Know your options – analysing human decision making in dynamic task environments with state space methods

Hendrik Oberheid[*], Andreas Hasselberg[*], Dirk Söffker[**]

* German Aerospace Center (DLR), Institute of Flight Guidance, Braunschweig, Germany
** Chair of Dynamics and Control (SRS), University of Duisburg-Essen, Germany

Corresponding Author:
Hendrik Oberheid
German Aerospace Center
Institute of Flight Guidance
Lilienthalplatz 7
38108 Braunschweig
hendrik.oberheid@dlr.de

Abstract

The paper discusses new methods for the analysis of human decision making in dynamic task environments (DTEs). It is argued that to understand human operator decisions, choices, and errors, researchers should seek to know a) the available options for each situation and their consequences and b) the options and consequences perceived by the operator in the DTE. On the example of an air traffic control scenario it is shown that both the technical part (a) and the human factors part (b) of this assumption are insufficiently supported by available methods. Taking a two-fold approach to this problem, the paper discusses state space methods to analyse spaces of physically possible solutions and describes an affordance test for operators to measure perceived affordances in dynamically changing situations. The overall approach is demonstrated on an en-route air traffic control study conducted in a microworld setting with N = 16 students, who executed a conflict detection and resolution task. Participants' subjective perception of affordances for conflict resolutions is contrasted with objective state space data. The results reveal differences in the difficulty and success of affordance assessment for different strategy types (speed, altitude, and lateral strategies). These differences are only partially reflected in participants' subjective difficulty assessment.

The proposed methods should lead to a more detailed understanding of human decisions and errors on dynamic tasks and help to identify areas, where assistance might be most useful.

Introduction

In many safety critical applications such as air traffic management, other transport applications, or process control, the decisions of human operators play a central role. Before changes to human-machine systems and operational procedures are introduced in the above mentioned domains, extensive human-in-the-loop simulations are usually performed. However, a recurring problem when trying to analyse the decision making of human operators in interactive simulations is the following: decisions are often evaluated with an incomplete understanding of available choices. While the decisions and actions that are actually taken by the operator are usually well recorded in meticulous simulation logs, information on the available alternative choices and potential alternative options for solving a specific problem situation are nearly always missing. Thus, information is usually only available on what actually happened. Alternative situations that could have been reached, in a positive as well as in a negative sense (including the proximity to dangerous states) often remain unclear, and technical analysis methods for exploring them are lacking.

Regarding the understanding of the human factor in the decision making process, the method support is not always more satisfying. Undoubtedly, it has become standard practice in many domains to test and query operators on situation awareness (SA) issues (Endsley, 1995), and good situation awareness can be argued to contribute to good decision making. However, SA is not concerned with successful problem solving itself (Pritchett & Hansmann, 2000). Measurement tools and metrics to find out directly about solution spaces which operators perceive in dynamic task environments seem less common and are not often employed. That way, it is difficult to know when human operators perceive the factual action opportunities of the dynamic system correctly or incorrectly and where the biggest potential may reside to assist them with supporting automation.

Given the ignorance of available solution spaces, current evaluation techniques often perform a statistical analysis of the decisions only. For an en-route air traffic control simulation, a typical result might report that a controller employed seven direction/heading clearances and three speed or altitude clearances ‘apparently’ to solve conflicts between aircraft. However, how many of these conflicts he could actually have solved more efficiently (or at all) by one strategy or the other usually remains unknown, as well as the fact which other options were perceived but dismissed in favour of the one actually implemented. It is pointed out, that due to the interactivity of typical air traffic control scenarios and interindividual differences in controller behaviour, even shortly after the scenario start, the factual action opportunities between controllers may be entirely different and cannot be predicted in advance of the exercise.

In an effort to remove some of these restrictions, this work describes methods allowing the evaluation of human decision making in relation to the specific solution spaces and action opportunities available in complex interactive simulations. From a human factors research point of view, reference is made to the concept of affordances (Gibson, 1979). These are used to describe if an action opportunity is present or not, and if it is perceived as such by the operator or remains hidden. The affordance concept is adapted to a specific air traffic control problem and a questionnaire method is used to test perceived affordances during a simulation. Regarding the technical approach, the reported method is based on the analysis of discrete state spaces computed from Coloured Petri Nets (CPN) (Jensen et al., 2007). On the basis of the state space, the feasibility of different action strategies (e.g. to solve traffic conflicts between aircraft vertically, laterally or by speed instructions) is explored. In (Oberheid, 2006) state space methods are used for the analysis of air traffic control and arrival management scenarios and supporting automation. Möhlenbrink and colleagues point out, that it is the strength of state space analysis to cope with the time-depended evolution of traffic scenarios caused by different controller behaviour (Möhlenbrink et al. 2008). The work of (Werther, 2006a; Werther, 2006b; Möhlenbrink & Schnieder, 2008) uses state space methods for the analysis of human operator decisions in DTEs and puts a focus on holistic human machine modelling and cognitive modelling.

The key new contribution of this paper presented here is that the physical view on affordance existence is contrasted with the subject's response on affordance perception, in order to make subjects' difficulties with assessing the feasibility of different strategies in dynamic situations objective and measurable.

Affordances for an air traffic control task

In this paper the term affordance is used to reason about factual and perceived action opportunities of air traffic controllers during conflict resolution task. In common language the verb 'afford' is used to describe if a situation allows for some action and if resources exist to make something possible. The noun 'affordance' as a psychological construct was originally coined by Gibson to describe all possible actions in an environment relative to an animal. This section outlines some of the discussion related to the affordance concept before developing an applied working definition for the ATC problem studied in this paper.

Originally, Gibson states that 'the affordances of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either good or ill' (Gibson 1979). As summarized by McGrenere and Ho (McGrenere & Ho, 2000) Gibsons view includes three fundamental properties of affordances which are (a) an affordance exists relative to the action capabilities of a particular actor (b) the existence of an affordance is independent of the actor's ability to perceive it and (c) an affordance does not change as the needs and goals of the actor change.

Turvey and Shaw (1979) brought the expression of effectivities into discussion. Effectivities are properties of animals that are complement to affordances and allow the animals to use affordances. They state that affordances and effectivities are inseparable, thus they can not exist without each other. If an affordance and an effectivity are present, a behaviour can be actualized. In an ATC example, the terms describe what is possible in the simulation (affordance) and what can be done/handled by the controller (effectivities), but their existence depends on another.

Chemero (2003) breaks with the definition of affordances as (animal related) properties of the environment. Here, affordances are relations between abilities of animals and features of the environment and both of them can exist separately. The author understands abilities of animals as functional properties, with the possibility of malfunction. For example one may fail to climb a step that affords climbing, even under the best conditions. Using this definition allows it to analyse the properties (what is possible in the simulation) and the capabilities (what the operator is capable of doing) separately.

Gaver (1991) separated affordances from the information available about them and defined four categories (perceptible affordance, false affordance, hidden affordance, correct rejection). A similar two-dimensional categorization will be employed in this paper. But instead of contrasting affordance existence with information availability about them, the perception of affordances is used as the second differentiating factor, leading to the matrix in Table 1.

	nonexistent affordance	existent affordance
perceived affordance	misperceived affordance	true affordance
unperceived affordance	correct rejection	hidden affordance

Table 1: Contrasting affordance existence and affordance perception

Working definition of affordances for the studied ATC problem

When it is possible in the simulation environment to avoid losing separation (feature of the environment) and the air traffic controller is capable of implementing that solution (capability) that is an *existent affordance*. In contrast to that, when the air traffic controller opines that a certain solution is possible, this will be denoted as *perceived affordance* (without considering the factual existence of an affordance). The combinations of these categories will be denoted as follows: An affordance which is perceived and which exists is a *true affordance*, an affordance which is not perceived but exists is a *hidden affordance* and an affordance which is perceived but does not exist is a *misperceived affordance*. If an affordance is neither perceived nor exists this is termed a correct rejection.

Microworld simulation environment MAGIE

The purpose of the microworld environment MAGIE (Micro Air Ground Integration Environment) is to test new procedures and assistance systems for air traffic control within a simplified and highly controlled setting (Oberheid et al., 2010; Weber et al., 2010). To allow the comparison of results from state space analyses and operators decision making in human-in-the-loop simulations MAGIE is based on a Coloured-Petri-Net-model realized with CPN Tools. This specific approach is suggested by Möhlenbrink et al. for the analysis of human decision making in dynamic task environments (Möhlenbrink et al. 2008; Möhlenbrink & Schnieder, 2010). The second part of the microworld is a graphical user interface (GUI) (see Figure 1), which is a representation of a radar display with some additional information elements, such as flightstrips and compass rose. Examples of Coloured Petri Nets coupled with graphical interfaces were discussed previously also in (Welters, 2005; Gamrad, 2006; Oberheid, 2006; Möhlenbrink et al., 2008).

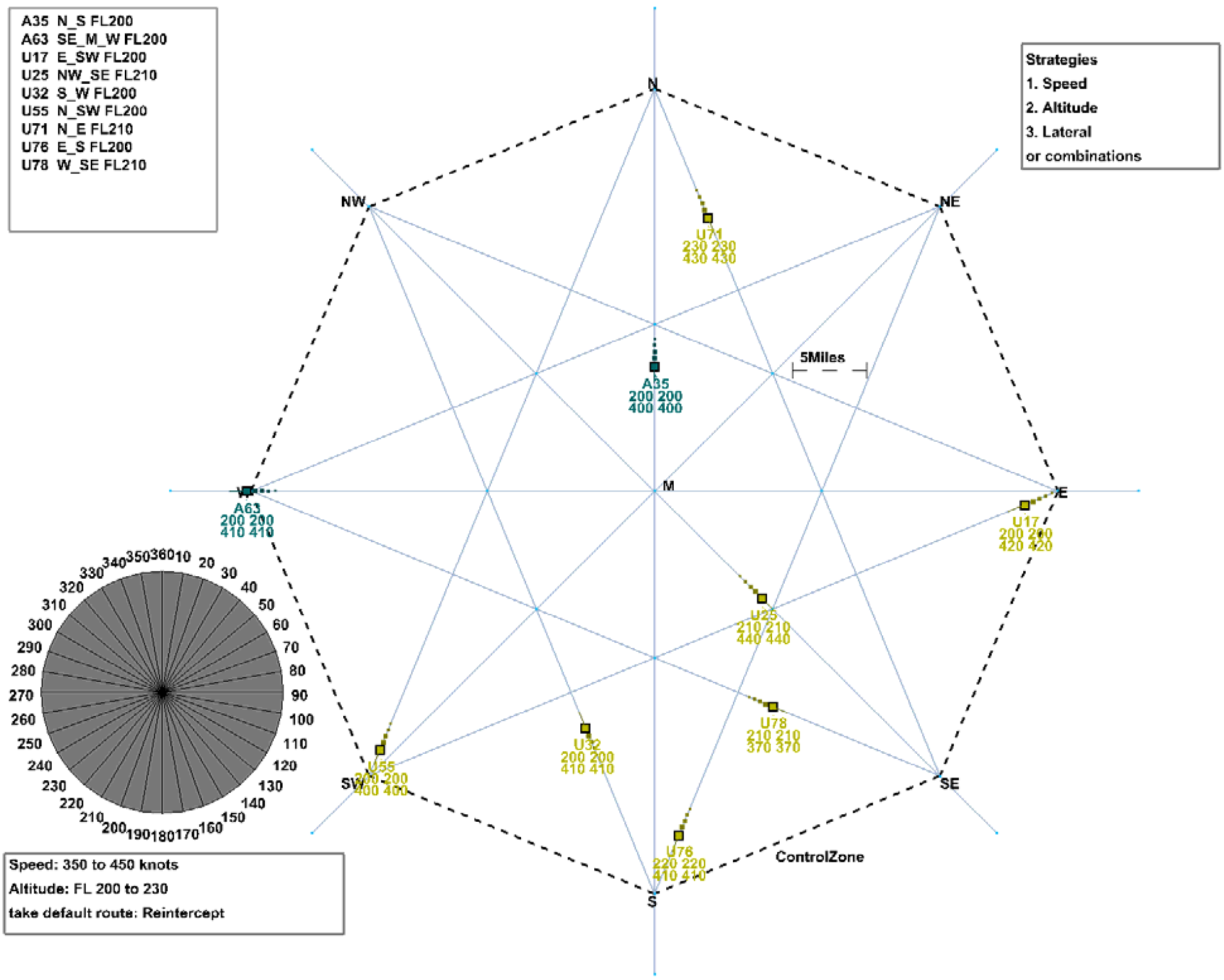

Figure 1: Microworld simulation environment GUI

Operator task

The simulation scenario studied in this paper features a generic en-route air traffic control sector. The sector has eight different entries/exit points. Each of these entries is connected to five exits. Different types of routes exist to build such connections, including straight and direct routes and bended routes via intermediate waypoints. This causes a high number of potential crossing points within the sector. For each aircraft, the waypoint of the route and the designated exit altitude can be read off the flightstrips represented in the upper left corner of the GUI. The current aircraft state (present and cleared speed, present and cleared altitude) is represented in the aircraft labels.

The main task of the operator/controller in the simulation scenario is the monitoring of aircraft for traffic conflicts and the control of avoidance manoeuvres when potential conflicts arise. In line with ICAO (International Civil Aviation Organization) radar separation rules, a separation violation/conflict is defined as a state where separation between two aircraft drops below 5 NM (9261 m) lateral separation and 1000 ft (304 m) vertical separation.

Conflict solution strategies

The operator can avoid conflicts by giving aircraft appropriate control clearances. Individual clearances can take the form of

- speed changes (decelerations or accelerations),
- altitude changes (descends or climbs) or
- lateral deviations from the route (left or right turns).

Conflict solution strategies can consist of combinations of these types of clearances for one or several aircraft. After altitude and lateral deviations, aircraft have to be guided back to their original route/altitude. This requires further monitoring and control commands of the controller.

To implement successful solution strategies, the physical aircraft dynamics (rate of acceleration, deceleration, climb and turn, minimum and maximum speed) have to be taken into account. Also the available altitude range of the sector has to be considered. Apart from the task of conflict detection and avoidance, aircraft have to be guided to their respective target altitude.

Conventional performance measures

A number of conventional performance measures can be gathered for the MAGIE simulation, which are based on a simple evaluation of event logs. Similar data are recorded by most simulators today. Examples of this are counting

- the number of separation violations (conflicts),
- the number of processed aircraft (throughput),
- the number of control commands given (controller activity), and
- the average flight path length (efficiency).

Need for new measures to analyse solution spaces

The simple performance measures listed above are useful to evaluate the overall effectiveness of a subject in dealing with a scenario. However, they offer little to answer the following questions:

- Which control commands exactly could have been used to avoid a conflict (what were the affordances of a situation)?
- When did solutions by certain strategies become impossible and until when could a safe solution be reached (when did affordances of the situation change)?
- What were the causes of erroneous decisions and individual separation violations (was a potential conflict detected in time but not solved successfully, potentially due to misperceived or hidden affordances)?

Answering such questions for dynamic situations is not technically supported in conventional simulation mode. In order to solve such problems, new types of analysis functions were implemented for the MAGIE Environment using state space methods.

State space methods

The Coloured Petri Net used in the simulation environment MAGIE was implemented with 'CPN Tools' as a graphical modelling tool (Jensen et al., 2007). The simulation mode of this software is used during trials with participants in order to execute sequential steps to determine the next state of the simulation. For analysis purposes the state space mode of this software is used to calculate possible model states and their connections to each other (Jensen, 2006).

Depending on the available computational resources CPN Tools can generate complete state spaces (representing all reachable states) or only partial state spaces (representing only a subset of reachable states). Due to the complexity of the MAGIE simulation the possible operator interventions are manifold which results in a high degree of freedom. As a consequence the size of the state space increases rapidly with expanding time horizon. This makes it impossible to calculate a complete state space beforehand. To avoid this problem, previous work developed techniques to calculate multiple partial state spaces separately. Thus, the handling of a complex scenario becomes possible (Oberheid et al., 2008; Gamrad et al., 2009).

The state space method requires four steps to analyse the effects of a specified strategy in a certain situation (see Fig. 2). First a logged state from a trial (at the time of an affordance test) is used as the initial state of a partial state space (Step 1).

To generate a state space for the MAGIE simulation, certain possible operator actions and degrees of freedom of operator choices have to be specified. For this purpose a simple controller strategy model is implemented. This model contains a list of all possible actions/clearances (change of speed, altitude and direction). Furthermore this model is equipped with a set of rules to define plausible actions. Some of the rules have to be activated for every state space calculation (Step 2). For example, if the state space method is used to analyse the effect of the strategy 'speed-down' on the aircraft 'U01' the activated rules narrow down the considered clearances to speed reductions concerning the aircraft 'U01'.

To allow a set of different controller actions and leave the choice to the model causes an essentially nondeterministic behaviour of the net. This property enables the software to calculate a partial state space, which describes the reachable states near the initial state (Step 3).

The following analysis is developed to identify conflict free paths from the initial node to nodes at least 600 seconds later. If such a path exists, the strategy is considered a feasible solution (Hasselberg et al., 2009) and interpreted as an existing affordance (Step 4).

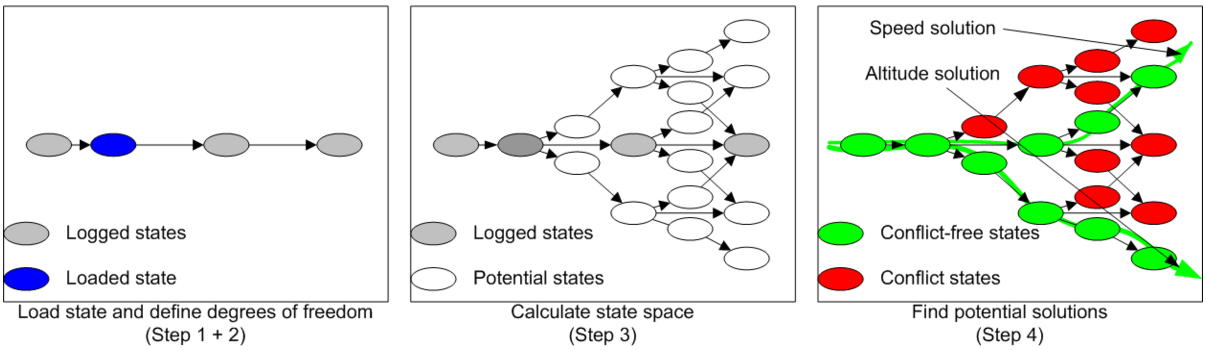

Figure 2: Steps to find potential solution using state spaces

Every state at the time of an affordance test is analysed with the state-space-method. The state space calculation is repeated depending on the amount of controllable aircraft and potential strategies.

Research question

The main purpose of the reported study is to demonstrate the usage of state space methods and affordance testing to examine influencing factors on operators' difficulty with correct affordance perception. The focus of the discussion is hereby on the methodological aspects rather than on a comprehensive valuation of the empirical results on the level of the air traffic control application.

Based on the introduced applied definition of affordances for the discussed ATC problem and the presented state space methods, the analysis categorizes human perception of potential solutions into true affordances, hidden affordances, misperceived affordance and correct rejects (see Table 1). The data within these four affordance categories is used to reason about the difficulty of assessing the feasibility of certain strategy types (affordance perception difficulty) for the application. It is argued that

- a strategy type is *easy* to assess if subjects' affordance ratings show a high number of true affordances and correct rejections.
- a strategy type is *difficult* to assess if subjects affordance ratings show a high number of misperceived and hidden affordances.

For the following study it is hypothesized that affordance perception for solution strategies via speed, altitude or lateral deviations will have different difficulties. This assumption is related to the different physical characteristics of aircraft reactions in each dimension (differences in vertical vs. horizontal change rate vs. turn rates). It is also connected to the different observability of the situation and changes in each dimension on the radar display (2D projection of 3D space, aircraft altitude encoded in labels).

Also it is assumed that the success of affordance assessment should depend on the quality and amount of available information during the assessment. Insufficient recall of the traffic situation (inadequate situation awareness) should hinder the subjects' ability to judge affordances for the specific situation correctly despite the general cognitive capability to so. Manipulation of information availability during the affordance test should therefore affect the percentage of affordances perceived correctly.

Further, the objective measures on difficulty of strategy assessment are contrasted with subjective difficulty measures in order to explore how far they correlate or contradict each other.

Method

Participants

Sixteen (3 female, 13 male) undergraduate students from the Technical University of Braunschweig participated in the study. Participants' age ranged between 21 and 28 years (M = 24.0, SD = 2.25). As an incentive for participation, each participant was paid 25 €. Additionally, subjects could earn a bonus of 30 € if they were among the best four performers of the study in terms of performance in conflict detection and solution.

Operator task

Participants controlled en-route air traffic in the MAGIE microworld simulation environment (see section Microworld simulation environment). The previously described route structure and GUI were used. Participants' task was to monitor the air traffic for potential conflicts and to avoid conflicts by appropriate control clearances (speed, altitude or lateral deviations with headings). Also aircraft had to be cleared to their respective sector exit altitude/flight level. Participants issued control clearances verbally in informal air traffic control phraseology.

Clearances were read back and entered into the simulation by the experimental instructor at a separate working station.

Measures – affordance test

Trigger process: Online affordance tests for conflict solution were administered a variable number of times during each trial, according to a specifically designed trigger procedure. Aim of the trigger procedure was the positioning of affordance tests in the middle of the participants' problem solution process. Therefore a test should be triggered after a potential conflict had been detected by the participant (so the participant should already be aware of the problem) but before a solution for the conflict had been implemented. In order to achieve this goal, participants were instructed to mark each detected potential conflict immediately by pressing the Ctrl button on the keyboard and clicking on the aircraft labels with the mouse. Only after the conflict was marked, the participants should begin with the solution process and issue control clearances for conflict avoidance.

Whenever a conflict was marked by the subject, a random process was initiated with a 60% probability of actually triggering an affordance test within a time interval of 2 to 20 seconds and a 40% probability of triggering no affordance test at all. By the applied stochastic dependencies the connection between marking conflicts was made less transparent. In that way, participants should be detained from waiting for the affordance test before starting their conflict resolution, as they would experience that sometimes no affordance test would appear at all. Also, whenever participants took action (i.e. issued clearances) before a test was triggered by the timer, the test would be cancelled. The trigger process is illustrated graphically in Figure 3.

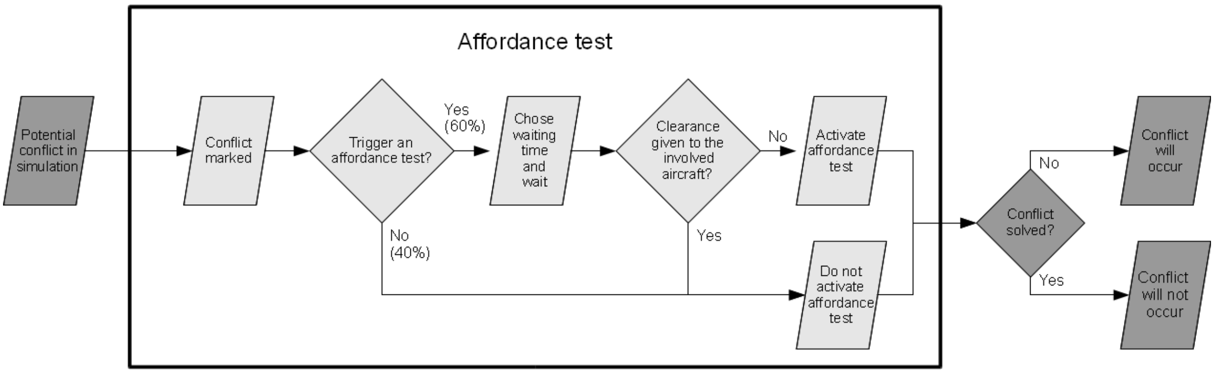

Figure 3: Trigger procedure for affordance testing

Questionnaire content: If an affordance test was activated, the simulation was frozen and an online questionnaire appeared on a separate screen. The questionnaire contained four groups of questions.

Subjective conflict probability: The first group of questions asked subjects to rank the probability of a separation violation if no control clearances for conflict avoidance were given. Responses were given on a five point scale with levels zero, low, medium, high and certain. When the answer to the initial question of conflict probability without control actions was ‘zero’ the rest of the questionnaire did not have to be filled out and was excluded from the data set.

Simple solution strategies: The second and third group of questions referred to the core issue of perceived affordances of the situation. Six different strategies (speed reduction/increment, altitude reduction/increment, left/right turn) were offered for each conflicting aircraft A (second group) and B (third group). Participants were asked to mark strategies which they perceived as *feasible* solutions for avoiding the conflict. They were instructed to mark *all* solutions they thought to be *feasible*, regardless of their preference for any particular strategy.

Combined strategies: The fourth group of questions combined strategies for aircraft A and B at the same time, with clearances for both aircraft. However, feasibility of combined strategies had to be rated only if the simple strategies for both A and B were thought unfeasible.

Questionnaire phases: The overall affordance questionnaire consisted of two phases represented on two consecutive pages. The pages showed the same set of questions, but the phases differed with regard to the amount of situational information which was available to the subjects on the primary radar display while the questionnaire was filled out on the secondary screen. During the first phase, only positional cues of aircraft were visible on the radar. Aircraft speed and altitude indicators in the aircraft labels were hidden. This could make it difficult for participants to give adequate responses if they could not recall the respective information from memory. After the first page of the questionnaire had been saved, however, speed and altitude information reappeared on the primary radar screen. Participants could then correct the first phase response (which was still visible) if the new information had changed their judgement of affordances. The motivation of the two-phased approach was to find out if the affordance judgement of subjects would depend heavily on a potentially imperfect mental traffic picture and situation awareness state (it should then be altered frequently when the full traffic picture became available again on the radar) or if problems were more related to the difficulty of judging aircraft dynamics themselves.

Procedure

Participants first read a set of instructions, which described the interface and experimental task. After that, a 20 minutes practice trial was started to familiarize participants with the task and environment. The practice trial contained a structured series of examples for different types of conflicts and conflict resolution strategies. Control clearances for speed changes, altitude changes and lateral avoidance manoeuvres by headings were practiced. Instructions contained a preference ranking for available strategy types, which was that speed changes were preferred to altitude changes and these were preferred to lateral changes. Participants were instructed to follow this preference ranking when multiple solutions were possible. All GUI functions were explained verbally in a standardized manner in addition to the written instructions. Affordance tests were practiced at least two times during the practice run.

After the practice trial, the two similar experimental scenarios were absolved (the order of the scenarios was controlled for) building the data set for the following statistical analyses and discussion. Each experimental session lasted 30 minutes plus a variable time for freezes from the affordance tests. For the experimental trials, participants were instructed that their performance evaluation for the bonus payment would depend both on the earliness of conflict detection and on the success of conflict avoidance.

In a post-experimental questionnaire, participants were asked about the difficulties with evaluating the possibilities of the three strategies. A scale from 0 (not difficult at all) to 5 (very difficult) was presented. Also, participants were asked to rank the strategies in their favourite order (1st rank = most favourite strategy, 3rd rank = least favourite strategy).

Results and discussion

Data pre-analysis, aggregation and exclusion A total number of 133 two-phased affordance tests were triggered and filled out by the 16 participants of our study (min = 2, max = 17, M = 8.31, SD = 3.9 tests per participant). In 38 cases (both phases in 17 affordance tests, one phase in 4 tests) participants rated the conflict probability without intervention as 0 (“the conflict will not occur”), rendering further assessment of conflict avoidance manoeuvres obsolete. The respective data (for both phases) were excluded from the analysis.

The data of 112 affordance tests and the total amount of 2376 questions about perceived affordances remained. This subjective data was compared with the objective data calculated with the state-space-analysis and responses classified into correct rejections and true affordances (correct perception) and hidden and misperceived affordances (false perception).

Affordance perception correctness as a function of information availability The influence of availability of speed and altitude information (questionnaire phase1 = hidden, phase2 = visible) on the correctness of affordance perception was analysed. The analysis of the data collected with the affordance test revealed that the perceived affordances changed in only 4.2 % of cases between the two phases. An ANOVA showed no main effect of phase on the correctness of affordance perception ($F(1, 2349) < 1$).

The data suggests that in the employed setting, the availability of speed and altitude information on the radar display during the test did not have the expected positive influence on the correctness of response. A possible explanation of this result is that the participants' situation awareness was sufficiently good, so that the hidden information was not needed when the test appeared. This may either reflect participants' preparation for the test or above average focusing and situation awareness for aircraft with conflict potential (Hauss & Eyferth, 2003). Another explanation may be that the participants refused to correct their own judgement. In any case, the results give no indication that cases of hidden or misperceived affordances were primarily due to insufficient level 1 situation awareness (Endsley, 1995) with regard to speed or altitude information. Rather, they point towards other cognitive difficulties or an imperfect mental model of the process (e.g. the aircraft dynamics) as a potential causal factor for false responses.

As no main effects and no secondary effects with phase were found, the data from the first questionnaire phase was excluded to avoid duplication of data. The following analysis is based on the results from the second phase only, with speed and altitude information visible.

Affordance perception as a function of strategy type The affordances for both directions of each conflict resolution strategy were grouped together as they were considered as equally perceivable. This yielded three distinct strategy types which are speed (either up or down), altitude (either up or down) and lateral (either left or right). In Figure 4 the rate of positive test responses (affordance perceived) differentiated by the three strategy types (speed, altitude, lateral) and the factual existence of the affordance (E = existent affordance, N = nonexistent affordance) according to state space calculations is shown. An ANOVA revealed a significant main effect of strategy type ($F(2, 1100) = 12.27$, $p < .001$) and affordance existence ($F(1, 1100) = 29.82$, $p < .001$) on affordance perception. Most importantly with regard to the formulated research question, however, a significant interaction effect of strategy combined with affordance presence was found ($F(2, 1100) = 7.33$, $p < .001$).

This interaction effect indicates the differences in the difficulty of correct affordance perception over strategies. In addition to the ANOVA, the effect size of affordance presence was calculated for every strategy using Cohens'd (Cohen, 1988). The calculated effect sizes together with their 95% intervals are also shown in Figure 4. According to Cohen, the effect size of affordance existence for the strategy type speed is classified as small ($d = .21$) while it is moderate for the strategy types altitude ($d = .66$) and lateral ($d = .66$). The bigger the effect, the easier it was to assess affordance existence for the strategy.

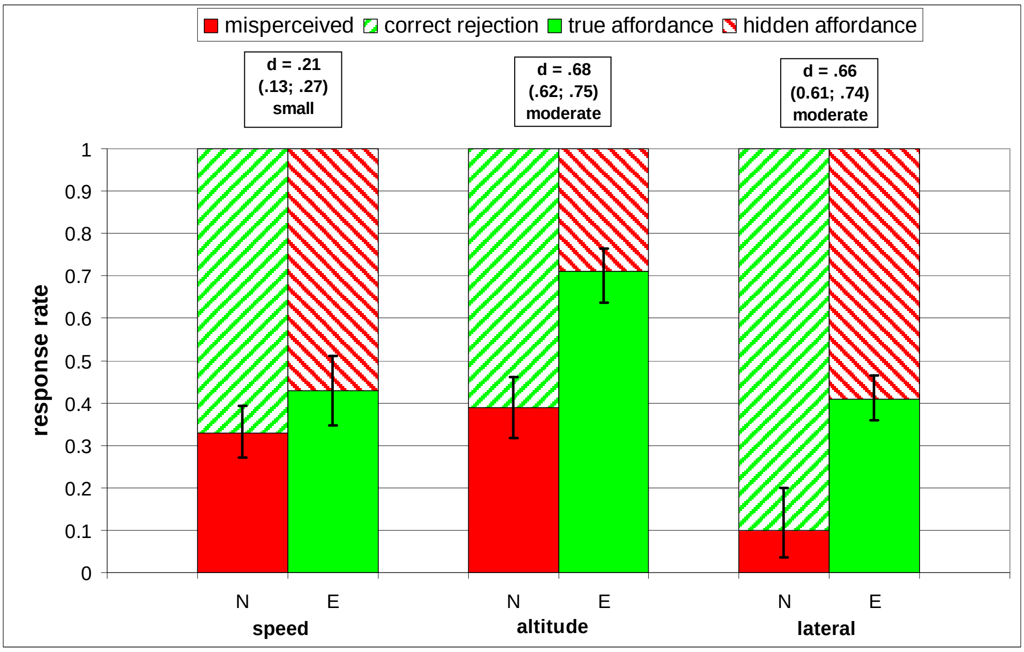

Figure 4: Affordance perception as a function of strategy and affordance existence

The significant secondary effect of strategy combined with affordance existence shows that the difficulty to perceive affordances varies between the strategies. The calculated effect sizes reveal that discrimination between situations where an affordance does or does not exist is much more successful for the vertical and lateral avoidance strategies than for the speed strategy. Speed affordances were therefore more difficult to assess then vertical and lateral solutions. In fact, regarding the test responses for speed strategy only, no significant effect of affordance existence is observed, leaving open if participants in the setting could make any sensible assessment of the speed affordances at all. With view to the design of potential assistance systems, participants should be more likely to profit from a decision support system for speed assistance than from assistance to lateral or altitude manoeuvres.

Data from post-experimental questionnaire Participants subjectively rated the affordance perception of the speed-change strategy as the most difficult one (M = 3.38, SD = 1.310), the lateral avoidance strategy as the second most difficult (*M* = 2.81, *SD* = 1.759) and the altitude-change strategy as the easiest one (*M* = 2.13, *SD* = 1.408), although potentially due to the low number of participants, the repeated measures ANOVA revealed no significant effect ($F(2, 30) = 2.52$, $p = .098$). The results regarding subjective strategy preference are as follows: The altitude-change strategy reached the highest rank (*M* = 1.44, *SD* = 0.727). The speed change strategy was rated on the second rank (*M* = 1.87, *SD* = 0.500) and the direction-change strategy is the least favourite strategy (*M* = 2.69, *SD* = 0.704). A χ^2 Test showed that the strategies were ranked significant differently ($p < .001$).

Relating objective difficulty to subjective difficulty and preference Contrasting the objective difficulty ranking (from comparing state space with affordance tests) with subjective difficulty and strategy preference rankings (from post-experimental questionnaire) revealed the following: For the altitude strategy a low objective and low subjective difficulty were found. As it would be expected, the subjective popularity of altitude strategies was thus high. For the speed strategy a high subjective and objective difficulty was found. However the popularity was still moderate (specifically, speed strategies were more popular than for the lateral strategies). For the lateral strategy objective difficulty ranked comparatively lower than subjective difficulty. Despite low objective difficulty, the popularity of lateral manoeuvres was lowest of all strategies.

For the specific application, a potential explanation of this mismatch lies in the usually higher number of follow-up clearances and potentially higher monitoring load during lateral solutions. On a more abstract level, however, the results also show that subjective and objective difficulties are not necessarily correlated. They are therefore no good predictors of each other and each of the measures play its own role in the overall system evaluation.

Conclusions

The paper points out methodological restrictions regarding the analysis of human decision making in dynamic task environments (DTE). These restrictions are related on the one hand to an often insufficient understanding of the available choices and their consequences and costs in the DTE. On the other hand, a lack of established tools and metrics to measure perception of solution spaces by the operators is diagnosed. Both limitations hinder the understanding of human decisions and errors on dynamic tasks and impede insights how the operator can be assisted most efficiently by new automation and decision support systems.

To approach this problem, this contribution argues from a technical and formal side for the application and intensive research on state space methods, in order to get a handle on the factual action opportunities and costs of complex dynamic tasks. On the human factors side, the paper proposes an affordance test procedure to query operators with regard to perceived solutions. A microworld en-route air traffic control study is conducted, which demonstrates how state space and affordance test data can be combined and contrasted to determine objective difficulty measures for assessment of potential solutions by the operator. It is shown that objective difficulties are not necessarily well reflected in participants' subjective difficulty ratings. So both results should be of independent value for the overall analysis.

Considering the results, a vast number of aspects regarding the application of state space methods for the analysis of human decisions in DTEs remain to be discussed. Among other things, the dynamic development of the solution space over time and its perception and also the nature of potential mappings from the structure of the physical state space to human operators' (problem-dependent) mental representation. However, the state space based approach should be able to support more effective analysis of operator decision in various domains. After all, it will be easier to judge the decisions when the options are known.

References

Chemero, A. (2003). An Outline of a Theory of Affordances. *Ecological psychology*, *15*, 181-195.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2. Ed., Hillsdale: Lawrence Erlbaum Associates.

Endsley, M.R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, *37*, 32-64.

Gamrad, D. (2006). *Entwicklung von Mustern höherer Petri Netze zur rechnergestützten Simulation und Analyse von Situations-Operator-Modellen*, Master's thesis, University Duisburg-Essen, Germany: Chair of Dynamics and Control.

Gamrad, D., Oberheid, H., & Söffker, D. (2009). Automated detection of human errors based on multiple partial state spaces. *6th Vienna Conference on Mathematical Modeling on Dynamical Systems MATHMOD Vienna*, Austria, 651-659.

Gaver, W.W. (1991). Technology affordances. *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*, 79-84.

Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

McGrenere, J., & Ho, W. (2000). Affordances: Clarifying and Evolving a Concept. *Proceedings of Graphics Interface 2000*, Montreal, Canada, 179-186.

Hasselberg, A., Oberheid, H., & Söffker, D. (2009). State-Space-Based Analysis of Human Decision Making in Air Traffic Control. *7th Workshop on Advanced Control and Diagnosis*, Zielona Gora, Poland.

Hauss, Y., & Eyferth, K. (2003). Securing future ATM-concepts' safety by measuring situation awareness in ATC: Gewährleistung der Sicherheit von zukünftigen Luftverkehrsmanagementkonzepten durch Messung von Situation Awareness in der Flugsicherung. *Aerospace Science and Technology*, *7*, 417–427.

Jensen, K., Christensen, S., & Kristensen, L.M. (2006). *CPN Tools - State Space Manual*, Technical report, University of Aarhus, Denmark: Dept. of Computer Science.

Jensen, K., Kristensen, L.M., & Wells, L. (2007). Coloured petri nets and cpn tools for modelling and validation of concurrent systems. *International Journal on Software Tools for Technology Transfer*, *9*, 213-254.

Möhlenbrink, C., & Schnieder, E. (2010). Fast and Frugal Heuristics conquer the airport. *AHFE International 2010*, 17.-20. July 2010, CRC Press / Taylor & Francis, Ltd. Miami, USA.

Möhlenbrink, C., Oberheid, H. & Werther, B. (2008). A Model Based Approach to Cognitive Work Analysis and Work Process Design in Air Traffic Control. In D. de Waard, F.O. Flemisch, B. Lorenz, H. Oberheid, and K.A. Brookhuis (Eds.), *Human Factors for assistance and automation* (pp. 401 - 414). Maastricht, the Netherlands: Shaker Publishing.

Oberheid, H. (2006). A Coloured Petri Net Model of Cooperative Arrival Planning in Air Traffic Control. *CPN'06 - Seventh Workshop and Tutorial on Practical Use of Coloured Petri Nets and CPN Tools*, Aarhus, Denmark,

Oberheid, H., Lorenz, B., & Werther, B. (2007). Supporting Human-Centered Function Allocation Processes Through Animated Formal Models. In D. de Waard, G.R.J. Hockey, P. Nickel, and K.A. Brookhuis (Eds.) , *Human Factors Issues in Complex System Performance* (pp. 361 - 371). Maastricht, the Netherlands: Shaker Publishing.

Oberheid, H. , Gamrad, D., & Söffker, D. (2008). Closed Loop State Space Analysis and Simulation for Cognitive Systems. *8th International Conference on Application of Concurrency to System Design (ACSD)*, Xi'An, China. 39-44.

Oberheid, H. , Weber, B., & Möhlenbrink, C. (2010). Eye movement analysis to evaluate ghosting *and* targeting aids for controller assistance. In D. de Waard, A. Axelsson, M.

Berglund, B. Peters & C. Weikert (Eds.), *Human Factors: A system view of human, technology and organisation (pp.* 291-311). Maastricht, the Netherlands: Shaker Publishing.

Pritchett, A.R., & Hansman, R.J. (2000). Performance-Based Measures of Flight Crew Situation Awareness Using Testable Responses, In J. Garland & M. Endsley (Eds.) (2000), *Experimental Analysis and Measurement of Situation Awareness* (pp. 189 - 209).

Turvey, M.T. & Shaw, R. (1979). The Primacy of Perceiving: An Ecological Reformulation of Perception for Understanding Memory. In L-G. Nilsson (Eds), *Perspectives on Memory Research: Essays in Honor of Uppsala University's 500th Anniversary (*pp. 167 – 222). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Weber, B., Deml, B., Friedrich, M., Schätzle, S., Oberheid, H., & Preusche, C. (2010). Vibrotaktiles Feedback zur Aufmerksamkeitslenkung bei komplexen Lotsentätigkeiten [Vibrotactile Feedback for Attention Direction in a Complex Air Traffic Control Task]. In M. Grandt & A. Bauch (Eds.), *Innovative Interaktionstechnologien für Mensch-Maschine Schnittstellen* (pp. 71-85). Bonn: DGLR.

Welters, M. (2005). *Building Logistic Games in CPN Tools*, Technical Report, University of Eindhoven, the Netherlands.

Werther, B. (2006a). *Kognitive Modellierung mit farbigen Petrinetzen zur Analyse menschlichen Verhaltens*, PhD thesis, DLR-Insitute of Flight Guidance, Braunschweig, Germany.

Werther, B. (2006b). *Colored Petri Net based modeling of airport control processes.* In: International Conference on Computational Intelligence for Modelling Control and Automation CIMCA06, IEEE Computer Society, Sydney (Australia),108-114.