The Dynamic Team Training Experiment: Improving Tactical Team Decision Making

Edzard Boland, Jelke van der Pal, Christopher Roos
National Aerospace Laboratory NLR
Amsterdam, The Netherlands
boland@nlr.nl, pal@nlr.nl, roos@nlr.nl

ABSTRACT

In order to enhance the decision making process, all air operation team players need to be alert and share information or concerns just in time. Earlier studies apply the mnemonic DESIDE (Detect, Estimate, Set, Identify, Do, Evaluate) to teach pilots to optimize their individual flight safety decision making process. In this paper, DESIDE will be applied by a team of three to test the effectiveness of the tactical decision making tool for teams. For that purpose, eight teams of Falcon 4 gamers were asked to perform an identical set of tactical missions in simulators. Each team consisted of two fighter jet pilots and one supporting Unmanned Aerial System (UAS) operator. In the control condition, a team received a short classical Crew Resource Management (CRM) training, similar in content to the CRM training received in the Royal Netherlands Air Force (RNLAF). In the experimental condition a team received a short training in the use of the DESIDE decision making tool (Murray, 1997). Over the course of five tactical missions, the subjects learned to apply DESIDE. In each debrief, the team was given feedback on the decision making process (according to the instruction received) and mission outcome by a former F-16 Weapons Instructor. The quality of the decision process and outcome was monitored and compared between the teams. The NOTECHS (NOn TEChnical Skills) behavioral marker system in combination with the RNLAF rating scheme was used to evaluate decision making. The results indicate that the DESIDE tool for decision making is effective for teams in a military context. It does generate improved decision making process quality in teams of F-16s and a UAS, performing air operation missions.

ABOUT THE AUTHORS

Edzard Boland (1972) has earned a MSc in Aviation Human Factors at Cranfield University in 2003. After working for the University as a research assistant he started working with the Royal Netherlands Air Force in 2005 as a Human Factors Specialist. His main areas of work included: CRM training design and delivery, pilot and air traffic controller selection and fatigue research. In 2008 he moved to the NLR where he is a training specialist within the department of Training, Simulation and Operator Performance. Research areas include performance based training, instructor competencies, lessons learned systems for training organizations and fatigue risk management.

Jelke van der Pal (1963) received his Ph.D. in 1995 (Educational Science and Technology, University of Twente, NL) and is Senior Scientist at the National Aerospace Laboratory NLR of the Netherlands. He has been active in aviation training R&D since 1996, which includes training analysis for fighter controllers, for the proposed pan-European military pilot school (Eurotraining), and for NATO-level networked simulator mission training. He has also performed simulator motion effectiveness studies for airlines and military applications. Finally, Dr. van der Pal has co-ordinated the European ADAPT-IT project, producing a training design tool that provides competency-based support for several phases of the Instructional System Design (ISD) cycle.

Christopher Roos (1984) obtained his MSc in Applied Cognitive Psychology at the University of Leiden in 2009. As Human Factors and Training specialist for the NLR he has carried out research on competencies, selection, training and automation issues in UAS operations, Ground Crew Station (GCS) Human-Machine Interface (HMI) design and other human factor and training aspects.

The Dynamic Team Training Experiment: Improving Tactical Team Decision Making

Edzard Boland, Jelke van der Pal, Christopher Roos
National Aerospace Laboratory NLR
Amsterdam, The Netherlands
boland@nlr.nl, pal@nlr.nl, Christopher.roos@nlr.nl

INTRODUCTION

The important role of good decision making in aviation safety is widely recognized. Decision making is a complex psychological construct in civil aviation, and more so in military aviation. Even though there are obvious commercial goals in civil aviation, the number one priority is safety. Risky situations are to be avoided. Safety, however, is just one of the outcomes in military aviation. Military pilots are constantly training for, and sometimes operating in, hostile environments, actually flying into high risk situations. Military air operation team members are usually dispersed in the air, flying single seat aircraft and coordinating with personnel on the ground, for example UAS operators. This dispersion enhances the complexity of an otherwise already complex team operation.

For current and foreseen air operations, there is an increasing focus on precision and effectiveness. Incidents, accidents, collateral damage, fratricide and loss of weapons need to be reduced. This requires, next to military specific measures such as real time updates on intel (intelligence), a highly flexible team operation, able to make correct and timely decisions. Teams have to deal with the situation at hand in a co-ordinated way, and be able to perform unplanned assignments. Higher operational demands require more precision and timely action from the team. Changes relative to the expected situation need to be detected and analyzed. Based on this analysis, options need to be generated that focus on achieving the mission objectives while minimizing the risks involved. How can training influence these cognitive processes?

This paper describes a new concept for air operations team decision making training intended to enhance team and operational efficiency and effectiveness. The effectiveness of this training concept has been evaluated in an experimental setup. With the experiment we tested the following hypotheses:

1. Using the DESIDE acronym for decision making within teams in highly complex, dynamic environments positively influences the quality of the decision making process over current standard CRM approaches.

2. A decision making process of higher quality will have a subsequent positive effect on mission output

Team work and team training

Almost any complex task or organizational goal requires a group of specialized individuals to succeed in an efficient and safe way. The teamwork required is usually facilitated by rules and procedures issued by the organisation or a higher authority (e.g., national legislation). In part, rules and procedures define team work: how and when information is gathered, processed, and coordinated and how to make decisions and take actions. While rules and procedures are useful and needed, they cannot cover all the dynamic and unpredictable situations in which teams may operate. Yet, teams need to be prepared for such ill-defined situations. An example of such a situation is, a two ship F-16 formation encountering air threats heading for friendly ground troops while en route towards a high priority moving ground target monitored by a UAV. This situation formed one of the missions executed by our participants.

Accurate competency descriptions are needed to ensure that training designed for team work focuses on the appropriate set of skills, knowledge and attitudes and ensure that learning environments provide appropriately realistic work conditions and circumstances.

Competencies are defined as integral sets of knowledge, skills, and attitudes as needed to perform operational tasks given defined operational conditions and standards (Abma & van Bavelgem, 2004). A team competency is a competency that is associated with the behaviour and the results of a team as a whole. Most team competencies do not supersede the individual level. When a person has 'good team skills', this reflects the skills for the individual, not the group. However, applying these individual competencies requires the team to function as a whole. Examples of team competencies are the attitude 'team cohesion' (degree to which team members value being a team member), or a 'shared mental model' knowledge structure such as knowing each other's roles and responsibilities.

Team related competencies may be very diverse and complex. To facilitate research and applications, many scholars have suggested team competency taxonomies. These taxonomies differ in terms of generality and completeness and their usefulness is often restricted to a particular application (kind of evaluation, kind of research, type of training). Most taxonomies divide team competencies in Technical (task-work related) & Non-technical (team-work related: organisational, social aspects) competencies (Bowers, 2000).

Technical competencies are specific to particular teams (e.g., execute a defence strategy for soccer team or setting up a local command centre for fire-fighters) and do not generalise over teams. Non-technical competencies tend to be more generic, although for particular professions certain non-technical team skills may be more important than others (e.g., 'problem diagnosis' for a soccer team versus a fire-fighter team). An example of a non-technical competency framework for a particular profession (civil cockpit crews) is provided by the NOTECHS project (Van Avermaete, 1998) specifying non-technical competencies and associated behaviors (see Table 1).

Table 1. The NOTECHS framework

Categories	*Elements*
Cooperation	Team building and maintaining Considering others Supporting others Conflict solving
Leadership and management Skills	Use of authority Maintaining standards Planning and coordinating Workload management
Situation awareness	System awareness Environmental awareness Assessment of time
Decision making	Problem definition/diagnosis Option generation Risk assessment Outcome review

The MilNOTECHS framework

MilNOTECHS was introduced with the DHC (Defence Helicopter Command, in The Netherlands in 2011. In 2012 it will also be implemented by the fixed wing transport squadrons. Plans exist for the introduction into the F-16 community as well.

MilNOTECHS is based on the civil variant, NOTECHS. Even though MilNOTECHS is a modification by the Royal Netherlands Air Force (RNLAF), 90% of the original content has been retained. Most changes have been made within the Situation Awareness (SA) and Decision Making (DM) categories. Several DM elements have been moved to the SA category. This has been done to create a clear boundary between SA and DM; reserving the actual making of the decision for DM, everything else, even 'outcome review' is SA.

Where the original NOTECHS suggested working with bipolar ratings (acceptable versus unacceptable), the MilNOTECHS adheres to the standard RNLAF grading system:

N (Not graded/demonstrated); example behavior has not been displayed and was not applicable in the graded situation.
U (Unsatisfactory/Unable); example behavior has not been displayed and this causes an unsafe situation.
F (Fair); example behavior is regularly not displayed and this causes possible risks for flight safety.
G (Good); example behavior is regularly displayed and flight safety risks do not arise or are neutralized in time.
E (Excellent); example behavior is continually displayed and enhances flight safety considerably.

Aeronautical Decision Making; Training and Tools

Several strategies (often embodied in mnemonics or acronyms) have been suggested to support the processes and procedures concerned with Aeronautical Decision Making (ADM), for example SHOR (Stimulus, Hypothesis, Options, Response) (Wohl, 1981) and DESIDE (Detect change, Estimate significance, Set safety objectives, Identify Options, Do, Evaluate) (Murray, 1997). Li and Harris (2005a) undertook a study to identify the best ADM mnemonic-based methods for training military pilots in decision making SHOR was rated as being the best ADM mnemonic in time-limited and critical, urgent situations. DESIDE was regarded as superior for knowledge-based decisions that required more comprehensive considerations but also had more time available.

In a later study, Li & Harris (2008) provided a short ADM training course to fighter pilots and evaluated their decision-making skills during a series of emergency situations presented in a full-flight simulator on the dimensions of situation assessment, risk management, and response time. Improvements were found in the quality of pilots' situation assessment and risk management, often at the expense of response speed. Li and Harris concluded that ADM is trainable and effective in improving decision making.

Instead of improving individual pilot decision making in emergency situations, the Dynamic Team Training (DTT) experiment aims to improve team decision making in a tactical environment, using the DESIDE acronym. The tactical mission scenarios were set up in such a way that most of the decisions required comprehensive considerations. The scenarios also provided participants enough time to apply the acronym or provided opportunities for the participants to create time. The strength of the DESIDE acronym is the inclusion of a deliberate decision whether or not to alter your mission objectives and to monitor if your actions have the expected results. For those reasons, DESIDE, not SHOR, was the selected ADM tool.

The DESIDE acronym

The DESIDE acronym can be used to improve the quality of decision making when sufficient time is available by following these six steps:

1. Detect change: monitor for any change that might influence the expected outcome of the flight.
2. Estimate the significance: estimate the significance of the change. With minimal risk a small corrective action might be sufficient. Otherwise you need to proceed to the next step.
3. Set safety objectives: modify your original objectives to obtain a safe/desired mission outcome. Be aware of the tendency to continue your flight as planned because you do not want to abandon your original flight objectives. A helpful tool to counter this tendency is to 'look ahead to how you might look back' when the risk leads to encountered danger (e.g.: loss of aircraft or lives).
4. Identify options: generate, evaluate and select an option that can successfully achieve the safety objectives.
5. Do: execute the option you selected in the previous step.
6. Evaluate: evaluate if your actions are achieving the safety objectives as expected.

METHOD

The Dynamic Team Training (DTT) experiment followed a randomized groups, repeated measures design. Participants were selected from the community around the 'Falcon 4.0' military flight simulation game. This community frequently organizes realistic virtual F-16 fighter jet campaigns. Members are familiar with F-16 controls and performing team based tactical sorties. Selected participants were all male, on average 44 years old and had accumulated an average of 1750 virtual flying hours. A total of 24 participants were divided over 2x (control and experimental) 4 teams of three persons. Each team consisted of two fighter jet pilots and one supporting Unmanned Aerial System (UAS) operator. Participants were randomly selected but roles were divided over the participants according to their experience and skill to form optimally functioning teams. In the experiment, each team was asked to perform a set of five tactical fighter jet missions. These missions all started with the two F-16s on the runway and the UAV loitering in the mission area. In the 20 to 35 minute missions that followed ground targets had to found, identified and destroyed, air threats had to be attacked or avoided, and unforeseen technical, tactical and weather issues had to be dealt with. In the control condition (CRM), the teams received a short classical Crew Resource Management (CRM) training, similar in content to the CRM training received in the RNLAF. Topics covered in this training were; human error, fatigue, stress, cooperation, communication, cooperation, situation awareness and decision making. Some topics, such as Fatigue, were only discussed briefly given their relevance to this particular experiment. In the experimental condition (DESIDE), the teams received a short presentation about the concept of CRM but only the topics 'situation awareness' and 'decision making' were covered in the same detail as in the control groups. The remaining training time focused on the use of the DESIDE decision making tool in a team setting and an explanation of some typical biases and decision making errors. The first two letters of the acronym, D(etect) and E(stimate) were to be used on an individual basis. If a pilot or the UAS operator estimated there was a consequence for the mission due to a detected change, they would communicate this to the team. For example, in one of the missions a mobile target had moved from its briefed position, a field, into a populated area. Attacking this target in this position would breach the Rules of Engagement. The UAS operator had the opportunity to spot this before the F-16 reached the target area. His Detect should be; target has moved into populated area. The next step, Estimate the consequences for the mission, should be; due to the Rules of Engagement this is not a valid target as long as it remains in this position. After communicating this message the remaining letters could be dealt with by the team in a way they considered appropriate. To prevent memory issues during the missions, all participants were encouraged to use a DESIDE checklist.

After the classroom training, team training continued with a series of five missions. Following each mission, the teams received feedback on their competencies. During the mission execution, the quality of the team performance, in particular their situation awareness and decision making, was monitored and rated using MILNOTECHS behavioral markers.

Participants

For the DTT experiment, teams of three people were formed. The participants were divided over the teams based on their knowledge with military fighter jet tactical operations and aircraft handling abilities. This information was collected beforehand from all of the participants. The most experienced participants would fulfill the role of Flight Lead, the least experienced participants the role of UAS operator. Participants with moderate experience would fulfill the role of Wingman.

The experiment was conducted with members from the 'Falcon 4.0 gaming community', based around the PC game 'Falcon 4.0'. This gaming community regularly organizes realistic virtual fighter jet campaigns. A number of simulator runs to test these gamers were organized in which a former RNLAF weapons instructor pilot evaluated aircraft and weapon systems handling and (tactical) procedural knowledge. His evaluation was used to decide if the level of performance of the participants was adequate for the experiment.

Experimental set up

The experiment was administered using F-16 research simulation facilities. Participants were screened on former formal Crew Resource Management (CRM) training and if applicable, excluded from the experiment. A good comparison could thereby be made between the effects of the different decision making tools. The teams of three participants would receive instructions and training on location before familiarization with the simulator and subsequent mission execution. The experiment consisted of a fighter jet campaign with a total of five separate missions. Every mission included an extensive briefing (to optimally prepare participants for the mission at hand) and debriefing (to facilitate the learning process). Instructions, training, familiarization and mission execution would, including a total of one hour of breaks, take up a total of 8 hours and 45 minutes.

Because the experiment would run for a prolonged period of time (up to several weeks) and participants were recruited from the same pool, there was a risk that teams that had undergone the experiment in the initial stages would inform later teams. To guard against this, participants were asked to keep strict secrecy about the training and missions until the experiment was completed.

Simulators and data collection instruments

To test the quality of the decision making process and the effects of that on the mission performance the experiment will be performed on flight simulators. The NLR simulator facilities 'Fighter 4 Ship' and 'Multi UAV Simulation Test bed (MUST)' were used. The same simulators could also be used for future validation on professional fighter pilots. The NLR Fighter 4-Ship (F4S) simulator was used for the tactical fighter jet (F-16) simulation. A (Medium Altitude Long Endurance (MALE)) UAS would be simulated using the NLR Multi UA Simulation Test bed (MUST). Both simulators were connected to the same tactical environment to enable the participants to receive inputs from each other, cooperate and work together as a team.

NLR Fighter 4-Ship (F4S)

The Fighter 4 Ship (F4S) is an F-16 fighter jet research simulator based at the RNLAF Volkel Air Force base that is comprised of four separate fighter jet simulators that are networked together to enable tactical team operations with up to four ships. F4S supports the whole mission-cycle by integrating the mission support systems, used in actual fast-jet operations by the Royal Netherlands Air force (RNLAF). The Fighter 4 Ship simulator is interoperable with the NLR MUST simulator that simulates a UAS ground control station, enabling the F-16 and UAS operators to work together for enhanced situational awareness (SA) and share certain tasks such as Battle Damage Assessment (BDA).

NLR Multi UAV Simulator Test bed (MUST)
The MUST is a desktop UAS (Unmanned Aerial System) simulator developed within NLR that facilitates UAS control by a single operator. The MUST was configured for this research to perform as a Medium Altitude Long Endurance (MALE) UAS (comparable to the General Atomics MQ-1B predator) with reconnaissance capabilities only. Because the MUST UAS simulator is interoperable with the F4S simulator, missions can be accomplished by coordination and sharing of information between the F-16 flight and the UAS operator.

Briefing/Debriefing

All briefing and debriefing was performed using the mission support systems that are used in actual fast-jet operations by the Royal Netherlands Air force (RNLAF). This facilitated the comparison of results from the study among the Falcon 4.0 participants with future further confirmation studies involving actual RNLAF F-16 fighter pilots. All briefing documents were prepared in close collaboration with a former F-16 weapons instructor pilot to ensure all necessary information was present for effective operations and contained campaign information, target folders and mission charts for both F-16 crew and UAS operator.

Communications

To facilitate the crucial team communication among the F-16 flight crew and between the UAS operator, a VHF communication system representative of the actual communication system currently in use in the RNLAF was used for all communications.

Missions

The dynamic team training experiment used preplanned events as a trigger for the use of the decision making tool (either CRM or DESIDE). These triggers were built into the missions that had to be executed in the form of malfunctions (fuel leaks), unexpected occurrences (bad weather, moving targets) and rare deviations from normal operations (red force defector). All five missions were designed in cooperation with an experienced weapons instructor pilot to contain all desired tactical elements and appropriate difficulty level. Two examples of missions are presented below:

<u>Mission 3:</u> Guard Fighter Area Of Responsibility (FAOR).
In this scenario, two threats appear. First, an air threat is reported by AWACS. Second, a Tactical Ballistic Missile (TBM), located close to a church will be observable by the UAV. When the UAV detects the TBM, the HQ communicates the attack order: to avoid collateral damage, the attack may commence only if the TBM starts moving out of residential area. An engagement with the air threats should be avoided because, as briefed, the TBM is a higher priority target and attacking it would involve less risk than attacking the air threats.

<u>Mission 5:</u> Attack SA-6 (target 1) and bridge (target 2).
In this final scenario a SA-6 Surface to Air Missile (SAM) system is the last tactical asset available to the enemy. If this asset is taken out, the war is won, and the second target (bridge) won't have to be taken out. The UAV detects the SA-6 but the target is concealed by a low level cloud deck. Target 2 (bridge) is not concealed by bad weather. The team has to decide on whether or not the SA-6 can be taken out, whether or not to take out the bridge, and how this can be done without causing unnecessary damage

In these scenarios, the DM Process Quality was primarily associated with the following behavioral markers: Environmental Awareness, Problem Definition, Option Generation and Risk Assessment.

Data collection

Data was collected in a twofold manner. The quality of the decision making process was measured on the basis of observations through standardized observation criteria by a training expert and a Subject Manner Expert (SME) on tactical fighter jet missions. The training expert measured the quality of the decision making process by using the MILNOTECHS method, specifically designed to measure non-technical skills among fighter pilots. The SME provided relevant insights (Instructor Pilot (IP) ratings) into the quality from a tactical point of view (e.g. risks taken).

One or more decisive moments were created per executed mission to ensure the participants had to make a decision on the course to follow. This would provide fixed reference points to allow the instructors to rate the performances of the teams. The focus was on the decisions made at these decisive moments, but decisions and actions before and after the decisive moments would also be taken into account as they could be of influence as well. This measure was used as input for the first hypothesis.

The SME also provided a measure of the mission success from a tactical point of view for the overall mission, as well as for a subset of mission elements (e.g. flight path taken, sub goals met). This measure was used as input for the second hypothesis.

As measure of the DM process quality a composite of two MILNOTECHS subcategories were used; 'Situation Awareness' and 'Decision Making'. The average ratings on these items for the team on every separate mission were used as input for GLM/repeated measures ANOVA (SPSS).

RESULTS

The experiment revealed a learning effect on the quality of the decision making process ($F = 11.57$, $p < 0.05$; see Figure 1). The experimental group receiving the DESIDE acronym decision making process tool showed better progress on decision making process quality than the control group that received a standard CRM decision making training ($F = 9.81$, $p < 0.01$; Bonferroni post hoc tests).

The average decision making process quality over the five missions did not differ significantly between groups: $M = 1.84$ for DESIDE versus $M = 1.41$ for CRM ($F = 2.27$, $p = 0.19$).

On individual missions 3 and 4, the DESIDE group received higher mean scores on decision making process quality than the control group (1.98 versus 1.44 on mission 3; 2.19 versus 1.93 on mission 4). These differences are not significant. However, for the last mission, the difference between the experimental group ($M = 2.38$) and control group ($M = 1.06$) is evident ($t = 2.53$, $p < 0.05$).

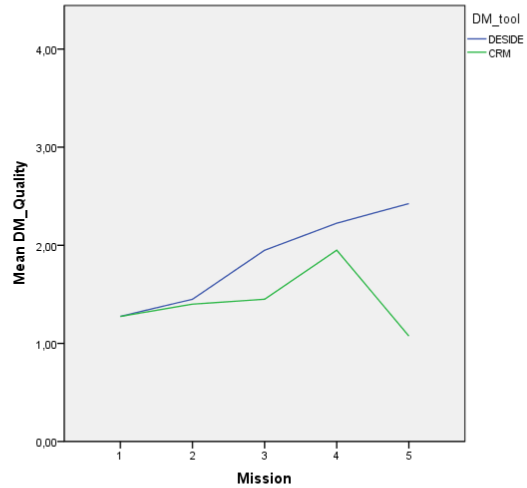

Figure 1. Mean decision making process quality for experimental and control group (1= unsatisfactory, 4 = excellent) over mission 1 to 5.

DISCUSSION

The results, although hypothesized, are remarkable since they are found with a sample size of only four teams per condition. A significantly different learning curve in decision making process quality was found for the DESIDE group in comparison to the control (CRM) group. However, no significant difference between the average decision making process quality in the control and experimental groups was found. With the limited amount of time available, this result was not unexpected.

Because the experiment had to be performed over the course of one day, only limited time was available for training and learning to use the decision making tool. The Falcon 4.0 gamers were not able to apply the decision making tool to the full extent, i.e., they did not overtly complete the decision process using all the letters of the acronym.

If we hypothesize that the differences between the two learning curves continues with subsequent training missions, this would result in a significant difference in decision making process quality. Using the DESIDE acronym to increase the decision making process quality therefore seems promising.

Limited overt use of the DESIDE tool may be attributed to high workload, caused by the effects of flying on a different and more realistic simulator, and executing more realistic task sets (compared to Falcon 4 devices and missions). Another factor that may have contributed to high workload was the required 'train as you fight' regime compared to the more ‘if there is chance to engage an air threat, don’t bother about tasking orders, rules of engagement, fuel levels, or mutual support and go for it’ work practice in gaming.

Observations by the instructors do support a high workload explanation. When overloaded, communication often is one of the first human capabilities to disappear (Kleinman & Serfaty, 1989). The gamers were able to keep on flying and performing their individual tasks, but were not always able to communicate their intentions and decision making considerations, which would include using the DESIDE tool. Instructors also noted occasional signs of mental overload, manifesting itself in physical (heavy breathing) and cognitive (acting disorderly) forms.

The observations of the instructors, however, are not supported by the self-ratings of the participants on the Rating Scale Mental effort (RSME; Zijlstra, 1993) as figure 2 illustrates. An RSME score of 70 indicates “effortful” which should not generate overload.

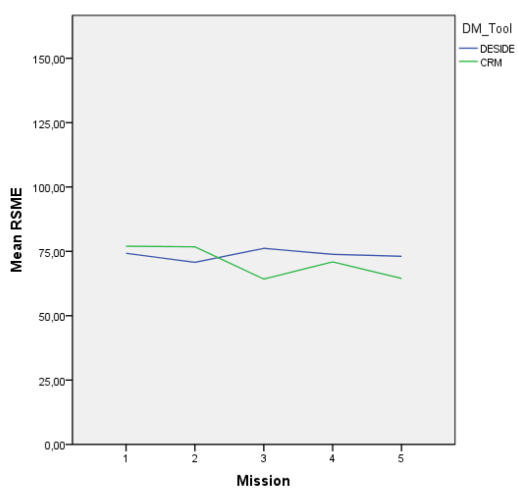

Figure 2. Mean RSME for experimental and control group over the five missions (scale from 0-150; 40 = somewhat effortful; 85 = very effortful; 115 = tremendously effortful).

Even though the decision making tool was not applied to the full extent, a marked difference is seen between the DESIDE and the control group on the latter mission, in particular in the ambiguous circumstances of mission 5. By giving the participants a tool to actively guide them through the decision making process, the quality of the decision making process increased. Even though high workload often seemed to prevent participants applying the full acronym during the missions, the first letter (D for Detect) was used many times. This could be seen as a considerable increase in SA compared to the control group in which most of these changes in the environment were not detected (as far as apparent from subsequent behavior). A considerable part of the increase in quality of the DM process can therefore be attributed to an increase in SA.

Several control variables have been measured to control for and help interpret the results: participant background characteristics, the Crew Alertness Rating Scale, the workload scale RSME and Stanford Sleepiness Scale. Full covariate analysis was not feasible due to the small sample size. These variables do not reveal any clear explanation for the effects. Participant background differs considerably between the 24 participants, obviously a much more heterogeneous group than the F-16 pilot community, but differences are not strong between the groups. Workload experienced (RSME) is not different between groups, although the DESIDE group reported significantly higher levels of sleepiness, higher experienced difficulty, and higher experienced stress levels after each mission. These self-reported ratings may indicate either a higher motivation level in the DESIDE group (not observed as such by the instructors) or may be the result of the additional task load resulting from (trying) to use the DESIDE tool.

For practical reasons the participants in this study were Falcon 4.0 gamers, who are quite knowledgeable with F-16 controls and mission flows. But they are not representative for the average fighter pilot. The DESIDE concept and the results of the study have been presented to operational pilots to acquire an indication of face validity of the concept. The F-16 pilots (one weapons instructor pilot, one experienced pilot, and one recently graduated pilot) received the same DESIDE decision making briefing and executed the first and fifth mission. After both missions they were debriefed by the same weapons instructor pilot used during the experiment. The fighter pilots indicated that, in their opinion, the DESIDE acronym could be a beneficial factor in further establishing the ‘speak up culture’. This will especially benefit novice fighter pilots that, due to the practice of sharing only important information over the busy radio channels, may be uncertain to determine whether information is irrelevant and consequently may not share relevant information in a timely manner. Using the DESIDE tool will make it more acceptable for all to have inexperienced pilots expressing a ‘detect’, even when the flight lead may decide to ignore the information. Furthermore, novice pilots may experience less difficulty in using a new way of decision making.

The reported downside of the DESIDE acronym was the cognitive load put on the pilot. The cognitive strain provided by deliberate and active use of a decision making tool might be handled better by an experienced pilot, who has more spare cognitive capacity for additional tasks. In that respect the experienced pilots would be the more appropriate target group for a DESIDE training.

For novice pilots learning to use the DESIDE tool, a role shifting paradigm might be applied. It may be good training value in its own right to take up the role of a UAS operator in a mission, and at the same time (after UAS familiarization), learning to apply the DESIDE tool in a team setting. The relatively low workload for the UAS operator may enable the novice fighter pilot to apply the DESIDE tool in a less demanding context.

Alternatively, training scenario's may be set up in such a way that (experienced) UAS operators joining a training mission would be more likely to detect and/or estimate consequences, and therefore initiate the DESIDE process. This would also ease the load on novice pilots.

The DESIDE DM tool could also help in creative problem solving and preventing Plan Continuation Errors (PCE) (Orasanu, et al. 2001). It is often difficult for humans to take a step back from an engaging task. This may prevent the execution of a conscious decision making process, possibly leading to a PCE. Even if the application of DESIDE would not lead to better mission quality (accomplish mission goals) it will most probably lead to better cost-benefit analyses due to the deliberate risk analysis and option generation steps engrained in the DESIDE process. That results in mission goals being accomplished with a minimum of risk.

Subsequent research should establish the ecological validity of the results. Does the effect also occur when training (novice) pilots and what are the effects after a prolonged period of training? While novice pilots may appear to be appropriate target audience for DESIDE training, they may also be susceptible for the same workload effects as the Falcon 4.0 gamers. The training method needs to be adapted to address this.

It is expected that the learning curve for the decision making process would continue to rise when the participants receive more training over a prolonged period of time in future studies. Also, application of the (complete) DM tool will probably increase in a more familiar environment, and therefore require lower work load.

Another aspect of interest for further study is the scope of applicability and the long-term effects of the DESIDE tool. DESIDE has been designed for structuring decision support for problems that allow a few minutes to be solved which fits the processes requiring team co-ordination in this study. Which type of teams, or missions would benefit by DESIDE and which would not? Also, when novice pilots learn to structure decision making using DESIDE, will it eventually internalize? Will the problem solving strategy routinize, become more implicit and therefore expedite the decision making process to such a level that it also enhances the (individual) decisions that need to be made in a shorter (e.g. ten second) time frame?

CONCLUSION

It has been found that the DESIDE tool for decision making is effective for teams in a military context. It does generate improved decision making process quality in teams of gamers performing air operation missions with an F-16 two-ship and a UAS operator. Further study is needed to investigate the validity of the effect and the scope of applicability of the DESIDE tool for team training in the professional military environment.

REFERENCES

Abma, H.J. & Van Bavelgem, J. (2004). Training Needs Analysis for Human Factors Practitioners in Aviation Maintenance. CR-2004-223. Amsterdam: National Aerospace Laboratory NLR

Avermaete, J. van, et al. (1998) NOTECHS: Non-technical skill evaluation in JAR-FCL. NLR-TP-9851

Bowers, C.A., Jentsch, F. & Salas, E. (2000). Establishing aircrew competencies: A comprehensive Approach for identifying CRM training needs. In O'Neil Jr., H.F. & Andrews, D.H. Aircrew Training and Assessment. Lawrence Erlbaum Associates: Mawwah., p. 67-84.

Kleinman, D.L. & Serfaty, D. (1989) Team Performance in distributed decision making. In R. Gilson, J.P. Kincaid & B. Goldez (red.), Proceedings of the Interservice Networked Simulation for Training Conference, Orlando. University of Central Florida 1989, p. 22-27.

Li, W.C., and Harris, Don (2008) 'The Evaluation of the Effect of a short Aeronautical Decision-Making Training Program for Military Pilots', International Journal of Aviation Psychology, 18:2, 135 — 152

Li, W.C., and Harris, D. (2005a). Aeronautical decision-making: Instructor-pilot evaluation of five mnemonic

methods. Aviation Space Environmental Medicine, 76, 1156–1161.
Murray, S.R. (1997). Deliberate Decision Making by Aircraft Pilots: A Simple Reminder to Avoid Decision Making Under Panic. The international Journal of Aviation Psychology, 7: 83-100.
Orasanu, J., Ames, N., Martin, L., & Davison, J. (2001). Factors in Aviation Accidents: Decision Errors. In E. Salas & G. A. Klein (Eds.), Linking expertise and naturalistic decision making (pp. 209-225): Mahwah, NJ: Lawrence Erlbaum Associates: 2001.
Wohl, J. G. (1981). Force management decision requirements for air force tactical command and control. IEEE Transactions on Systems, Mans, and Cybernetics, SMC-11, 618–639.
Zijlstra, F.R.H. (1993). Efficiency in work behavior. A design approach for modern tools. PhD thesis, Delft University of Technology. Delft, The Netherlands: Delft University Press.