisable_by}(G18.3, L1) \ \& \ \text{is_in_list}(T4.4, L1) \ \& \ \text{is_realisable_by}(G18.4, L2) \ \& \ \text{is_in_list}(T4.4, L2)$.

4.7 The specification of the authority relations

The following types of authority relations are distinguished: superior-subordinate relations on roles with respect to tasks, responsibility relations, control for resources, authorization relations. Roles may have different rights and responsibilities with respect to different aspects of task execution, such as execution, passive monitoring, consulting, making technological and managerial decisions. For example, Safety Investigator role is assigned responsible for execution of and making technological decisions with respect to task T4.4, Safety Manager is responsible for monitoring, consulting and making managerial decisions related to T4.4. Formally, $\text{is_authorized_for}(\text{Safety Investigator}, \text{execution}, T4.4) \ \& \ \text{is_authorized_for}(\text{Safety Investigator}, \text{tech_decisions}, T4.4) \ \& \ \text{is_authorized_for}(\text{Safety Manager}, \text{monitoring}, T4.4) \ \& \ \text{is_authorized_for}(\text{Safety Manager}, \text{consulting}, T4.4) \ \& \ \text{is_authorized_for}(\text{Safety Manager}, \text{manag_decisions}, T4.4)$.

4.8 The specification of the workflows

Workflows describe temporal ordering of processes of an organization in particular scenarios. A workflow starts with the process BEGIN and ends with the process END; both have zero duration. The (partial) order of execution of processes in the workflow is defined by sequencing, branching, cycle and synchronization relations (referred to as ordering relations) specified by the language from Popova and Sharpanskykh (2008). A workflow that describes the execution of the safety occurrence reporting and the report handling task is depicted in Fig. 4a.

This workflow is executed every time when an occurrence has been identified by a controller. In the following the workflow is considered briefly. After a controller decides to report an observed occurrence, she creates a notification report, which is provided to the Safety Investigation Unit (SIU). Depending on the occurrence severity and the collected information about similar occurrences, Safety Investigator role in SIU makes the decision whether to initiate a detailed investigation. During the investigation, accumulated organizational data and knowledge about safety related issues is used. As the investigation result, a final occurrence assessment report is produced, which provides feedback to the controller-reporter. Furthermore, often final reports contain recommendations for safety improvement, which are advised to be implemented by the ANSP.

A workflow for the safety monitoring task is depicted in Fig. 4b. This workflow is executed continuously in a cycle, i.e., whenever the current instance of the workflow finishes, a new instance of the workflow starts.

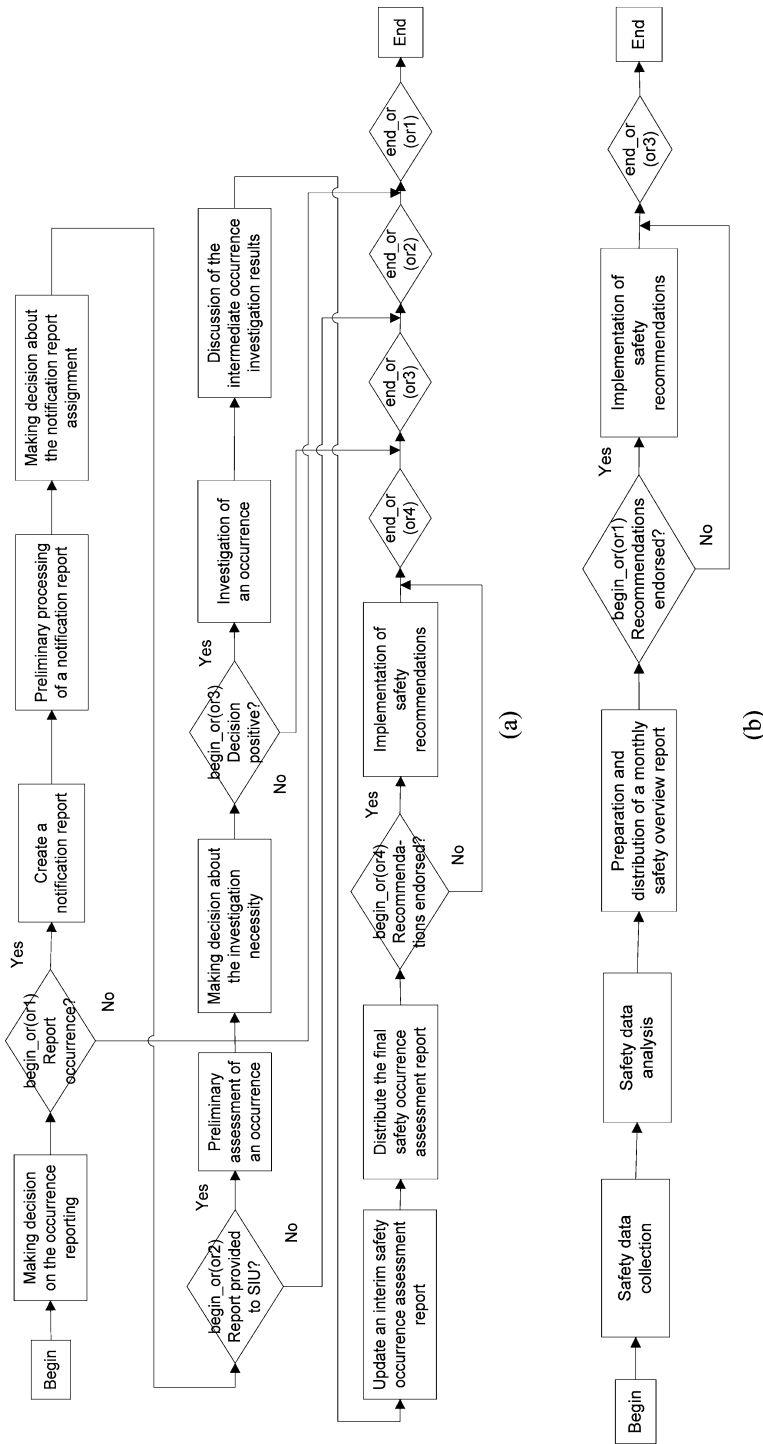


Fig. 4 The workflow (a) defines the execution of the safety occurrence reporting and the report handling task initiated by a controller and the workflow (b) defines the execution of the safety monitoring task; the graphical modeling language used is adopted from Popova and Sharpanskykh (2008)

4.9 The identification of the generic and domain-specific constraints

Generic constraints ensure internal consistency of an organizational specification. Domain specific constraints restrain behavior of individuals in a particular organization. In particular, organizational reward policies can be formalized as domain-specific constraints. For example, the reprimand policy for reporting can be formalized by a set of constraints using a function repr that maps the number of occurrences of some type to a reprimand value $[0, 1]$: $\text{repr}(1, A) = 1$; $\text{repr}(1, B) = 0.5$.

5 Modeling agents

The specification of a formal organization forms a part of an overall organizational description. Another part describes the characteristics and behavior of agents. In this section first a theory of human behavior and cognition that forms a basis for agent models is introduced in Sect. 5.1. The proposed agent architecture realizing the theory is described in Sect. 5.2. Organizational learning based on learning of agents is described in Sect. 5.3. Finally, it is demonstrated in Sect. 5.4 how cognitive and behavioral states of the agents are related formally to safety culture indicators.

5.1 Theory of human behavior and cognition

In Cognitive Science the behavior of a human is often considered from the *external* and *internal perspectives* (Kim 1996). From the external perspective, behavior of the human can be described by correlations between input and output states over time, without any reference to internal or mental properties of the human. Such a view is considered within the philosophical perspective of behaviorism (Kim 1996). A specification of externally observable behavior has the advantage that it can be directly related to empirical information, however, it may have a high complexity. From the internal perspective the behavior of the human can be characterized by a specification of more direct (causal) temporal relations between mental states of the human, based on which an externally observable behavioral pattern is generated. Such a perspective is taken within the functionalist tradition (Kim 1996). The mental states and their direct causal temporal relations can be used for simulation and analysis of the human behavior, but they may well have a hypothetical status, as they are difficult to observe. In this paper we consider both externally observable behavior and internal cognitive processes of a human, which are related to each other in the agent architecture described in Sect. 5.2.

The behavior of a human involves interaction with other humans and with the environment. A human perceives information by *observation* and produces *communication* and *actions*. We distinguish passive and active observation processes. For example, when some object is observable by a human and the human continuously keeps track of its state, passive observation occurs. For passive observation, no initiative of a human is needed. Active observation is always concerned with the human's initiative and a particular observation focus. To represent communication between humans, speech act theory is used in this paper (Searle 1969). This theory allows representing a wide diversity of illocutionary speech acts constituting communication.

Based on psychological theories (e.g., Goldman 1967) it is postulated that a human forms *beliefs* about observation and communication acts received. Besides beliefs about single states, a human forms beliefs about dependencies between states. Beliefs are used in human reasoning and decision making.

It is widely recognized in the literature (Pinder 1998) that the behavior and dynamics of internal processes of a human are influenced by *personal traits* and *capabilities*. In general, one can specify a great variety of personal traits (e.g., the big five model Goldberg 1993). In particular, cultural traits are often recognized as influential for the human behavior (Hofstede 2005). In the case study, cultural traits are specified for each air traffic controller based on the cultural classification framework by Hofstede (2005). More specifically, the following indexes from the framework are used: *individualism* (IDV) is the degree to which individuals are integrated into groups; *power distance index* (PDI) is the extent to which the less powerful members of an organization accept and expect that power is distributed unequally; and *uncertainty avoidance index* (UAI) deals with individual's tolerance for uncertainty and ambiguity. Human capabilities include *knowledge* and *skills*. Knowledge of a human is a set of reasoning procedures, which are used in particular for execution of organizational tasks. Human beliefs may be modified using knowledge. *Skills* describe developed abilities of a human to use effectively and readily his or her knowledge for tasks performance.

In social science behavior of individuals is considered as *goal-driven* (Robbins 2004; Pinder 1998; Locke 2001; van de Walle 1997). It is also recognized that individual goals are based on *needs* (Pinder 1998). Different types of needs are distinguished: (1) *extrinsic needs* associated with biological comfort and material rewards; (2) *social interaction* needs that refer to the desire for social approval and affiliation, in particular own group approval and management approval; (3) *intrinsic needs* that concern the desires for self-development and self-actualization, in particular contribution to organizational safety-related goals and self-esteem, self-confidence and self-actualization needs. Different needs have different priorities and minimal acceptable satisfaction levels for individuals in different cultures.

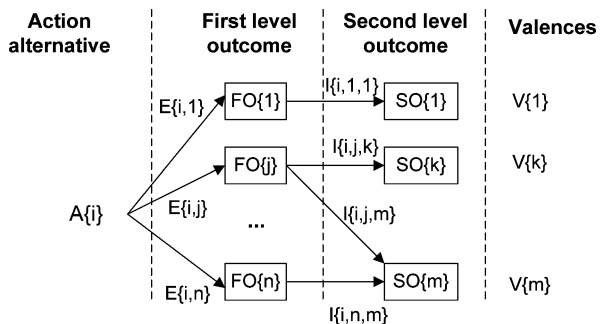
According to the air traffic domain literature, commitment of air traffic controllers to safety has a significant impact on the controllers' safety-related behavior (Ek et al. 2007). In Griffin and Bateman (1986) two important aspects of organizational commitment of a human are distinguished: (a) belief and acceptance of the values and goals of the organization (safety-related goals in particular); (b) willingness to exert effort on behalf of the organization, which is an element of the human's maturity w.r.t. to organizational tasks. The acceptance of the organizational goals is influenced by the maturity degree with respect to organizational tasks (Hersey et al. 2001). Thus, the priority of safety-related goals in the role description and the maturity level with respect to ATC tasks are the primary constituents of the organizational commitment of air traffic controllers in the case study considered. According to Griffin and Bateman (1986) the commitment increases when an agent has influence on decisions related to the object of commitment. Furthermore, also the environment (own team, management) influences commitment (Burt 1987). The more cohesive the group, the more influence it has on the human. Similarly, the more control and monitoring the management exerts, the higher its influence on the human.

In the theory of situational leadership (Hersey et al. 2001) the human’s maturity w.r.t. to a task is defined as an aggregate of the human’s experience, willingness and ability to take responsibility for the task. The human’s willingness to perform a task is determined by the human’s confidence and commitment, which are necessary for the task execution. The ability of a human to perform a task is determined by his or her capabilities and the adequacy of the mental models for the task. In particular, in accordance with Endsley and Garland (2000) the adequacy of the mental models for the air traffic control tasks depends on the sufficiency and timeliness of training provided to the controller and the adequacy of knowledge about safety-related issues. Based on the domain knowledge we assume that such information is contained in reports that resulted from safety-related activities: final occurrence assessment reports that are based on occurrence investigations and monthly safety overview reports. The maturity value changes over time as a result of gaining new knowledge and skills, and changing self-confidence of a human.

Rational behavior of a human is determined by the human’s decision making, in which all internal states and characteristics discussed previously (i.e., traits, capabilities, beliefs, goals, needs, commitment, maturity) are involved (Janis and Mann 1977). In this paper a refined version of the expectancy theory by Vroom (Pinder 1998) has been used to model decision making of a human. Advantages of the expectancy theory include: (a) it can be formalized; (b) it allows incorporating the organizational context; (c) it has received good empirical support. According to this theory, when a human evaluates alternative possibilities to act, he or she explicitly or implicitly makes estimations for the following factors: *valence*, *expectancy* and *instrumentality* (see Fig. 5).

Expectancy refers to the individual’s belief about the likelihood that a particular act will be followed by a particular outcome (called a first-level outcome). Its value varies between 0 and 1. *Instrumentality* is a belief concerning the likelihood of a first level outcome resulting into a particular second level outcome; its value varies between -1 and $+1$. Instrumentality takes negative values when a second-level outcome has a negative correlation with a first-level outcome. A second level outcome represents a desired (or avoided) state of affairs that is reflected in the agent’s needs. In the proposed approach the original expectancy model is refined by considering specific types of individual needs described above. Valence refers to the strength of the individual’s desire for an outcome or state of affairs; it is also an indication of the

Fig. 5 An example of the expectancy model by Vroom (Pinder 1998); here E ’s denote expectancies, I ’s denote instrumentalities and V ’s denote valences



priority of needs. Values of expectancies, instrumentalities and valences change over time, in particular due to individual and organizational learning.

In the Vroom's model the force on an individual to perform an act is defined as:

$$F_i = \sum_{j=1}^n E\{i, j\} \cdot \sum_{k=1}^m V\{k\} \times I\{i, j, k\} \quad (1)$$

Here $E\{i, j\}$ is the strength of the expectancy that act $A\{i\}$ will be followed by outcome j ; $V\{k\}$ is the valence of the second level outcome k ; $I\{i, j, k\}$ is perceived instrumentality of outcome j for the attainment of outcome k for act i . The action alternative with the highest force is chosen to be performed by the human.

A human has the ability to learn and adapt. In general, learning of a human has many aspects and facets. In this paper by learning a change of human's skills and/or knowledge is understood. No assumption about the necessity of a performance improvement due to learning is made. In particular, forgetting is also seen as a special kind of learning. Humans in organizations vary in the abilities and desires to learn. In particular, the human's learning behavior depends on the agent's traits and needs. Some traits mediate the attainment of skills. For example, extroversion and agreeableness are important for building interpersonal skills. High priority human's needs that cannot be satisfied due to skill and/or knowledge limitations result into generation of goals to learn the required capabilities.

Two types of individual learning are identified by Meinolf et al. (2001) as particularly relevant in an organizational context: instrumentalist and social learning. By instrumentalist learning a human is able to form beliefs about dependencies between his or her own states, observed states of the environment, and observed states of other humans (such as expectancies and instrumentalities).

The second individual learning type—social learning—occurs by observation of and communication with other humans. Communication may be formal (i.e., as prescribed by a formal interaction specification in Sect. 4) and informal. To model informal communication of agents, the Burt's social contagion theory has been used (Burt 1987). According to this theory, the intensity of informal communication between agents is dependent directly on: (1) similarity of the formal communication patterns of the agents' roles; (2) equality of the formal power statuses of the agents in the organization; (3) physical possibilities to communicate; (4) degree of acquaintance of the agents with each other. For a shift of controllers the factors (1)–(3) have a high degree of evidence, whereas the factor (4) depends on the shift composition (stable versus variable).

5.2 Agent architecture

The proposed agent architecture is a particular refinement and instantiation of the component-based Generic Agent Model (Brazier et al. 2000) (see Fig. 6). In the Generic Agent Model agents are represented by autonomous components interacting with the environment and other agents by sending and receiving information through input and output interfaces. Agents perceive information by observation and generate output in the form of communication or actions. Since agents can play organizational roles, interaction among them is specified using the interaction ontologies of roles.

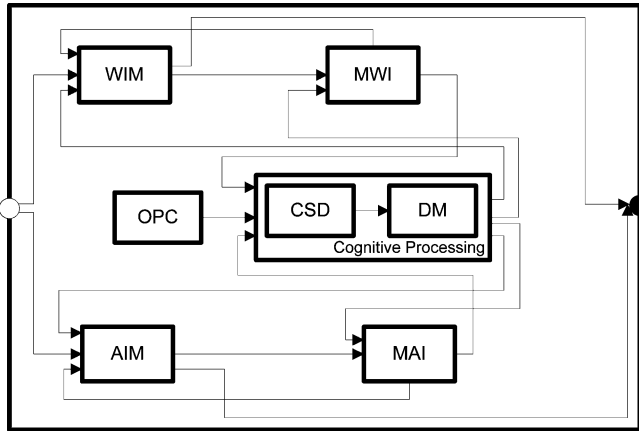


Fig. 6 Component-based Agent architecture

In contrast to roles agents have internal states. The internal dynamics of an agent is realized in the architecture proposed by the components described in the following: The component World Interaction Management (WIM) takes care of interaction with the world, the component Agent Interaction Management (AIM) takes care of communication with other agents. The component Maintenance of World Information (MWI) maintains beliefs about the world, and the component Maintenance of Agent Information (MAI) maintains beliefs about other agents. The component Own Process Control (OPC) coordinates the internal agent processes. The component Agent Specific Task in the original Generic Agent Model is specialized into the Cognitive Processing component which consists of the Cognitive State Determination (CSD) component and the Decision Making component (DM). The functionality of the components is specified by direct causal relations between internal and interaction states of the agent. The specification of the components described in the following sections relies heavily on the concepts, relations and mechanisms from the theory of human behavior and cognition described previously in Sect. 5.1.

5.2.1 World and agent interaction management

WIM and AIM components create beliefs about agent’s input and output states, which are stored in MWI and MAI components respectively. A belief generation rule for an observation of an agent *a* is specified as follows:

$$\forall p:\text{STATE_PROPERTY } \text{observation_result}(p) \ \& \ \text{current_time}(t) \ \rightarrow \ \text{belief}(p, t)$$

Information about observed safety occurrences is stored by agents as beliefs: e.g., $\text{belief}(\text{observed_occurrence_with}(\text{ot: OCCURRENCE_TYPE, ag:AGENT}), t:\text{TIME})$. An agent forms also beliefs about dependencies between its own states, observed states of the environment, and observed states of other agents:

$$\text{belief}(\text{occurs_after}(p1:\text{STATE_PROPERTY, } p2:\text{STATE_PROPERTY, } t1:\text{TIME, } t2:\text{TIME}), t:\text{TIME}),$$

which expresses that state property p_2 holds t' ($t_1 < t' < t_2$) time points after p_1 holds.

Such beliefs are created based on individual and organizational learning; mechanisms for generation of such beliefs are considered in Sect. 5.3.

In the considered case, each controller agent creates the belief about the dependency between providing a notification report on an occurrence of some type to his/her supervisor and receiving a final assessment report on the occurrence (i.e., feedback) from a safety investigator agent. Moreover, often final assessment reports include recommendations for organizational and environmental improvement, which may be observed by the controller-reporter when implemented. These dependencies are formalized as:

```
belief(occurs_after(output(ag_controller,
communicated_from_to(ag_controller, ag_supervisor, inform,
notification_report_for_B)),
output(ag_investigator,communicated_from_to(ag_investigator, ag_controller, inform,
final_assessment_report_for_B))), 24h, 1440h)
belief(occurs_after(output(ag_investigator, communicated_from_to(ag_investigator,
ag_controller, inform, final_assessment_report_for_B)),
output(env, observation_result(recommendations_implemented_for_B))),
360h,8640h))
```

5.2.2 Own process control

OPC component contains characteristics and attitudes of the agent, which do not change or change very slowly over time, in particular personal traits and capabilities. These characteristics play a role in the cognitive processing of the agent. For the case study the ranges of the uniformly distributed cultural indexes for air traffic controller agents from ANSP3 are defined in OPC based on the indexes for a Western European culture provided in Hofstede (2005): IDV: [0.7, 0.9]; PDI: [0.3, 0.5]; UAI: [0.4, 0.6]. Furthermore, for the controller agents from ANSP3 the priority values of needs provided in Table 3 are assumed. The values were determined based on the cultural classification framework from Hofstede (2005); each value is from [0, 1]. According to Locke (2001) and van de Walle (1997) the behavior of a human is goal-directed and is steered by his or her needs. Thus, needs exert a significant influence on the cognitive dynamics of a human, in particular on decision making.

Table 3 The minimal acceptable satisfaction values of needs and the ranges of the uniformly distributed basic valences of needs of controller agents from the ANSP3 used in the simulation

Need	Minimum acceptable satisfaction level
1. Extrinsic needs	1
2. Own group approval	0.6
3. Management approval	0.5
4. Contribution to organizational safety-related goals	0.7
5. Self-esteem, self-confidence and self-actualization needs	0.9

5.2.3 Cognitive state determination

Component Cognitive State Determination is responsible for the update of the agent's cognitive states used in its decision making. The complete specification of the agent's cognitive dynamics is provided in Appendix A (supplementary material). In this section, the dynamics of the commitment to safety and maturity level considered in Sect. 5.1 are described formally.

The dynamics of cognitive states is specified by causal networks. The relations in the causal networks are formalized by temporal equations. Some of these equations are instantaneous, whereas others are difference equations over two successive time steps. In such a way one can specify mutual influences of states and loops in cognitive processes of agents. In particular, based on the theoretical findings from Sect. 5.1 the commitment value is calculated based on a feedback loop: the agent's commitment influences the team commitment, but also the commitment of the team members and of the management influence the agent's commitment:

$$e6_{a,t} = w1 \cdot e1_{t-1} + w2 \cdot e2_{a,t-1} + w3 \cdot e3_{a,G,t-1} + w4 \cdot e4_{a,t-1} + w4 \cdot e5_{a,t-1},$$

here $e1_t$ is the priority of safety-related goals in the controller's role description, $e2_{a,t}$ is the perception of the commitment to safety of management, $e3_{a,G,t}$ is the perception of the average commitment to safety of the team, $e4_{a,t}$ is the perceived influence degree of controller a on safety arrangements, $e5_{a,t}$ is the controller's maturity level w.r.t. the task; $w1$ - $w4$ are the weights.

The perception of agent a of the average commitment to safety of its team G at time point t ($e3_{a,G,t}$) is based on the perception of commitment to safety of the team supervisor s and of other team members and is calculated as:

$$e3_{a,G,t} = w16 \cdot e13_{a,t-1} + w17 \cdot e41_{a,G,t-1}$$

The perception of commitment to safety of the team supervisor s is determined by:

$$e13_{a,t} = w14 \cdot e14_{s,t-1} + w15 \cdot e2_{a,t-1},$$

here $e14_{s,t}$ is the level of development of the managerial skills of the supervisor of the agent's team G (i.e., $is_in_team(a, G) \wedge is_supervisor_of(s, G)$); $e2_{a,t}$ is the perception of the commitment to safety of management.

The perception of commitment to safety of other team members in G is calculated as:

$$e41_{a,G,t} = \sum_{i \in T} e6_{i,t} / |T|,$$

here $T = \{ag \mid is_in_team(a, G) \text{ AND } \exists ag:AGENT \ is_in_team(ag, G) \text{ AND } is_supervisor_of(s, G) \text{ AND } a \neq ag \text{ AND } ag \neq s\}$, i.e., the set of all agents from G except s and a; $e6_{i,t-1}$ is the commitment of agent i at time point t-1; $w14$ - $w17$ are weights.

An agent evaluates the management's commitment to safety by considering factors that reflect the management's effort in contribution to safety (investment in personnel

and technical systems, training, safety arrangements):

$$e2_{a,t} = w6 \cdot e7_{t-1} + w7 \cdot e8_{t-1} + w8 \cdot e1_{t-1} + w9 \cdot e9_{t-1} + w10 \cdot e70_{G,t-1} \\ + w11 \cdot e10_{t-1} + w12 \cdot e11_{t-1} + w13 \cdot e12_{t-1}$$

here $e7_t$ is the sufficiency of the amount of safety investigators, $e8_t$ is the sufficiency of the amount of controllers, $e1_t$ is the priority of safety-related goals in the controller’s role description, $e9_t$ is the availability of up-to-date technical systems for controllers, $e10_t$ is the sufficiency and timeliness of training for changes, $e11_t$ is the regularity of safety meetings, $e12_t$ is the quality of developed and implemented safety management system (SMS), $e70_{G,t}$ is the averaged perceived degree of influence of the controllers on safety arrangements:

$$e70_{G,t} = \sum_{ag \in GR} e4_{ag,t} / |GR|, \\ \text{where } GR = \{g | \text{is_in_team}(g, G) \text{ AND } \text{is_in_team}(a, G)\}$$

here $e4_{ag,t}$ is the perceived influence degree of controller agent ag on safety arrangements.

Based on the theoretical findings from Sect. 5.1, the maturity level of a controller agent a at time point ($e5_{a,t}$) is calculated by a difference equation as:

$$e5_{a,t} = w22 \cdot e19_{a,t-1} + w23 \cdot e20_{a,t-1} + w24 \cdot e21_{a,t-1} + w25 \cdot e23_{a,t-1}$$

here $e19_{a,t}$ is the agent’s self-confidence w.r.t. the ATC task (depends on the number of occurrences with the controller); $e20_{a,t}$ is the agent’s commitment to perform the ATC task; $e21_{a,t}$ is the agent’s development level of skills for the ATC task; $e23_{a,t-1}$ is the adequacy of the mental models for the air traffic control (ATC) tasks.

The adequacy of mental models of the agent for the ATC task ($e23_{a,t}$) is defined as:

$$e23_{a,t} = w26 \cdot e10_t + w27 \cdot e22_{a,t}$$

here $e10_t$ is the indicator for sufficiency and timeliness of training for changes. The adequacy of knowledge about safety issues ($e22_{a,t}$) is defined by:

$$e22_{a,t} = w28 \cdot \sum_{s_i \in S1} s_i / |S1| + w29 \cdot \sum_{s_i \in S2} s_i / |S2|,$$

where $S1 = \{q1 | \text{has_state}(a, \text{belief}(\text{final_report}(\text{occurrence}(o, cl, ag1), q1))) \text{ holds at } t\}$; $\sum s_i / |S1|$ is the average quality of the final occurrence assessment reports received by the agent up to the time point t ; $S2 = \{q2 | \text{has_state}(a, \text{belief}(\text{monthly_report}(m, q2))) \text{ holds at } t\}$, $\sum s_i / |S2|$ is the average quality of the monthly safety overview reports received by the agent up to the time point t , $w22$ - $w29$ are the weights. All evidence (e) variables vary in the range $[0, 1]$.

The completeness and correctness of knowledge of a controller agent are determined to a great extent by the quality of reports obtained by the agent. The model

includes factors that influence the quality of a notification report for an observed occurrence, such as the maturity level w.r.t. the ATC task, the quality of technical systems and the acceptability of the workload. The quality of the notification report influences the quality of the occurrence investigation.

The quality of a notification report is influenced by the quality of perception and comprehension of a situation by a controller in the context of execution of an air traffic control task. The perception and comprehension of a controller are dependent on:

- The quality of technical systems ($e17_t$), which ensures (timely) provision of the required data and minimal distortion of these data:

$$e17_t = w30 \cdot e25_t + w31 \cdot e26_t + w32 \cdot e9_t$$

here $e25_t$ is sufficiency of the amount of maintenance personnel, which is dependent on the number of technical systems that need to be maintained and on the degree of exploitation of these systems by controllers, $e26_t$ is the quality of formal procedures for system checks and repairs depending on the clarity of identification of role responsibilities, on the frequency of system checks and on scheduling and timely performance of system repairs, and $e9_t$ is the availability of reliable and ergonomic technical systems for controllers required for the effective and efficient execution of air traffic control tasks.

- Acceptability of the workload level ($e18_{G,t}$): it is assumed that quality of perception and comprehension of a controller degrades when his/her workload level increases above some threshold:

$$e18_{G,t} = w39 \cdot e8_t + w40 \cdot e14_{a,t},$$

here a denotes the supervisor of team G , i.e., $is_supervisor_of(a, G)$, $e8_t$ is the sufficiency of the number of controllers, evaluated by determining if the number of controllers in each sector is sufficient to handle the traffic in the sector; $e14_{a,t}$ is the level of development of managerial skills of the supervisor. (Highly) developed management skills of the supervisor contribute to a work distribution under which each team member has an acceptable for him/her workload level.

- Information contribution provided by other agents ($e16_{o,ot,a,t}$). Other agents may also provide useful information to a controller, which may enhance his/her comprehension of a situation in the context of an air traffic control task. Among these agents are controller(s) assigned to the same sector, controllers of the adjacent sectors, and the team supervisor. To provide information useful for the occurrence recognition, the controller-information provider should possess knowledge relevant to the occurrence (e.g., knowledge about similar observed occurrences). The more knowledge related to the occurrence the controller possesses, the higher the effect of his/her information contribution. In this model it is assumed that the knowledge about three or more similar occurrences is the maximal amount of knowledge. This number was determined roughly based on the frequencies of occurrences of different types in ANSP3. Furthermore, the motivation to provide this knowledge depends on the commitment of the controller to safety. Similarly,

the information contribution by the team supervisor depends on the amount of knowledge about safety-related issues, and on his/her commitment to safety.

$$e16_{o,ot,a,t} = w33 \cdot (w34 \cdot e6_{a,t} + w35 \cdot e27_{o,ot,a,t}) + w36 \cdot (w37 \cdot e28_{o,ot,a,t} + w38 \cdot e13_{a,t})$$

here $e6_{a,t}$ is the commitment to safety of controller agent a, $e13_{a,t}$ is the a's perception of the commitment to safety of the team supervisor, $e27_{o,ot,a,t}$ is the amount of knowledge on similar safety-related issues of a controller:

$$e27_{o,ot,a,t} = 0, \quad \text{when } \max_{ag \in H} |S_{ag}| = 0;$$

$$e27_{o,ot,a,t} = 1, \quad \text{when } \max_{ag \in H} |S_{ag}| > \text{sufficient_similar_occur};$$

$$e27_{o,ot,a,t} = 1 / \max_{ag \in H} |S_{ag}|, \quad \text{when } \max |S_{ag}| > 0 \quad \text{and}$$

$$\max |S_{ag}| \leq \text{sufficient_similar_occur}$$

where $H = \{a \mid \text{is_in_team}(ag, gr) \ \& \ \text{is_in_team}(a, gr) \ \& \ a \neq ag \text{ holds at } t\}$, $S_a = \{\text{has_state}(a, \text{belief}(\text{occurrence}(o1, ot, a1)))\}$;

$e28_{o,ot,a,t}$ is the amount of knowledge on similar safety-related issues of the supervisor:

$$e28_{o,ot,ag,t} = 0, \quad \text{when } |S| = 0;$$

$$e28_{o,ot,ag,t} = 1/|S|, \quad \text{when } |S| > 0 \text{ and } |S| \leq \text{sufficient_similar_occur};$$

$$e28_{o,ot,ag,t} = 1, \quad \text{when } |S| > \text{sufficient_similar_occur};$$

where $S = \{\text{has_state}(a, \text{belief}(\text{occurrence}(o1, ot, a1))) \mid \text{is_in_team}(ag, gr) \ \& \ \text{is_supervisor_of}(a, gr) \}$

- Maturity level w.r.t. ATC task ($e5_{a,t}$).

Thus, the quality of a notification report produced by agent a for an occurrence o of type ot is calculated as:

$$e15_{o,ot,a,t} = w18 \cdot e16_{o,ot,a,t} + w19 \cdot e5_{a,t} + w20 \cdot e17_t + w21 \cdot e18_{G,t}$$

For calculating the quality of other types of reports we refer to the Appendix A (supplementary material).

5.2.4 Decision making

The agent's decision making considers the evaluation of the forces for two alternatives: to report and to not report. A decision making model for reporting an occurrence is provided in Fig. 7. The agent chooses to perform the alternative with a greater force. In the following the basis for calculation of the variables of the decision making model for reporting is discussed. The precise, elaborated details of the

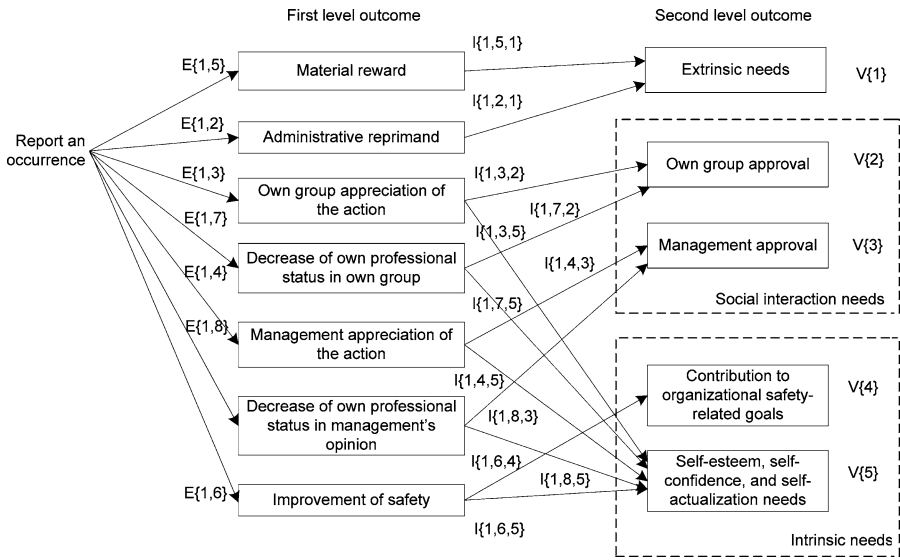


Fig. 7 Decision making model for reporting an occurrence

mathematical model for both reporting and not reporting can be found in Appendix A (supplementary material).

The factors $E\{1, 5\}$, $E\{1, 2\}$, $I\{1, 5, 1\}$ and $I\{1, 2, 1\}$ are defined based on the ANSP’s formal reprimand/reward policies. In particular, $E\{1, 2\} = 1$ for an observed occurrence, which completes a set of occurrences, for which a reprimand is defined; $E\{1, 2\} = 0$ for all other observed occurrences. The values of $E\{1, 3\}$ and $I\{1, 3, 2\}$ depend largely on the average commitment of the team of controllers to safety, and $E\{1, 8\}$ and $I\{1, 4, 3\}$ depend on the management commitment to safety. In particular:

$$E\{1, 3\}_{a,t} = (1 - e66_{a,t}) \cdot (dw1 \cdot e57_{G,t} + (1 - dw1) \cdot e3_{a,G,t})$$

where G is such that $is_in_team(a, G)$ holds, $e3_{a,G,t}$ is the perception of agent a of the average commitment to safety of its team G , $e66_{a,t}$ is the agent’s perception of confidentiality of reporting, $e57_{G,t}$ is the perception of the positive attitude to reporting in team G :

$$e57_{G,t} = \sum_{s_i \in S1} s_i / |S1|,$$

where $S1 = \{v \mid has_state(ag, belief(force(report(o, cl), v), t)) \ \& \ is_in_team(ag, G)\}$. Agents learn by observing the behavior of other agents. In the case study a controller agent observes occurrence reporting of other agents from his/her shift, and based on that forms the beliefs about the shift’s averaged attitude to reporting of different types of occurrences ($e57_{G,t}$).

With each set of occurrences, in which a controller agent was involved during an evaluation period (e.g., a month), a measure of severity is associated, calculated as the sum of the severities of the occurrences from the set. The factors $E\{1, 7\}$, $E\{1, 8\}$,

$I\{1, 7, 2\}$, $I\{1, 4, 3\}$ depend mostly on the severity of the set of occurrences of the controller known to his/her team and known to the management. In particular,

$$E\{1, 7\}_{a,t} = (1 - e66_{a,t}) \cdot (e65_{a,t} \cdot e61_a + (1 - e35_{G,t}) \cdot (1 - e61_a))$$

$$I\{1, 7, 2\}_{a,t} = dw8 \cdot (e65_{a,t} \cdot e61_a + (1 - e35_{G,t}) \cdot (1 - e61_a)),$$

here agent a is in team G , $e66_{a,t}$ is the agent’s perception of confidentiality of reporting, $e65_{a,t}$ is the total severity of a set of occurrences of the controller occurred during the evaluation time interval, which are known to the team of the controller, $e61_a$ is the individualism index (IDV) value of the controller agent; $e35_{G,t}$ is the intensity of informal interaction in the team of controllers.

To calculate the intensity of informal interaction in a team G of controllers ($e35_{G,t}$) according to the Burt’s social contagion theory introduced in Sect. 5.1, a measure of degree of acquaintance (acq) of agents $ag1$ and $ag2$ worked n shifts together is introduced, calculated as follows:

if $n = 0$ and t is the current time point, then $acq(ag1, ag2, t) = 0$;

else if $acq_shifts > n > 0$ and t is the current time point, then $acq(ag1, ag2, t) = e35_{max}/n$

else if $n \geq acq_shifts$ and t is the current time point, then $acq(ag1, ag2, t) = e35_{max}$, where $e35_{max}$ is the maximum intensity of informal interactions in the team of controllers.

Then, $e35_{G,t} = \sum acq(ag1, ag2, t) / (|G| \cdot (|G| - 1))$.

Informal interaction relations enhance the knowledge of controllers about safety issues and observed occurrences, and may contribute to the proactive identification of issues by the controllers. The degree of contribution of informal discussions of controllers in a team G to the quality of input for proactive identification of issues by the controllers is specified by:

$$e34_{G,t} = w49 \cdot \sum_{a \in \{ag | is_in_team(ag, G)\}} e3_{a,G,t} / |G| + w51 \cdot \sum_{s_i \in S1} s_i / |S1|$$

$$+ w52 \cdot \sum_{s_i \in S2} s_i / |S2| + w50 \cdot e35_{G,t}$$

where $S1 = \{q \mid \exists t1 \text{ has_state}(a, \text{belief}(\text{notification_report_for}(\text{occurrence}(o, cl, a1), q), t1)) \text{ holds}\}$,

$S2 = \{q \mid \exists t2 \text{ has_state}(a, \text{belief}(\text{monthly_report}(m, q), t2)) \ \& \ \text{is_supervisor_of}(a, G) \text{ holds}\}$

$e3_{a,G,t}$ is the perception of agent a of the average commitment to safety of its team G .

$E\{1, 6\}$ is based on the agent’s beliefs about the dependencies between previous reporting of similar occurrences and improvement of safety that followed:

$$E\{1, 6\}_{o,cl,a,t} = dw3 \cdot e39_{o,cl,t} + dw4 \cdot e59_{cl,a,t} + dw5 \cdot e2_{ag,t}$$

$$+ (1 - dw3 - dw4 - dw5) \cdot e43_{a,t}$$

Here $e39_{o,cl,t}$ is the severity of the occurrence o ; $e2_{ag,t}$ is the perception of the commitment to safety of management; $e59_{cl,a,t}$ is the average quality of the final safety occurrence assessment reports for occurrences similar to the observed occurrence provided to controller agent a :

$$e59_{cl,a,t} = \sum_{s_i \in S1} s_i / |S1|,$$

where $S1 = \{ v \mid \text{has_state}(a, \text{belief}(\text{final_report}(\text{occurrence}(o, cl, a), v), t)) \}$ $e43_{a,t}$ is the average quality of the monthly safety overview reports received by the agent:

$$e43_{a,t} = \sum_{s_i \in S1} s_i / |S1|,$$

where $S1 = \{ v \mid \text{has_state}(a, \text{belief}(\text{monthly_report}(m, v), t)) \}$

$I\{1, 3, 5\}$ and $I\{1, 7, 5\}$ are based on the agent’s IDV index, which indicates the degree of importance of team’s opinions for the agent.

$I\{1, 4, 5\}$ and $I\{1, 8, 5\}$ are based on the agent’s PDI index. In particular,

$$I\{1, 8, 5\}_{a,t} = dw12 \cdot e62_a \cdot e60_{a,t}$$

where $e62_a$ is the power distance index (PDI) value of controller agent a , $e60_{a,t}$ is the total severity of a set of occurrences of the controller occurred during the evaluation time interval, which are known to the management.

Furthermore, also the values of the basis valences (the degrees of importance of particular needs taken alone, see Fig. 7) of a controller agent depend on its indexes:

$$\begin{aligned} v\{1\}_a^b &= 1 & v\{2\}_a^b &= 1 - e61_a \\ v\{3\}_a^b &= 0.7 \cdot e62_a + 0.3 \cdot e64_a & v\{4\}_a^b &= 0.3 + 0.7 \cdot e64_a \end{aligned}$$

$e61_a$ is the value of individualism (IDV) index of controller a , $e62_a$ is the power distance index (PDI) value of controller agent a , $e64_a$ is the uncertainty avoidance index (UAI) value of controller agent a .

The values of valences change over time depending on the degree of satisfaction of the agent’s needs: the more a need is satisfied, the less its valence:

$$v\{need\}_{a,t} = \begin{cases} v\{need\}_a^b \cdot \frac{\text{sat_accept}_{need,a}}{\text{sat}_{need,a,t}}, & \text{sat}_{need,a,t} \geq \text{sat_accept}_{need,a} \\ v\{need\}_a^b + v\{need\}_a^b \cdot \frac{\text{sat_accept}_{need,a} - \text{sat}_{need,a,t}}{\text{sat_accept}_{need,a}}, & \text{sat}_{need,a,t} < \text{sat_accept}_{need,a} \end{cases}$$

here $\text{sat_accept}_{need,a}$ is defined in Table 3, $\text{sat}_{need,a,t}$ is the current satisfaction value of a need.

Individual learning influences expectancies and instrumentalities of the decision making models of a controller agent significantly. In particular, the agent’s expectancies $E\{1, 5\}$, $E\{1, 2\}$ and $E\{2, 6\}$ change depending on the received (observed) reprimands and rewards for occurrences reported by the agent (or by another agent from

the shift) (instrumentalist (social) learning). $E\{1, 3\}$ and $E\{2, 2\}$ are adjusted by the agent based on the observed shift's averaged attitude to reporting of different types of occurrences (social learning). $E\{1, 6\}$ and $E\{2, 5\}$ are adjusted based on the feedbacks from the safety investigator agent on the previously reported occurrences and the observed implementation of safety recommendations for previous reports (instrumentalist learning), and safety information informally provided by other controllers during breaks (social learning).

5.3 Organizational learning

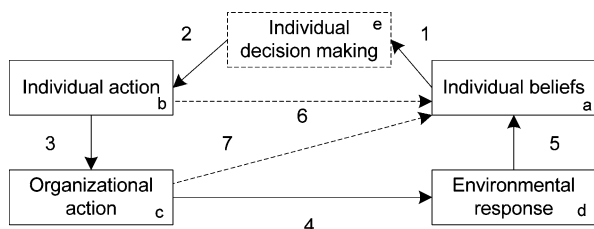
In the context of organizations a number of studies have been performed, which investigated the link between individual and organizational learning (Meinolf et al. 2001), viewed in particular from the adaptation and knowledge perspectives.

From the adaptation perspective organizational learning is defined as the process of organizational adaptation to internal/external changes based on individual learning. Within this perspective the model of organizational learning by March and Olsen (1975) (see Fig. 8) gained much attention. In this model an individual action is based on certain individual beliefs. This action, in turn, may lead to an organizational action, which may produce some environmental response. The learning cycle is completed when the environmental response affects individual beliefs. In such a way an individual learns through instrumental learning by observing the effects of own actions. In such a way beliefs about (causal) relations between states can be created. For example, if an agent receives a reprimand resulted from its reporting of an occurrence of some type, then it will create a belief about a relation between these two states (i.e., 'reporting an occurrence of the type' and 'getting reprimand'). This belief will influence decision making of the agent on whether to report occurrences of the same type in the future. The learning process may be made more subtle by assigning degrees of confidence to the beliefs, as proposed in Sharpanskykh (2009).

The model also addresses the issue of incomplete learning cycles, when learning is impaired because of the weakening or breakage of one or more links. For example, link 3 in Fig. 8 is broken, when the individual affects organizational action in an ambiguous way; link 4 is broken when no real basis for the connection between organizational action and environmental response can be identified.

Many studies (e.g., Meinolf et al. 2001) have shown that information about incomplete learning cycles is also remembered by individuals and used in their decision making, thus influencing the organizational learning. In such a way organizational learning does not always lead to the organizational improvement. To capture such effects, the March and Olsen's model is extended in this paper as shown in Fig. 8

Fig. 8 Model of organizational learning adapted from March and Olsen (1975) with the modifications indicated by dotted lines



by dotted lines. Also, the individual decision making process, implicit in the original model, is made explicit.

From the *knowledge perspective* organizational learning is defined as accumulation of organizational knowledge. From this perspective a significant part of the organizational knowledge is captured in the organizational routines (e.g., procedures, regulations, guidelines). Furthermore, much of not formalized organizational knowledge can be learned through informal social interaction within the organization (individual social learning).

An interaction between the adaptation and knowledge perspectives can be established through individual beliefs. More specifically, an individual forms beliefs about formally and informally provided organizational knowledge. A relevant subset of these beliefs can be used further in the individual decision making about the action choice as shown in Fig. 8.

In the model considered organizational learning leads to safety improvement, when controller agents choose to report occurrences. In this case from the knowledge perspective organizational learning leads to the accumulation of information about observed safety occurrences. From the adaptation perspective, safety improvement is supported by complete learning cycles.

5.4 Safety culture modeling

Based on the cognitive and behavioral states described in the previous sections precise expressions for the safety culture indicators identified in Sect. 3.1 have been determined.

SCI1: Average reporting quality of controllers This indicator is determined based on the decisions of controller agents whether to report observed occurrences:

$$SCI1 = \sum_{a \in \text{CONTROLLER}, ot \in \text{OCC_TYPE}} n_{ot,a} / (|\text{CONTROLLER}| \cdot |\text{OCC_TYPE}| \cdot m_{ot,a}),$$

where $n_{ot,a}$ is the number of occurrences of type ot observed and reported by an agent controller a during the simulation; $m_{ot,a}$ is the number of occurrences of type ot observed by an agent controller a during the simulation; CONTROLLER is the set of all names of the controller agents; OCC_TYPE is the set of all names of the occurrence types.

SCI2: Average quality of the processed notification reports

$$SCI2 = \sum_{a \in \text{CONTROLLER}, t \in \text{TIME}, ot \in \text{OCC_TYPE}, o \in \text{OCCUR}} e15_{o,ot,a,t} / |\text{OCCUR}|,$$

where $e15_{o,ot,a,t}$ is the quality of the notification report for an occurrence o of type ot reported by an agent a at time point t ; OCCUR is the set of names of all occurrences.

SCI3: Average quality of the final safety occurrence assessment reports

$$SCI3 = \sum S/|S|,$$

where $S = \{v \mid \exists a:CONTROLLER, o:OCCUR, cl:OCC_TYPE, t:TIMEhas_state(a, belief(final_report(occurrence(o, cl, a), v), t))\}$

SCI4: Average quality of the monthly safety overview reports received by controllers

$$SCI4 = \sum S/|S|,$$

where $S = \{v \mid \exists a:CONTROLLER, m:MONTH, t:TIMEhas_state(a, belief(monthly_report(m, v), t))\}$

SCI5.1: Average commitment to safety of controllers

$$SCI5.1 = \sum_{a \in CONTROLLER, t \in TIME} e6_{a,t} / (|CONTROLLER| \cdot |TIME|),$$

here $e6_{a,t}$ is the commitment to safety of an agent a at time point t ; $CONTROLLER$ is the set of all names of controller agents, and $TIME$ is the set of all time points.

SCI5.2: Average commitment to safety of a team as perceived by controllers

$$SCI5.2 = \sum_{a \in CONTROLLER, t \in TIME, G \in TEAM} e3_{a,G,t} / (|CONTROLLER| \cdot |TEAM| \cdot |TIME|),$$

here $e3_{a,G,t}$ is the perception of the average commitment to safety of the team G at time point t , $CONTROLLER$ is the set of all names of the controller agents, $TEAM$ is the set of all team names, $TIME$ is the set of all time points.

SCI6: Average commitment to safety of a supervisor as perceived by controllers

$$SCI6 = \sum_{a \in CONTROLLER, t \in TIME} e13_{a,t} / (|CONTROLLER| \cdot |TIME|),$$

here $e13_{a,t}$ is the perception of commitment to safety of the team supervisor of the team of agent a ; $CONTROLLER$ is the set of all names of the controller agents, $TIME$ is the set of all time points.

SCI7: Average commitment to safety of management as perceived by controllers

$$SCI7 = \sum_{a \in CONTROLLER, t \in TIME} e2_{a,t} / (|CONTROLLER| \cdot |TIME|),$$

here $e2_{a,t}$ is the perception of the commitment to safety of management of agent a at time point t , $CONTROLLER$ is the set of all names of the controller agents, $TIME$ is the set of all time points.

6 Model results

First, simulation results obtained based on the developed model are discussed in Sect. 6.1 These results reflect the estimated quality of the organizational safety culture. Next, a sensitivity analysis approach to identify main sources of deficiencies in the safety culture is proposed in Sect. 6.2. Moreover, it is demonstrated how results of such analysis can be used for structured identification of safety culture improvement options. Model validation results are considered in Sect. 6.3.

6.1 Simulation

Based on the developed model 3000 simulation trials were performed in the Matlab environment, where each trial represents three years of operations. The probabilities of different types of occurrences used in the simulation are specified based on the Annual Safety Report 2007 of ANSP-3.

To perform simulation the model has been initialized based on the obtained information about the organization by assigning values to:

- independent inputs of the causal networks in the model (the complete list is provided in Appendix A, supplementary material). These variables were evaluated using part of the results of the safety culture survey questionnaire of ANSP-3, as detailed in supplementary material. In particular the questions used to determine the values of the input variables were separated from the questions to determine the reference results of the safety culture indicators in the model validation phase (Sect. 6.3);
- weights, which reflect relative contribution of states in aggregations for other states in the causal networks (the model contains 60 weights);
- task durations, which refer to the (limits of) task durations in workflows, the model contains upper and lower bounds for 16 tasks;
- decision making parameters, which are related to the decision making models from Sect. 5.2 (28 parameters);
- other (55 parameters).

The model contains a number of stochastic processes, which represent the randomness in the happening of the four types of occurrences, the identification not reported occurrences by a team of controllers, the identification of not reported occurrences by the management, the feedback provision to a controller for a reported occurrence, and the duration of task processing. Details of these intrinsic stochastic processes are in Appendix A (supplementary material).

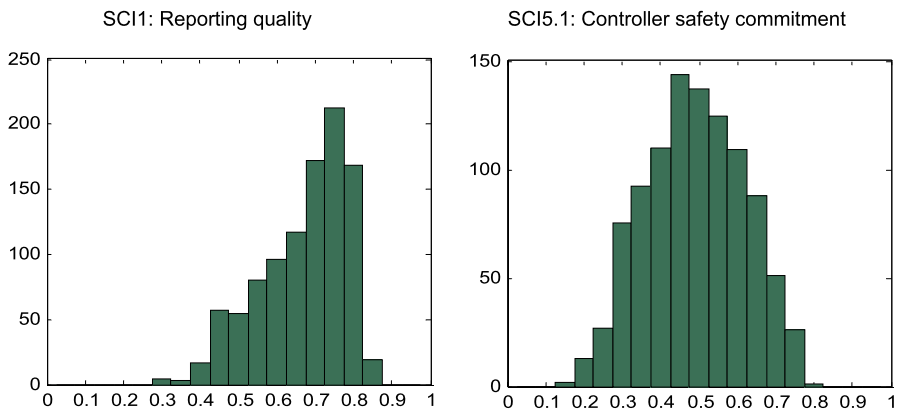
For the complete set input variables and parameters used in the simulation we refer to Appendix A (supplementary material).

The obtained means and variances of the safety culture indicators are provided in Table 4. It can be observed that the variances are very small, which indicate that the randomness due to the intrinsic stochastic processes has almost no effect on the safety culture indicators.

The results in Table 4 are the model predictions of the safety culture indicators given the fixed estimates of the input variables. To assess the effect of variation in the

Table 4 Simulation results for the safety culture indicators of ANSP-3, based on 3000 simulation trials of three years of operation each

SCI	Description	Mean	Variance
SCI1	Average reporting quality of controllers	0.74	5e-3
SCI2	Average quality of the processed notification reports	0.54	4e-4
SCI3	Average quality of the final safety occurrence assessment reports	0.2	7e-3
SCI4	Average quality of the monthly safety overview reports received by controllers	0.58	6e-5
SCI5.1	Average commitment to safety of controllers	0.56	2e-3
SCI5.2	Average commitment to safety of a team as perceived by controllers	0.55	1e-3
SCI6	Average commitment to safety of a supervisor as perceived by controllers	0.6	8e-4
SCI7	Average commitment to safety of management as perceived by controllers	0.61	4e-5

**Fig. 9** Examples of distributions of safety culture indicators (Monte Carlo simulation results); all input variables were varied over their full range, except the national culture variables which were associated with the Western European culture; the *vertical axis* is the number of simulation runs

input variables, we performed additional Monte Carlo simulations in which all input variables were varied over their full range, except the national culture variables which were associated with the Western European culture. Figure 9 provides examples of histograms for the Monte Carlo simulations results for two safety culture indicators (1000 simulation trials each). As indicated by these typical results in Fig. 9, the variance of the safety culture indicators due to the variation in model inputs is much higher than the variance due to the intrinsic stochastic processes shown in Table 4. A more detailed analysis of the sensitivity of the safety culture indicators for the input variables is provided next in Sect. 6.2.

6.2 Sensitivity analysis

Sensitivity analysis can be used to identify the most influential organizational factors for each safety culture indicator. Values of these factors determine to a large extent the safety culture indicator value. Thus, the cause of a deficiency in the value of an indicator can be attributed to the corresponding organizational factors. The sensitivity analysis method used in this study is Monte Carlo filtering (Saltelli et al. 2008). The aim of Monte Carlo filtering is to identify the model parameters of which the variation according to associated credibility intervals lead to significant differences in attained model output classes. It consists of two steps, which are presented next.

Step 1: MC simulations For the complete set of model parameters, lower and upper bounds of credibility intervals of their values were determined based on our knowledge of the ANSP3 organization and on our knowledge about the uncertainty in modeled aspects. Next, 8000 Monte Carlo simulation trials were performed where in each simulation the parameters were chosen uniformly within their credibility interval bounds. For each input factor x_i two sets of values were determined: $x_i|B$, containing all values of x_i from the simulations that produced a High safety culture indicator, and $x_i|\underline{B}$, containing all x_i values that produced a Low or Medium safety culture indicator.

Step 2: Smirnov test A Smirnov two sample test was performed for each input factor independently. The applied test statistics are

$$d(x_i) = \sup_Y \|F_B(x_i|B) - F_{\underline{B}}(x_i|\underline{B})\|,$$

where F_B and $F_{\underline{B}}$ are marginal cumulative probability distribution functions calculated for the sets $x_i|B$ and $x_i|\underline{B}$, respectively, and where Y is the output. A low level of $d(x_i)$ supports the null-hypothesis $H_0: F_B(x_i|B) = F_{\underline{B}}(x_i|\underline{B})$, meaning that the input factor x_i is not important, whereas a high level of $d(x_i)$ implies the rejection of H_0 meaning that x_i is a key factor.

It was determined at what significance level α , the value of $d(x_i)$ implies the rejection of H_0 , where α is the probability of rejecting H_0 when it is true. In the sensitivity analysis, the classification High/Medium/Low for the importance of each factor was used. If $\alpha \leq 0.01$, then the importance of the corresponding factor x_i is High; if $0.01 < \alpha \leq 0.1$, then the importance of the corresponding factor is Medium, and if $\alpha > 0.1$, then the importance of the corresponding factor is Low.

In Table 5 some examples of the importance of input variables are given for each safety culture indicator, according to above methodology.

To analyze the reproducibility of the sensitivity analysis results two similarity measures were introduced for two sets of safety culture indicators S^1 and S^2 obtained through sensitivity analysis by performing the same number of simulations twice; a set S comprises indicators that belong to the class High (subset S_H) and indicators that belong to the class Medium (subset S_M):

$$i(S^1_H, S^2_H) = |S^1_H \cap S^2_H| / |S^1_H \cup S^2_H| \text{ (similarly for } S^1_M \text{ and } S^2_M)$$

$$i(S^1_H, S^1_M, S^2_H, S^2_M) = |(S^1_H \cup S^1_M) \cap (S^2_H \cup S^2_M)| / |(S^1_H \cup S^1_M) \cup (S^2_H \cup S^2_M)|$$

Table 5 Importance of some input variables classified by categories (H)igh/(M)edium/(L)ow for the ANSP3’s safety culture indicators

Var	Description	SCI 1	SCI 2	SCI 3	SCI 4	SCI 5.1	SCI 5.2	SCI 6	SCI 7
e1	Priority of safety-related goals in the role description	M	M	L	M	H	H	H	H
e4	Influence of a controller on safety activities	H	L	L	L	H	H	H	H
e7	Sufficiency of the number of safety investigators	M	L	H	H	H	H	H	H
e10	Sufficiency and timeliness of training for changes	L	L	L	L	H	H	H	H
e12	Developed and implemented SMS	M	M	L	L	H	H	H	H

Table 6 Similarity indexes for sets $S_H^1, S_H^2, S_M^1, S_M^2$ for different numbers of simulation trials

Index	Number of simulations			
	2000	4000	6000	8000
$i(S_H^1, S_H^2)$	0.81	0.88	0.96	1
$i(S_M^1, S_M^2)$	0.63	0.72	0.93	1
$i(S_H^1, S_M^1, S_H^2, S_M^2)$	0.83	0.86	0.97	1

The similarity indexes for the sets obtained by performing 2000, 4000, 6000 and 8000 simulation trials twice are provided in Table 6. As can be seen from the table, by performing 8000 simulations stable results were obtained, i.e., the same set of the safety culture indicators was obtained by the sensitivity analysis for both simulation data samples. Thus, the sensitivity analysis results for ANSP3 provided in Table 7 obtained by performing 8000 simulations are reproducible.

A *total safety culture sensitivity index* is defined by firstly setting a value 0 for Low sensitivity, a value 0.5 for Medium sensitivity and a value 1 for High sensitivity, and subsequently summing those values over all safety culture indicators that need to be improved for a particular variable. In Table 7 the safety culture sensitivity indexes for the evidence input variables of the ANSP-3 model are provided. Input variables with the total safety culture sensitivity indexes greater than or equal to 4 are considered to be major organizational factors with the greatest influence upon the safety culture indicators that require improvement.

For all other types of parameters in the model, the same sensitivity analysis was performed. Overall, the importance of the other parameters for the total set of safety culture indicators is much lower than the importance of the input variables. Only for one weight, which describes the relation between the commitment of a supervisor to safety and the perception of the commitment to safety in a team, a total safety culture index of 4 is achieved. All other parameters are less important.

Based on the sensitivity analysis of the model for ANSP3, eight Major Organizational Factors (MOFs) with the greatest influence on the organizational safety culture were identified:

Table 7 Total safety culture sensitivity index for the evidence input variables of the ANSP-3 model. The major organizational factors are highlighted in *grey*

Input evidence variable	Total SC sensitivity index
e1 Priority of safety-related goals in the role description	5.5
e4 Influence of a controller on safety activities	5
e7 Sufficiency of the number of safety investigators	6.5
e8 Sufficiency of the number of controllers	8
e9 Availability of reliable and ergonomic technical systems for controllers	5.5
e10 Sufficiency and timeliness of training for changes	4
e11 Regularity of safety meetings	2.5
e12 Developed and implemented SMS	5
e14 Level of development of managerial skills	7
e19 Self-confidence for ATC task	3.5
e20 Commitment to perform ATC task	3
e21 Development level of skills for ATC task	1
e25 Sufficiency of the number of maintenance personnel	3.5
e26 Quality of formal procedures for system checks and repairs	0.5
e35 Intensity of informal interactions in the team of controllers	0
e36 Quality of the formal safety occurrence assessment procedure	0
e40 Quality of the communication channel between controllers and safety investigators	1
e44 Average commitment of the agents involved in the safety analysis	1.5
e61 Individualism index of a controller	0
e62 Power distance index of a controller	0
e63 Masculinity index of a controller	0
e64 Uncertainty avoidance index of a controller	0
e71 Formal support for confidentiality of reporting	0

MOF1: Sufficiency of the number of controllers

MOF2: Level of development of managerial skills of supervisors

MOF3: Sufficiency of the number of safety investigators

MOF4: Priority of safety-related goals in the role description

MOF5: Availability of reliable and ergonomic technical systems for controllers

MOF6: Influence of a controller on safety activities

MOF7: Developed and implemented Safety Management System (SMS)

MOF8: Sufficiency and timeliness of training for changes

Since MOFs exert a significant effect on the organizational safety culture, they can serve as a good basis for identification of organizational improvement options. Based on the MOFs identified for ANSP3, five organizational improvement options (OIOs) were identified:

- OIO1:** *More involvement of controllers in safety assessment for development of new systems and procedures (based on MOF4, 5, 6)*
- OIO2:** *Improve workload of controllers by developing explicit rules for balancing safety and capacity in nominal and non-nominal conditions (based on MOF1, 4)*
- OIO3:** *Improve the quality of management by supervisors (based on MOF2)*
- OIO4:** *Improve coherence and communication in the safety management system (based on MOF3, 7)*
- OIO5:** *Improve the communication about and training for changes (based on MOF8)*

6.3 Model validation

As a basis for the validation, results of the safety culture survey questionnaire for ANSP3 and of the related workshops administered by the EUROCONTROL organization were used. The validation of the model results was performed in two ways (see Fig. 10):

- The level of validity of the results was determined by the comparison of the model-based results and the results of the safety culture survey questionnaire at ANSP3. The questionnaire comprises a set of statements about potential enablers and disablers of safety culture in an ANSP. The questionnaire results include mean scores for the level of agreement of employees of ANSP3 to the statements on a scale from 1 (“fully disagree”) to 5 (“fully agree”).
- The degree of agreement was determined between major organizational factors affecting safety culture indicators and related improvement options which were inferred from a sensitivity analysis of the organizational model, and key issues and related improvement options stemming from the safety culture survey workshop results.

As a basis for the comparison of the obtained model results with the survey results, we introduced three classes for the values of the model results and the survey results: Low, Medium and High.

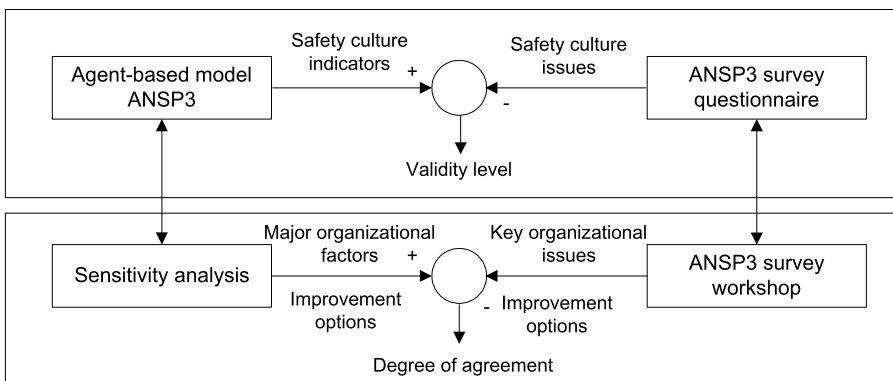


Fig. 10 Schematic representation of the validation process

Table 8 The classified values of the safety culture indicators obtained from the model and the survey questionnaire data

Safety culture indicator	Model	Survey
SCI1: Average reporting quality of controllers	Medium	Medium
SCI2: Average quality of the processed notification reports	High	–
SCI3: Average quality of the final safety occurrence assessment reports	Medium	Medium
SCI4: Average quality of the monthly safety overview reports received by controllers	Medium	Medium
SCI5.1: Average commitment to safety of controllers	Medium	Medium
SCI5.2: Average commitment to safety of a team as perceived by controllers	Medium	Medium
SCI6: Average commitment to safety of a supervisor as perceived by controllers	Medium	Low
SCI7: Average commitment to safety of management as perceived by controllers	Medium	Medium

The model results were mapped to these classes using the Monte Carlo simulations results for full variation of the input variables as presented in Sect. 6.1 and Fig. 9. In particular for each safety culture indicator, the ranges of the class Low followed from the lowest 30% of its values, the ranges of the class Medium followed from the subsequent 55% of the values and the ranges of the class High followed from the highest 15% of the values.

A definition of the class ranges of the survey results had been done before we received the questionnaire results of ANSP-3, on the basis of survey questionnaire results of another (independent) ANSP-2. We analyzed the distribution of these survey results and we chose the class ranges such to achieve a distribution equal to the distribution used for the model results, i.e. Low for the lowest 30%, Medium followed from the subsequent 55% and High for the highest 15%.

Using these mappings, the model results (Table 4) and questionnaire results were classified as shown in Table 8. More details on this validation can be accessed in the companion paper (Stroeve et al. 2011).

The comparison of the class labels for the model- and survey-based safety culture indicator values shows that the results are consistent for six out of seven indicators. Only the result for indicator I6 is lower in the survey than in the model. The indicator I2.1 was not relevant for the survey study.

To determine the degree of agreement between major organizational factors affecting safety culture indicators and related improvement options, results of the safety culture workshop, which was held by EUROCONTROL personnel at ANSP3, were used. The purpose of this workshop was to perform a deeper investigation of the issues identified by the safety culture questionnaire at ANSP3 by conducting interviews with ANSP3's employees. The model results were obtained without any knowledge about the workshop results. The model and workshop results were compared with the assistance of the workshop organizers by identifying for each Major Organizational Factor from the model (from Sect. 6.2) the range of related results of the safety cul-

ture workshop. In such a way, a conclusion about the agreement between the results of the model and the workshop was reached.

It was concluded that from the eight model-based Major Organizational Factors there is some agreement for two factors (MOF1, 4) and good agreement for the remaining six factors (MOF2, 3, 5, 6, 7, 8). Also, the model-based improvement options (from Sect. 6.2) were evaluated. Together with the workshop organizers it was concluded that the model-based recommendations are consistent with the recommendations of the survey study. The latter recommendations tend to reflect the larger detail in the organizational context as has emerged in the workshop at ANSP3. In addition to the list of consistent recommendations, the survey study identified a number of recommendations that are not or only partly addressed in the model-based study. Recommendations that were not addressed reflect aspects that are out of the scope of the model, e.g. on-the-job-training or learning processes of the Engineering department. Recommendations that were only partly addressed mostly reflect aspects for which the organizational context is known in more detail via the workshops at ANSP3.

7 Conclusions

In this paper a new, agent-based approach for structured analysis and improvement of organizational safety culture is proposed. Agent-based modeling and analysis have been used previously to study the efficiency or safety of air traffic scenarios (e.g. Wolfe 2007; Tumer and Agogino 2007; Blom et al. 2006), however agent-based modeling at the organizational level to study safety culture is a new subject both in air traffic management and in the area of multi-agent systems.

The main added value of the proposed approach is that it defines explicit formal relations between safety culture indicators and organizational processes and structures, thus enabling identification of important organizational aspects impacting safety culture by sensitivity analysis techniques.

The agent-based organizational model for ANSP3 was designed using a structured methodology. The model of the formal organization was developed based on the framework from Sharpanskykh (2008). The specification of agents and their behavior were defined based on the identified modeling requirements, on a number of theories from social science and psychology, and on empirical basis from air traffic. The agent architecture proposed provides more extensive tools for modeling cognitive processes of different types than the agent architectures and frameworks based on traditional modeling paradigms (such as BDI (Bordini et al. 2009)). In contrast with the activity-based architectures (such as Brahms (Sierhuis et al. 2007)) agent reasoning is performed based not only on beliefs and world facts, but also using other internal states such as goals, needs, commitments, personal traits. Moreover, the reasoning models described in the paper represent formalized theories from psychological and sociological literature and are not limited to deduction procedures from formal logics. As has been demonstrated in Byrne and Kirlik (2005), the cognitive architecture ACT-R, which constitutes a unified theory of cognition, can be used for modeling low level operations of air traffic controllers. However, ACT-R is less suitable for specification of processes of the social level and of aggregated organizational concepts

and processes, which are of importance for our study. As a result many of the safety culture indicators from the study cannot be related to ACT-R modeling concepts. Furthermore, ACT-R models do not scale well, which would constitute a computational problem for our study.

Every modeling step was based on empirical data (formal documents, interviews, questionnaire results) provided by EUROCONTROL and ANSP3. In such a way the correctness of many assumptions underlying the model of ANSP3 was ensured as much as possible.

The validation of the model results was performed in two ways. The comparison of the model-based safety culture indicators with the survey-based safety culture issues showed that mostly valid model results have been obtained. Next it was found that, the model results obtained based on the sensitivity analysis are mostly consistent with the results of the safety culture survey workshop. Both the model and the workshop used outcomes of the survey questionnaire as input, but the processes for achieving their results were completely different and independent. As such, the consistency in both validation phases is a good indication of the model's validity. The validity of the model results depends greatly on the validity of the estimation of the model inputs. Our experience shows that the formal organizational documentation and interviews provide a limited basis for such estimation. The highest validity level of the estimation was achieved with the use of anonymous survey data from ANSP3, different from the ones used for the model analysis and validation.

The limitation in the detail of the organizational context and the scope of the model in comparison with the survey-based results, on the one hand, a fundamental modeling issue, in the sense that a model is always an abstraction of reality. On the other hand, the range of organizational aspects that are considered in detail in the model may be enhanced. In the current study, the model development was focused on the occurrence reporting cycle and other processes such as management actions, engineering activities and traffic management actions by controllers were modeled at a high (abstract) level.

Safety culture professionals from EUROCONTROL recognized a high potential of the proposed approach. In particular, the approach may further enhance safety culture questionnaires and be used to prepare safety culture survey workshops. In the future the approach may be applied for safety culture analysis of ANSPs from other national cultures.

Furthermore, another promising direction for future research is to integrate the developed approach into existing agent-based approaches such as Tumer and Agogino (2007) and Blom et al. (2006), which address systems at the operational level and do not consider the organizational and cultural contexts. In such a way explicit reciprocal relations between organizational structures, processes and safety culture on the one side and operational safety indicators (e.g., risks) on the other side can be defined. The identification of such relations enables understanding and profound analysis of the system behavior at different aggregation levels (e.g., individual, team, organization).

Acknowledgements This research was supported by the EUROCONTROL CARE Innovative III project. We are grateful to Barry Kirwan, Marinella Leone and Eve Grace-Kelly of EUROCONTROL for their cooperation and contribution to the safety culture model workshop, and to the interviewees at ANSP-3 and EUROCONTROL Head Quarters for their cooperativeness.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Bedford T, Cooke RM (2001) Probabilistic risk analysis: foundations and methods. Cambridge University Press, Cambridge
- Blom HAP, Stroeve SH, De Jong HH (2006) Safety risk assessment by Monte Carlo simulation of complex safety critical operations. In: Redmill F, Anderson T (eds) Developments in risk-based approaches to safety. Springer, London, pp 47–67
- Bordini RH, Dastani D, Dix D, El Fallah Seghrouchni A (eds) (2009) Multi-agent programming: languages, tools and applications. Springer, Berlin
- Bosse T, Jonker CM, Meij L, van der Treur J (2007) A language and environment for analysis of dynamics by simulation. *Int J Artif Intell Tools* 16:435–464
- Brazier FMT, Jonker CM, Treur J (2000) Compositional design and reuse of a generic agent model. *Appl Artif Intell J* 14:491–538
- Burt RS (1987) Social contagion and innovation: cohesion versus structural equivalence. *Am J Sociol* 92(6):1287–1335
- Byrne M, Kirlik A (2005) Using computational cognitive modeling to diagnose possible sources of aviation error. *Int J Aviat Psychol* 15(2):135–155
- Ek A, Akselsson R, Arvidsson M, Johansson CR (2007) Safety culture in Swedish air traffic control. *Saf Sci* 45(7):791–811
- Endsley M, Garland D (2000) Situation awareness, analysis and measurement. Lawrence Erlbaum, Mahwah
- EUROCONTROL/FAA AP15 (2008) Safety culture in air traffic management: a white paper (December 2008)
- Fisher M (1996) A temporal semantics for concurrent METATEM. *J Symb Comput* 22(5):627–648
- Goldberg LR (1993) The structure of phenotypic personality traits. *Am Psychol* 48:26–34
- Goldman AI (1967) A causal theory of knowing. *J Philos* 64(12):357–372
- Griffin RW, Bateman TS (1986) Job satisfaction and organizational commitment. In: Cooper CL, Robertson I (eds) International review of industrial and organizational psychology 1986. Wiley, New York, pp 157–188
- Hersey P, Blanchard KH, DE Johnson (2001) Management of organizational behavior: leading human resources
- Hofstede G (2005) Cultures and organizations. McGraw-Hill, New York
- Hopkins A (2006) Studying organizational cultures and their effects on safety. *Saf Sci* 44:875–889
- Horii T, Yan J, Levitt RE (2004) Modeling and analyzing cultural influences on project team performance. *J Comput Math Organ Theory* 10(4):305–321
- Janis I, Mann L (1977) Decision making: a psychological analysis of conflict, choice, and commitment. The Free Press, New York
- Kim J (1996) Philosophy of mind. Westview Press, Boulder
- Kunz JC, Levitt RE, Jin Y (1998) The virtual design team: a computational simulation model of project organizations. *Commun ACM* 41(11):84–92
- Lee S, Ravinder U, Johnston JC (2005) Developing an agent model of human performance in air traffic control operations using apex cognitive architecture. In: Proceedings of the 2005 winter simulation conference, INFORMS college on simulation, December 4–7, Orlando, Florida, USA, pp 979–987
- Leveson N, Dulac N, Zipkin D, Cutcher-Gershenfeld D, Barrett B, Carrol J (2005) Modeling, analyzing, and engineering NASA's safety culture. Phase 1 Final Report
- Locke EA (2001) Motivation by goal setting. In: Handbook of organizational behavior, vol 2. Springer, Berlin, pp 43–54
- March JG, Olsen JP (1975) The uncertainty of the past: organizational learning under ambiguity. *European Journal of Political Research* 3:147–171
- Meinolf D, Berthoin Anthal A, Child J, Nonaka I (2001) Handbook of organizational learning and knowledge. Oxford University Press, Oxford
- Mohaghegh Z, Mosleh A (2009) Incorporating organizational factors into probabilistic risk assessment of complex socio-technical systems: principles and theoretical foundations. *Saf Sci* 47:1139–1158

- Pinder CC (1998) *Work motivation in organizational behavior*. Prentice-Hall, New York
- Popova V, Sharpanskykh A (2008) Process-oriented organisation modelling and analysis. *Enterp Inf Syst J* 2(2):157–176
- Robbins SP (2004) *Organizational behavior—concepts, controversies, applications*, 4th edn Prentice Hall, New York
- Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S (2008) *Global sensitivity analysis: the primer*. Wiley-Interscience, New York
- Searle J (1969) *Speech acts*. Cambridge University Press, Cambridge
- Sharpanskykh A (2008) *On computer-aided methods for modeling and analysis of organizations*. PhD thesis, VU Amsterdam
- Sharpanskykh A (2009) Individual decision making in a learning organization. In: *Proceedings of the 4th international conference on intelligent computing and information systems*. ACM Press, New York, pp 535–544
- Sharpanskykh A, Stroeve S (2010) Can we predict safety culture? In: *Proceedings of the 9th international conference on autonomous agents and multiagent systems, AAMAS'10*, ACM Press, New York, pp 1739–1746
- Sharpanskykh A, Treur J (2010) Automated analysis of compositional multi-agent systems. *Int J Agent-Oriented Softw Eng* 4(2):174–221
- Sierhuis M, Clancey W, van Hoof R (2007) Brahms: a multiagent modeling environment for simulating work processes and practices. *Int J Simul Process Model* 3(3):134–152
- Stroeve S, Sharpanskykh A, Kirwan B (2011) Agent-based organizational modelling for analysis of safety culture at an air navigation service provider. *Reliab Eng Syst Saf*. doi:10.1016/j.ress.2010.12.017
- Tumer K, Agogino A (2007) Distributed agent-based air traffic flow management. In: *6th international joint conference on autonomous agents and multiagent systems (AAMAS 2007)*, Honolulu, Hawaii.
- Uttal B (1983) The corporate culture cultures. *Organ Cult J* 17:66–72
- van de Walle D (1997) Development and validation of a work domain goal orientation instrument. *Educ Psychol Meas* 57:297–301
- Wolfe SR (2007) Supporting air traffic flow management with agents. In: *AAAI spring symposium: interaction challenges for intelligent assistants*, Stanford, CA



Alexei Sharpanskykh is a research fellow at the Agent Systems Group, VU University Amsterdam. He received his PhD degree at the VU University Amsterdam in the area of Artificial Intelligence. Currently he is doing research in agent-based modeling and analysis of complex socio-technical systems in the context of a number of projects in the areas of air traffic, incident management, and logistics.



Sybert H. Stroeve was born in Hardenberg, The Netherlands in 1967. He holds a MSc degree in Electrical Engineering from University of Twente, Enschede, The Netherlands (1991) and a PhD degree in Mechanical Engineering from Delft University of Technology, Delft, The Netherlands (1998). He did research at Delft University of Technology on modelling of the neuromuscular control of human arm movements and at Radboud University Nijmegen on neural network systems. Currently he is Senior Scientist at the Air Transport Safety Institute of National Aerospace Laboratory NLR, Amsterdam, The Netherlands. In this position he has conducted research in human performance modelling, organizational modelling and dynamic risk modelling for multi-agent environments. He has applied this research in a range of safety assessment studies for air traffic applications.