Cooperation under Guidelines

How Guidelines for Cooperation Affect Interaction Behavior of Airport Experts

Anne Papenfuß
Institute of Flight Guidance
German Aerospace Center (DLR)
Braunschweig, Germany
anne.papenfuss@dlr.de

Anna Biasotto
Institute of Psychology
Universität Heidelberg
Heidelberg, Germany

***Abstract*— Collaborative decision making is a key component for novel air traffic management concepts. They should enable pro-active decision making that is flexible enough to take into account needs and priorities of several stakeholders. Experience from training of cockpit teams shows the benefit of models structuring the decision making process in terms of decision quality. Guidelines for cooperation were developed to enhance decision making in two exemplary airport management tasks. Qualitative interaction analysis was conducted to investigate the impact of these guidelines on interaction behavior of three teams of airport experts. Results show, beside strong inter-team differences, that guidelines have the potential to focus team decision making on more thorough situation assessment. The paper proposes metrics to analyze and quantify compliance of teams with a given process. Furthermore, ideas for advancement of the guidelines are derived.**

Keywords-human factors; collaborative decision making; interaction analysis; working methods

I. INTRODUCTION

Collaborative Decision Making (CDM) is a key component for European air traffic management (ATM) concepts. In airport management, through collaborative decisions the goals and constraints of several stakeholders should be considered in decisions made at the day of operation. The expected benefit of CDM processes is that decisions found take into account more constraints thus serve better the circumstances of the situation and therefore are of higher quality. Consequently, CDM processes are an enabler for more performant airport operation.

Beside airport management, other operational areas could benefit from the positive effect of “higher-quality decisions” and greater flexibility through collaborative decision making. For example, within the Pilot Common Project of SESAR Joint Undertaking [1], CDM is named as component for enhancing airport throughput, system wide information management and integrating network management into airport management

In today’s ATM system, the working methods of operators are characterized by standard operating procedures. The rationale is that standard operating procedures guarantee safety by providing predictability and minimizing influences of personality.

In contrast to this, working methods like collaborative decision making are described as being more flexible in order to react upon specific situations and to take into account needs and priorities of other stakeholders. Those features contradict to some extent the idea of rather predictable, inflexible standard procedures. At least, ideas have to be developed on how to integrate the two work design philosophies.

Summarizing these trends, we state a need for research looking into the working procedures required for those flexible and collaborative decision making concepts. In the domain of airport management, substantial experience was gathered. For instance, state of the art working processes at airports were analyzed via job-shadowing. Furthermore, initial ideas for guidelines for cooperation were presented to experts.

The adjustment of plans within an operations center between stakeholders with different goals was focus of the project “Collaboration within Control Centers (COCO)”, financed and executed by the German Aerospace Center DLR. A simulation study was conducted to assess the influence of guidelines for decision making on perceived quality of the decision making process. This paper explores the theoretical foundation of these guidelines and the observable impact on the decision making process of airport experts. Furthermore, metrics to analyze and quantify compliance of teams with a given process are proposed.

II. BACKGROUND

A. Collaborative Decision Making in Airport Performance Management

Airport operations have a large potential for optimization and thus influence performance of overall ATM. Airports are complex systems with multiple interconnections between numerous processes owned by a multitude of stakeholders [2] Each stakeholder at an airport plans its processes and actions according to individual goals and standards and corporate business plans. But most stakeholders miss information about intentions, goals and actions of other (cooperating or competing) parties at the same airport. Relevant information is not available, available but incorrect or available but too late (cf.[3]).

Hence, harmonizing plans between different stakeholders at an airport is rather time consuming and difficult, especially regarding partly conflicting goals of airport stakeholders and their unwillingness to share all information about their plans. Assessing the impact of other parties actions on one´s own plan is therefore difficult and an integrated view of total airport operations is missing, cf.[2].

Airport collaborative decision making (A-CDM) was developed [4], to foster a more proactive behavior and share a minimum set of relevant data, like the Target Off-Block Time, TOBT, between all stakeholders. Performance benefits of A-CDM could be demonstrated, as more and more European airports adopt the concept [5].

This concept was further developed to a solution called Total Airport Management (TAM) [6]. TAM enhances the airside-focused A-CDM concept by integrating landside processes and developing ideas for highly collaborative decision making in Airport Operation Control Centers (APOC) (cf.[7]). Within the context of the Single European Sky ATM Research Program SESAR, processes and use cases for airport operations control centers (APOC) were developed and validated (cf.[8-10]).

Up to now most research focused on technical solutions for the socio-technical system APOC (e.g.[2, 11]). Ideas to foster pro-active, collaborative behavior mainly focused on the competitive roles of several airlines and involved the development of negotiation protocols and bonus-malus-systems [12]. Research on Performance Based Airport Management could show that airport experts in general see the potential for collaboration but they see the need for a mandatory framework or rules for collaboration and cooperation [3]. As one step towards feasible working procedures within an APOC, guidelines for cooperation were developed and evaluated in a high-fidelity simulation with airport experts [13]. The guidelines aimed at improving the flow of relevant information between stakeholders in order to enable each participant in the APOC to adapt his/her plans according to the traffic situation.

B. Models of Decision Making in Aviation

Decision Making is daily business in ATM. Good decisions are the basis for safe and performant air traffic. Laboratory and field studies have shown that human decision making behavior is not necessarily rational. Especially under time pressure, humans tend to make ad-hoc decisions guided by expectations and preferences, they tend to stick to sometimes inadequate goals and follow heuristics instead of analyzing a situation in depth.

Orasanu [14] identified for her descriptive model of aviation decision making the two major components “situation assessment” and “choosing a course of action”. Situation assessment includes the definition of the problem, assessing the risks associated with it and the time available for solving the problem.

The selection of actions distinguishes between the application of rules, the choice between several options as well as the creation of novel solutions. The course taken depends on the understanding if the situation, how much time is available as well as the availability of rules, all together called the situational constraints and affordances of the situation. With this model, observed decision making behaviors of pilots, e.g. during flight accidents, could be explained. For instance, if situation assessment was not sufficient pilots chose an action based on an ill-defined problem which might contribute to the fatal outcome of an accident.

To enhance decision making competencies of pilots, prescriptive models for crew training were developed. They should provide the crew with a structure for the decision making process which suits better to the complex socio-technical environment of the cockpit.

One model developed from Lufthansa and the German Aerospace Center, is the FOR-DEC model [15]. The model distinguishes the six phases Facts, Options, Risks and Benefits, Decision, Execution and Check. It was developed to structure the judgement and decision making processes within the cockpit and was used to train crew resource management (CRM). It was developed to take into account the complexity and dynamics of the cockpit environment, to be applicable to a wide range of situations, and to separate the phases of collecting information about the situation and evaluate possible solutions. The model can be run in several cycles until the desired goal is reached, furthermore sub-cycles and feedback loops are possible. The phases Facts, Options and Risks can be subsumed as “situation assessment”, the Decision, Execution and Check-Phases can be mapped to the “course of action” stage of the ADM model.

In team exercises, the authors found that the application of the FOR-DEC model helped teams to structure their communication and group interaction processes [15]. FOR-DEC is widely known in the pilot community but pilots report that in real life situations, especially under time pressure or where options are clear, following the model feels artificially [16].

C. Guidelines for Collaborative Decision Making

Based on the experience made with decision making in the cockpit, as well as the results from job-shadowing, the need for a structure for the collaborative decision making process in an APOC was postulated. Guidelines were developed to provide the team with a workflow that incorporates the findings of the CRM domain with respect to separate situation assessment and decision making phases. Additionally, for each team member expected activity within the phases was highlighted. The rule for this assignment was that stakeholders who inherit the relevant and reliable information for this phase must provide this information at that time to the other stakeholders. For example, the groundhandler has information about the maximum capacity for turn arounds and thus is able to analyze whether an event will cause an over-demand and thus delays.

On a more abstract level, the guidelines were designed to 1) provide the relevant information at the right time and 2) create transparency on dependencies between stakeholders’ individual planning. In contrast to the cockpit environment, time pressure and risk assessment are not a major issue in the APOC environment. But additional "affordances" exists, as the choosing of actions is influenced by conflicts between the stakeholders, especially conflicts of goals and conflicts of power.

Six phases were differentiated. They are depicted in Figure 2. The guidelines were first described in terms of information and decisions required within each phase. Afterwards, the required actions per stakeholder were broken down for these phases to generate instructions per phase. It must be highlighted, that the guidelines tested in this study were developed for two specific tasks and therefor do not claim to have universal validity. The two tasks are "prioritizing departures" and "stand- and gate allocation".

First, information about the event that might disturb the airport operations is received (phase 1). Next, dependent on the task, either airport (AP) or groundhandler (GH) analyses the impact on his/her operations (phase 2) as their resources are the bottleneck. Phase one and two can be mapped directly to the situation assessment step of the ADM.

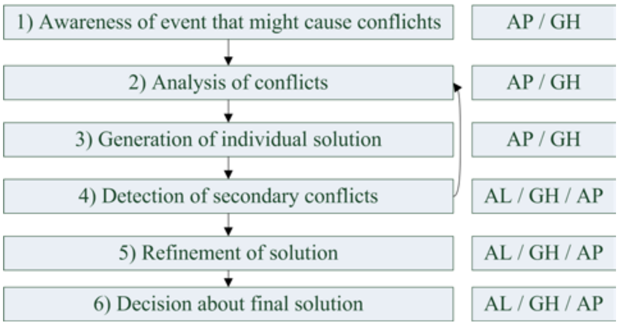

Figure 1. Phases of the guidelines for cooperation and activity of stakeholders

Following, principal constraints for the final solution are elaborated by going through the phases 3) generation and distribution of a first, individual solution by either groundhandler or airport and phase 4) the evaluation of the impact of this solution on each stakeholders plan. Within this phase the Airlines (AL) are required to check the influence on their operations. Phase four might trigger a new task, so there is a feedback to phase 2. For instance, the airport might conclude that s/he should start a stand and gate allocation task because the impact of departure prioritization is too big. Phases five and six can be related to the course of action, as the solution is refined to get rid of the problems detected in phase 4. This might require that phase 5 is run several times. Finally, phase six finishes the process by decision of all stakeholders to agree on a final solution.

III. RESEARCH QUESTIONS

Experience with prescriptive models for aviation decision making in the cockpit showed the potential of these models to structure the teams decision making behavior. Thus, it was of interest, to what extent the proposed guidelines for an APOC team influenced the collaborative decision making process. Results from a simulation based evaluation of the guidelines show that operators rated efficiency of the decision making process when following the guidelines significantly better compared to an unstructured process [13]. The guidelines provided effective procedures to guide team functioning, provided the teams with clear agreements about how decisions were made so that the teams worked constructively on issues until they were resolved [13]. Beside these positive results, additional analysis was needed to understand whether teams actually behaved in conformance with these guideless or came up with ad-hoc adaptions of the prescriptive decision process.

An exploratory and descriptive analysis of two research questions leads the following analysis:

We are interested to see, to what extent the proposed phases of the guidelines for cooperation manifested in the observable interaction behavior of the APOC teams. First, do teams follow the proposed transition of phases and the proposed activity patterns of stakeholders?

Second, are there distinctive features that can be observed when teams apply the guidelines for cooperation compared to an unstructured decision making process?

IV. METHOD

1) *Data Collection*

A high fidelity simulation in the Airport and Control Center Simulator (ACCES) facility of DLRs Institute of Flight Guidance was conducted. Experts from german speaking airports participated in this study. The set-up used in this study is described in detail in [13]. A generic airport simulation [17, 18], build upon Eurocontrol's A-CDM implementation standard [4], was used to create three different scenarios where an event disturbs the scheduled Airport Operations Plan (AOP). In order to achieve best possible punctuality, two airlines, one groundhandler and the airport representative collaboratively decide on a new AOP. Within 45 minutes the TOBT of 34 flights in a two hour time window need to be adjusted so that punctuality is improved and the individual goals of all stakeholders are reached.

Four teams of four airport experts participated in this study, each running through three scenarios. Within the first two runs, participants were asked to make their decisions based on their experience and individual working methods. For the third run, participants were introduced to the guidelines, trained and asked to follow these guidelines in their decision making process. The material used in this study consisted of the second and third runs of three teams. Thus, the sample consists of one run with free structure (free) and one run with guidelines (guided) of each group.

All interactions between participants were present in the form of audio files. During the experiment, each participant was equipped with a microphone that recorded his interactions. Using the audio-editor Audacity, the four soundtracks were time synchronized and put into a shared project file.

B. *Data preparation.*

The steps undertaken to prepare the raw audio data for data analysis are summarized in Figure 2.

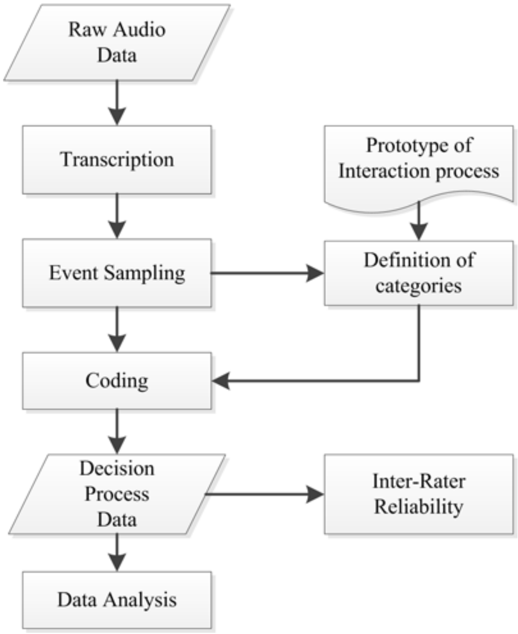

Figure 2. Process of data preparation from raw audio data to decision process data.

Transcription. All utterances were transcribed completely and literally, including incomplete sentences and repetitions. Nevertheless, filler words and hecklings were omitted because the focus of the analysis was on the content. Dialect (pronunciation as well as choice of words) was adapted to standard German. Breaks and moments of silence were marked by a dash. Straying from the topic, typing, making jokes as well as questions raised to the experiment supervisor or the observers (e.g. regarding the interface) were not transcribed content wise. Instead, only the action was noted. Other peculiarities such as laughing, a joking or ironic tone, mumbling or loud thinking were mentioned in brackets and italic font after the content [19].

Event sampling. The course of conversation was divided into events whenever the speaker, the addressee or the topic changed, following approaches from discussion coding [20]. Therefore, consent or rejection of an utterance was classified as a separate event. Furthermore, the starting time in minutes and seconds was noted for each event. These data were arranged in the form of a table with one row per interaction and the columns time, content, speaker and addressee.

Prototype of Interaction process. As the guidelines were present in the form of individualized ‘checklists’ for each individual stakeholder, they had to be combined into one general process flow. The chronological order of the expected interactions within each phase was determined. This task was conducted for each of the two tasks “prioritizing departures” and “stand and gate allocation” and resulted in two prototypical flow charts of the interaction process.

Definition of categories. Initially, all six phases of the guidelines were established as categories. In order to evaluate this concept and determine further possible categories, one run was chosen and worked through. Thereby, the following final coding categories were established: phase 2 (conflict detection), phase 3 (generation of initial solution), phase 4 (identification of subsequent conflicts), phase 5 (optimization of solution), priority Flights (P), system (S) and other (R).

The categories ‘phase 2’, ‘phase 3’, ‘phase 4’ and ‘phase 5’ correspond to the phases of the guidelines. The original categories ‘phase 1’ (information about event) and ‘phase 6’ (implementation of final solution) were excluded, as no verbal interactions were found pertaining to these two categories. This finding was supported by the fact that the prototype of the interaction process does not intend any interaction in phases 1 and 6. In phase 1, the stakeholder with the bottleneck receives external information about the event via the system. Thus, no interaction with other stakeholders is required. Phase 6 starts, when the final solution has been agreed upon. In this phase, the final solution is documented in the system. Thus, no interaction between stakeholders is required.

The category ‘priority flights’ was introduced, because this was a topic of relevance in each of the six runs. As this topic was not in line with any of the phases, it was encoded as a separate category. Events, in which priority flights were identified, were assigned to this category. However, all events that addressed measures for priority flights were allocated to the respective phase of the guidelines.

Coding. In the next step, all events were assigned to a category. Each event was compared to the prototype of the interaction process and allocated to the phase, which resembled this event the most. In order to ensure internal consistency of the process of category allocation, certain rules were defined: First, an event could only be allocated to phase 3, if the stakeholder with the bottleneck was the one compiling an initial solution. This would be groundhandler or airport, depending on the specific subtask. No suggestions or requests by other stakeholders are allowed in this phase. Second, responses to the initial solution (phase 3) were characterized as phase 4, when the focus was on a general problem analysis. However, suggestions and requests pertaining to specific aircraft were assigned to phase 5 (optimization). Third, when a conflict was induced by a decision or system input, this event was allocated to the category phase 4 (analysis of induced conflicts). Fourth, after the interaction was in full swing, it could only go back to phase 2 (conflict detection) and phase 3 (initial solution), if the stakeholders switched to the other task. This is due to the fact that phases 2 and 3 are characterized as initial reactions to the problem at hand. Last, the content was prioritized over the speaker. For instance, if an airline took on a moderating role and suggested that the groundhandler should now work on an initial solution (phase 3), this event was allocated to phase 3, even though the Airline was not supposed to be active in this phase. The psychometric properties of the process of coding were examined by means of stability (intra-rater reliability) and reproducibility (inter-rater reliability).

Krippendorff's alpha (α) amounted to .91 for the intra-rater reliability, which measured stability.

C. Data Analysis

The interaction process of the six runs was analyzed qualitatively. Furthermore, metrics were developed to derive quantitative results representing the conformance of the decision making process with the proposed guidelines for cooperation and to quantify features of the decision making process.

Phase progression. Each event was allocated to one of the guideline's phases. Thus, the chronological sequence of a runs' events depicts the phase progression of this run. For each run, the phase progression was displayed graphically. In addition, several sections of each run were further analyzed. Here, the main focus was on the initial part of the interaction, sections where the task switched and long phases of optimization. For these sections, it was examined whether the phases were completed in the linear order intended by the guidelines or whether the interaction oscillated between different phases.

Compliance of transitions with guidelines. The guidelines specify which phase transitions are allowed (guideline-consistent) and which are prohibited (non-consistent with guidelines). For each run, the absolute and relative amount of prohibited phase transitions was determined. For each group, a Pearson's Chi-square test [21] was computed with IBM SPSS Statistics 24 to test if the type of run (free vs. guided) and transition compliance were significantly associated. All values exceeded the necessary case number of five observations.

Focus of decision making process. For each run, the number of events pertaining to phases 2, 3 and 4 were summed. Then, for each run, the portion of events pertaining to these three phases was compared to the portion of events pertaining to phase 5. For each group, a Pearson's Chi-square test [21] was calculated. All observed values exceeded the critical value five.

V. RESULTS

A. Qualitiative Analysis of Phase Progression

Due to the limited space in this paper, a detailed description of the decision making process is given for team two only. It has to be mentioned that the detailed phase progression differed between all teams. Nevertheless, commonalities between all groups could be found that are reported subsequently.

Phase progression and communication activity of the free structured decision making process is depicted in Figure 3; Figure 4 shows the process with guidelines. The x-axis codes the time of events and the y-axis shows the phase coded for the events. Furthermore, grey dots visualize the speaker of the event. Topmost are events from the Airport (AP), followed by groundhandler (GH) and the two airlines (AL).

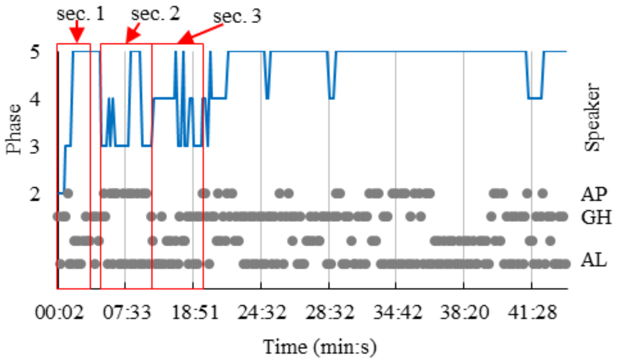

Figure 3. Phase progression and communication activitiy of run "free structure"

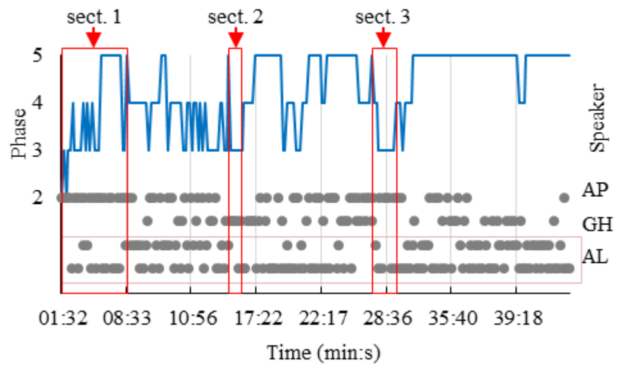

Figure 4. Phase progression and communication activitiy of run "guidelines"

First, it can be observed that in both conditions phase progression contains several loops from phase 3 to 5. In the freely structured decision making process (Figure 3), three sections are analyzed in detail. In section 1, after conflict detection (phase 2), both airlines rescheduled their flights. These actions were interpreted as optimization measures for specific flights and classified as phase 5. After the airlines concluded their actions, the groundhandler began to solve his conflicts (section 2). Then, however, it was decided to have the airport solve his induced gate conflicts first. Hence, the groundhandler was skipped and the task switched to planning gates and stands. Whilst developing the initial solution, the airport was interrupted several times by the airlines making offers to swap flights or by adjustments of the TOBT of a flight. These two examples of airline interference explain why the graph of the interaction detoured to phase 5 during section 2.

Following, the groundhandler got to resolve his conflicts (section 3). Thus, the task changed to departure prioritization and the interaction went back to phase 3. However, while the groundhandler was compiling an initial solution, the airlines made several inquiries, for instance suggesting a specific flight. In addition, the two airlines compared their number of conflicts (phase 4) and discussed specific actions for optimization (phase 5). From minute 22 that decision making process was characterized by a long-lasting period of phase 5 which was interrupted three times by short periods of phase 4. This pattern of long optimization phases was found in all runs analyzed in this study.

With guidelines applied by the team, the decision making process of the stakeholders was structured as depicted in Figure 4. Three sections are highlighted. In section 1, the airport detected conflicts (phase 2) and compiled an initial solution (phase 3). Here, phases 2 and 3 alternated, because the airport took an iterative approach: He detected a conflict, solved it and communicated the solution and then checked for further conflicts. Whilst the airport was working on the initial solution, the airlines already disclosed their conflicts. For instance, one airline communicated an induced conflict for the priority flight, an action typical for phase 4. While working on the initial solution, the airport asked the airlines for their approval to change a gate. This was interpreted as an optimization procedure, as it targeted a specific flight (phase 5). These two examples explain why the graph of the phase progression (see Figure 4) oscillated between initial solution (phase 3), analysis of induced conflicts (phase 4) and optimization (phase 5). Nevertheless, it should be mentioned that the airport insisted several times on going back to concluding the initial solution (phase 3

In the second phase, the groundhandler worked on an initial solution for the task departure prioritization. Thus, the phase progression went back to phase 3. Here, the moderating role of one airline became apparent, who directed this approach. Whilst the groundhandler was compiling the initial solution, the airlines discussed measures for optimization, but decided not to interfere. This was identified as a discussion about strategy, as it broached the issue of how to proceed.

During the optimization, one of the groundhandler's changes (min 25:08) entailed severe conflicts for the other stakeholders (section 3). Instead of optimizing this situation, the airport was asked to reinstate his initial solution (min 27:19), leading to a repetition of phase 3. In this situation, the moderating role of one airline could be observed, who made two propositions regarding strategy.

Summarizing, in the run with free structure, the first actions were taken by the two airlines that changed their TOBTs. In contrast, in the run with guidelines, the first actions were taken by the airport, who had the best overview of the capacity bottleneck. In the run with free structure, there was initial confusion on how to proceed. Stakeholders were uncertain, whether airport or groundhandler should start with an initial solution. In contrast, in the run with guidelines, all actions were implemented consecutively and in the order intended by the guideline.

B. Quantitative Metrics

1) Compliance with proposed phases of the guidelines

No global difference was found between the two types of runs (free vs. guidelines) in terms of the number of guideline compliant phase transitions. Descriptive data of number of events and percentage of compliant transitions can be found in Table 1 together with further descriptive data of the teams' interaction process.

TABLE I. OVERVIEW ON DESCRIPTIVE DATA OF DECISION MAKING PROCESS

	team 1		**team 2**		**team 3**	
	free	***guide***	***free***	***guide***	***free***	***guide***
n events	229	265	226	235	212	205
n evt. phase 2-4	65	74	70	115	48	69
n evt. phase 5s	164	191	156	120	164	136
% of compliant trans.	90.8	96.1	93.3	88.5	91.0	91.2
evt. in phase 5 [%]	71.6	72.1	73.0	51.1	77.4	66.3

In all runs, more than 88% of all transitions were compliant with the guidelines, even in the unstructured decision making processes. The average percentage of compliant transitions is virtually identical, with 91.71 % in the unstructured runs and 91.95% in the runs with guidelines.

TABLE II. CHI-TEST STATISTICS FOR COMPLIANCE OF TRANSITIONS WITH GUIDELINES

	χ^2	**p**	**odds ratio**
team 1	6.09	.011	2.58
team 2	3.27	.049	0.55
team 3	.004	.543	1.05

Introducing the guidelines had a different effect in all three groups. Chi-Square tests were calculated, as well as odds ratios, representing the likelihood that for each team in runs with guidelines transitions are compliant with the guidelines. In groups one and two, there was a significant association between the type of run and whether or not a phase transition was guideline conforming. However, the nature of the relationship was opposite. In group one, the odds of a phase transition being guideline-consistent was higher when guidelines were present. In group two, the odds of a phase transition being guideline-consistent was lower when guidelines were present. In group three, no significant association between type of run and whether or not a phase transition was guideline-consistent, was found. Consequently, accumulated over all phase transitions of the three groups, no association between the type of run and whether or not a phase transition was guideline-consistent was found $\chi^2(1) = .093$, $p = .419$.

2) Focus of decision making process – Long optimization phase versus Situation Assessment

Each of the six runs concluded with a long-lasting period of phase 5 with correction loops. No general difference regarding the length of this section was found. While in groups two and three this section started later when guidelines were present (group two: min 22:03 vs. 25:08; group three: 12:02 vs. 19:05), the opposite was true for group one (min 15:48 vs. 13:56).

Regarding the distribution of interaction events between the different phases, in all runs the majority of events were categorized as "optimization of solution" (phase 5), raning from 51% to 77%. The absolute and relative numbers can be found in Table 1. Chi-square tests were calculated to assess if guidelines had an effect on the amount of interaction events in the optimization phase itself. The number of events in phase 5 was compared to summarized number of events in phases 2, 3 and 4. Parameters of the test for each group are summarized in table 3. Odds ratio refer to the likelihood that an event did not belong to phase 5 in decision making process with guidelines compared to unstructured decision making process. No significant effect was found for team 1 where the length of phases was not affected by the guidelines. A significant effect was found for teams 2 and 3. When applying the guidelines, in both teams more interaction events where categorized as belonging to phases 2, 3 and 4 of the guidelines for cooperation.

TABLE III. CHI-TEST STATISTICS FOR AMOUNT OF EVENTS IN PHASE 5 "OPTIMIZATION OF SOLUTION"

	χ^2	p	odds ratio
team 1	.013	.494	1.02
team 2	15.47	<.001	2.14
team 3	6.27	.008	1.73

VI. DISCUSSION

This study examined to what extent airport experts implemented a proposed structure (guidelines for cooperation) in their decision making process. An approach for the collection and analysis of interaction behavior was presented. Qualitative and quantitative analysis was conducted to get insight into the decision making processes.

A. Influence of guidelines on interaction behavior

The qualitative analysis of the phase progressions revealed that teams in an unstructured decision making process discussed on prioritization of tasks and oscillated between overall adjustment of plans (phase 3) to detailed refinement of single flights (phase5). When applying the guidelines, they did not follow the process consequently but content analysis showed they were aware of the general structure and priority of the tasks to be conducted in phase 3. Nevertheless, with regard to the quantitative results derived, the impact of guidelines was not as visible as expected.

First, some features of the expected interaction process under guidelines could also be observed in freely structured runs. Secondly, teams did not strictly follow the proposed process of the guidelines. Third, interaction behavior analysis revealed strong team differences. Regarding the first argument, the phases of the proposed guidelines could also be identified in the runs with free structure. The steps of decision making process, especially phases 3, 4 and 5, seemed to be intuitive and were also followed when guidelines were not present. This is positive, as the activities proposed by the guidelines are not completely artificial to the expert.

The guidelines mainly influenced the sequence, as they were designed for. Nevertheless, the quantitative metric of conformant transitions was not sensitive to these changes.

When experts were not guided in their decision making process, refinements (phase 5) were conducted very early or even where the starting point for decision making. Mainly, Airlines pushed this phase by providing information or adaptions for their priority flights. In contrast to this, in runs with guidelines the decision making process started in conformance with the guidelines. The observed behavior in the "unguided" interaction process shows that decision making in natural environment tends to focus on heuristics and come up with quick initial solutions. This finding is in line with experiences made in crew resource management.

Additionally the perspective on the problem was dominated by the airlines; which reflects the situation at airports where airlines are the major customer and thus have a lot of power. The guidelines do not propose to change this situation, but the analysis of the situation should also include the perspective of the other stakeholders in order to find a solution in the APOC which increases overall airport performance. In two of three teams, differences in the decision making process in terms of increasing the focus on situation assessment was found. This is result is quite promising, as the teams tested the guidelines for the first time.

B. Metrics to analyze and quantify compliance of teams with a given process

Phase progression of the decision making process was used to analyze the teams' interaction process. Content analysis revealed causes why teams left the proposed sequence of the phases. The insights gained by this analysis can be used to advance the guidelines in a next step. Furthermore, quantitative metrics were calculated, like percentage of compliant transitions and odds ratio which describes the likelihood of interaction within a certain phase. These metrics can be used to quantify the conformance of team behavior with proposed work procedures. With regard to teamwork and team process, research so far did not produce widely accepted methods and metrics to assess these processes. We demonstrated how team process metrics can be derived and used for statistical analysis by analyzing the interaction process.

There are some limitations with regards to metrics and quantitative analysis. First, the sample size of this study is a limit. Second, the duration of the runs was fixed to 45 minutes. It can be assumed that guidelines lead to shorter time needed to come to a decision every stakeholder agrees with. In this study we could not analyze such an effect. Third, the analysis revealed general strong differences in the interaction behavior between teams, both with regards to their freely structured and guided decision making process. Literature suggests a broad range of factors influencing teamwork process. In the setting of this study, it is likely that even all participants where in their role, their specific experience varied because of the different airports they came from. Furthermore, the teams consisted of different personalities. Extroverted persons tend to talk more, so even if guidelines foresee no activity for these persons their personality might lead to an active interaction behavior.

C. Advancing the guidelines

The detailed analysis conducted in this paper provided insights into decision making process in complex tasks with conflicting goals, hidden information profiles, influenced by different personalities. The limitations of a simulation study are discussed above. Nevertheless, the results provided evidence how to advance the guidelines.

First, results showed that airlines communicated their constraints, represented by priority flights, early and during the situation assessment phases. This finding matches with experiences made with the FOR-DEC model. There, it was observed that pilots thought the process to be too lengthy in case a suitable solution was “obvious”. The guidelines should foresee a phase where constraints that affect situation assessment are shared.

Second, following the approach of the FOR-DEC design, the structure of decision making process should be easy to remember. More emphasis should be on the necessity of a proper situation assessment. The challenge in the collaborative environment is that no single stakeholder has access to all information relevant for situation assessment. In the set-up used for this study, traffic situations were pre-defined and consequences on operations included in the instructions of the simulation run. Nevertheless, explicit assessments like “how many flights are affected” versus “what capacity is available” are valuable to help the team to understand the “affordance” of the situation.

Third, it is of interest to develop guidelines and to understand collaborative decision making in situations with higher risks and higher time pressure as in the APOC environment. Also, decision making processes with several teams contributing to decisions, thus inter-team-collaboration, should be addressed as future air traffic management concepts foresee those processes. For instance, inter-team collaboration is required in remote tower centers when airport control is distributed dynamically, dependent on situational factors. In flight-centric air traffic control, teamwork and task distribution between teams is regarded as one factor to achieve more efficient air traffic control [22]. This paper proposes an approach how to analyze and evaluate collaborative and team processes. Whilst these analysis methods are labor intensive when conducted manually, evolving methodologies like speech recognition, natural language processing and process mining bear the potential for more automated analysis.

REFERENCES

[1] *Commission Implementing Regulation (EU) No 716/2014 of 27 June 2014 on the establishment of the Pilot Common Project supporting the implementation of the European Air Traffic Management Master Plan Text with EEA relevance,* E. Commission 716/2014, 2014.

[2] K. G. Zografos, M. A. Madas, and Y. Salouras, "A decision support system for total airport operations management and planning," *Journal of Advanced Transportation,* vol. 47, pp. 170-189, 2013.

[3] A. Papenfuss, N. Carstengerdes, and Y. Günther, "Konzept zur Kooperation in Flughafen-Leitständen," presented at the 57. FACHAUSSCHUSSSITZUNG ANTHROPOTECHNIK, Rostock, Germany, 2015.

[4] EUROCONTROL. (2008, 07.08.2009). Airport CDM Implementation: the manual. *3*. Available: http://www.eurocontrol.int/airports/gallery/content/public/pdf/cdm_implementation_manual_2009.pdf

[5] EUROCONTROL, "A-CDM Impact Assessment: Final Report," 2016.

[6] Y. Günther, A. Inard, B. Werther, M. Bonnier, G. Spieß, A. Marsden, M. Temme, D. Böhme, R. Lane, and H. Niederstrasser, "Total Airport Mangagement: Operational Concept & Logical Architecture," EURONCONTROL, DLR, Braunschweig, Betrigny2006.

[7] G. Spies, F. Piekert, A. Marsden, R. Suikat, C. Meier, and P. Erisken, "Operational Concept for an Airport Operations Center to enable Total Airport Management," presented at the 26th International Congress of the Aeronautical Sciences, Anchorage, Alaska, 2008.

[8] H. Bogers, M. Linde, J. R. Matas Sebatia, I. Álvarez Escotto, J. I. Martin Espinosa, B. Ciprián Tejero, ..., and J. A. Ubeda Torres, "D16 - OFA 05.01.01 Consolidated OSED Edition 3, Part 1," Brussels, Belgium2015.

[9] H. Bogers, M. Linde, J. R. Matas Sebatia, I. Álvarez Escotto, J. I. Martin Espinosa, B. Ciprián Tejero, ..., and J. A. Ubeda Torres, "D16 - OFA 05.01.01 Consolidated OSED Edition 3, Part 2," Brussels, Belgium2015.

[10] N. Carstengerdes, A. Papenfuss, R. Suikat, S. Schier, Y. Günther, F. Piekert, and A. Marsden, "EXE-06.03.01-VP-757 Validation Report. D142," Brussels, Belgium2016.

[11] N. Carstengerdes, M. Jipp, F. Piekert, A. Reinholz, and R. Suikat, "Validation of the Total Airport Management Suite- Research Report," Deutsches Zentrum für Luft- und Raumfahrt, Braunschweig2012.

[12] A. E. v. Dongen, "Literature Review On Negotiation Protocols: Total Airport Management," National Aoerspace Laboratory NLR NLR-CR-2008-145, 2008.

[13] A. Papenfuss, N. Carstengerdes, S. Schier, and Y. Günther, "What to say when: Guidelines for Decision Making. An evaluation of a concept for cooperation in an APOC," presented at the 12th USA/Europe Air Traffic Management R&D Seminar, Seattle, U.S., 2017.

[14] J. M. Orasanu, "Flight Crew Decision-Making," in *Crew Resource Management (Second Edition)*, B. Kanki, G. , R. Helmreich, L. , and J. Anca, Eds., ed San Diego: Academic Press, 2010, pp. 147-179.

[15] H. J. Hörmann, "FOR-DEC - A Prescriptive Model for Aeronautical Decision Making," in *21st Conference of the European Association for Aviation psychology (EAAP(*1995, pp. 17 - 23.

[16] G. Hofinger, S. Proske, H. Soll, and G. Steinhardt, "FOR-DEC & Co: Hilfen für strukturiertes Entscheiden im Team," in *Entscheiden in kritischen Situationen: Neue Perspektiven und Erkenntnisse*, R. Heimann, S. Strohschneider, and H. Schaub, Eds., ed: Verlag für Polizeiwissenschaft, 2014, pp. 119 - 136.

[17] S. Schier, F. Timmermann, and T. Pett, "Airport Management in the Box - a Human-in-the-loop Simulation for ACDM and Airport Management," presented at the Deutscher Luft- und Raumfahrt Kongress 2016, Braunschweig, 2016.

[18] S. Schier, T. Pett, O. Mohr, and S. Yeo, "Design and Evaluation of User Interfaces for an Airport Management Simulation," presented at the AIAA Modeling and Simulation Conference, Washington D.C., 2016.

[19] P. Mayring, *Qualitative Inhaltsanalyse: Grundlagen und Techniken* 13 ed. Weinheim: Beltz, 2010.

[20] C. C. Schermuly, T. Schröder, J. Nachtwei, and W. Scholl, "Das Instrument zur Kodierung von Diskussionen (IKD)," *Zeitschrift für Arbeits- und Organisationspsychologie,* vol. 54, pp. 149-170, 2010.

[21] A. Field, *Discovering Statistics Using SPSS*, 4th ed. London: Sage Publications, 2011.

[22] B. Birkmeier, S. Tittel, and B. Korn, "Controller Team Possibilities for Sectorless Air Traffic Management," in *Integrated Communications Navigation and Surveillance (ICSN) Confernece*, Herndon, VA, USA, 2016.