

A GENERIC ASSESSMENT TOOL FOR EVALUATING C2 EXERCISES

Marcel van Berlo & Jan Maarten Schraagen
Netherlands Organization for Applied Scientific Research (TNO)
TNO Human Factors (TNO HF)
Soesterberg, The Netherlands

Abstract

The commander of the Royal Netherlands Airbase Volkel has commissioned a study to obtain a clear insight into the process of command and control (C2), with the objective of assuring the quality of the output of C2. To enable this goal, TNO HF developed a generic assessment tool for evaluating the performance of C2 teams: the Command & Control Process Measurement Tool (C2PMT). The C2PMT comprises concrete and clearly observable performance indicators on the basis of which the process of C2 teams can be assessed. These specific performance indicators are based on interviews with key commanders of Airbase Volkel, and on a review of the relevant literature. A prototype of the C2PMT was successfully tested during a three-day exercise. In this paper, the development of the C2PMT will be described. First, the problems and questions of the Airbase will be presented. Secondly, team performance and performance indicators, as identified by the literature and field studies that have been conducted, will be commented upon. Thirdly, the development and prototyping of the C2PMT will be discussed. The final section concludes with future research and development issues.

Biographical Sketches:

Marcel van Berlo is an educational scientist and a researcher within the Team Training Research Group at TNO Human Factors. He has experience in research on developing methodologies for instructional systems development, especially with respect to team performance and troubleshooting. The field of research is primarily aimed at technologically sophisticated learning environments, like (multi-media) computer-based training, simulator and simulation. His Ph.D.-research includes the development and validation of a methodology for designing team training systems.

Jan Maarten Schraagen holds a Ph.D. in cognitive psychology and is program manager of the Distributed Decision Making Research Group at TNO Human Factors. He has experience in fundamental research in human memory and problem solving, applied research in team decision making, fault diagnosis, and in-car information presentation. He has expertise in experimental methodology, knowledge elicitation techniques, and use of simulators in nautical training. Currently, he is on leave at the Naval Air Warfare Center Training Systems Division (NAWCTSD), Orlando, Florida.

A GENERIC ASSESSMENT TOOL FOR EVALUATING C2 EXERCISES

Marcel van Berlo & Jan Maarten Schraagen
TNO Human Factors (TNO-HF)
Soesterberg, The Netherlands

BACKGROUND

The Royal Netherlands Airbase Volkel is one of the Dutch bases deploying F-16 squadrons during military conflicts. In order to maintain operational readiness, this Airbase conducts several exercises each year. Emphasis is placed especially on 'command & control' (C2) as this is viewed as the key element of effective team performance. However, the current exercises have several drawbacks:

- Many persons are involved. Consequently, the preparation is time consuming, the execution is large-scale and therefore difficult to supervise, and, as a result, the debrief and feedback to the participants is troublesome. In addition, because of the large scale it is not possible to train on a more frequent basis.
- Preparing actions and making decisions is automatically followed by actually executing these actions by the lower-echelon personnel. Because this is labor intensive and time consuming, the commander of the Airbase is looking for alternative means of training the higher echelon C2 function separately and thereby more frequently.
- Although a lot of information is gathered during an exercise, it is still uncertain to what extent the most adequate performance indicators are used by the evaluators, especially with respect to command and control. This has the risk of an incomplete assessment of the quality of C2, and evaluators not being able to give specific feedback and the necessary recommendations for improving the C2 process.
- Because of a lack of standardization in evaluating exercises, it is hardly possible to monitor the C2 team's progress in time.

Command and control is evaluated in three areas, namely 'operations' (OPS), 'logistics' (LOG) and 'survival to operate' (STO). OPS

relates to preparing pilots and aircraft conducting the operational missions; LOG relates to all logistic activities required for getting aircraft airborne; STO relates to the passive and active defense of the base, and to damage control activities in order to maintain or restore the base's operational readiness. The standards an operational unit (up to the level of an entire base) should meet are described in the STEM (SHAPE¹ Tactical Evaluation Manual). Being valid for all of NATO's Airforce units, the STEM is by nature generic. Although standards are formulated with respect to the final outcome of the mission accomplishment, the process of how to achieve the results is hardly described. For every separate area (OPS, LOG and STO) one general norm is stated on which the base can achieve a grading being 'excellent' (EX), 'satisfactory' (ST), 'marginal' (MA) or 'unsatisfactory' (UN). In order to assess to what extent this norm has been met, several *amplifying notes* are formulated. These amplifying notes describe in more concrete form the demands operational units should meet, and could therefore be a tool for the evaluators during base-exercises. However, evaluators encounter many difficulties in applying these amplifying notes because they vary in completeness, concreteness and clarity. Consequently, the evaluators are largely thrown back on their own subjective interpretation of the amplifying notes in order to determine the gradings. Ultimately, this has the risk of training benefits and lessons learned remaining unclear.

Facing an upcoming Operational Evaluation (OPEVAL) by NATO, the commander of Airbase Volkel stated the need to obtain a clear insight into the process of C2, with the objective of assuring the quality of the output of C2. In this respect, the following questions are relevant:

¹ Supreme Headquarters Allied Powers Europe.

- (1) which standards for C2 can be developed?
- (2) how can these standards be measured?
- (3) how can C2 exercises be evaluated?
- (4) which recommendations can be provided in order to improve the C2 process?

Answering these questions resulted in the development of an assessment tool for evaluating the performance of C2 teams: the Command & Control Process Measurement Tool (Van Berlo & Schraagen, 2000a/b/c). Before describing this tool, the next section discusses team performance and performance indicators as identified by key Airforce personnel and in literature studies that have been conducted.

TEAM PERFORMANCE AND PERFORMANCE INDICATORS

In order to gain more insight into (a) the hierarchical command structure, (b) the tasks, responsibilities and authorities of the commanding officers, and (c) the performance of the several command teams, structured interviews have been conducted with ten key commanding officers. Each interview lasted about 1.5 hours and was conducted by two TNO-HF researchers and one project member of Airbase Volkel. A report was made up of every interview, which was subsequently offered to the respective commanding officer with the request of checking the correctness of the contents.

The most relevant results of the interviews are related to the flaws in the C2 process and the characteristics of good C2 teams. A summary of these results is described next. Flaws in the C2 process that have been identified fall in the categories of information exchange, communication, supporting behavior, initiative/leadership (see Smith-Jentsch, Zeisig, Acton, & McPherson, 1998), task/responsibilities/authorities, (re)planning/adaptation, and physical workspace. Some examples of the respective categories are summarized below.

Information exchange:

- one receives little information from other personnel, hence much effort has to be put in gathering the information oneself;
- it is not always clear where to get the relevant information

Communication:

- communication lines are occupied as a result of non-relevant information exchange

Supporting behavior:

- much information is redirected to the best performing team member, resulting in an information overload of this particular team member

Initiative/leadership:

- no clear priorities are stated by higher commanding officers

Task/responsibilities/authorities:

- it is not always clear who has which task, responsibilities and/or authorities

(Re)planning/adaptation:

- taking over the command by an alternate command post is often laborious

Physical workspace:

- members of a command team are physically separated by walls hindering adequate communication

In addition to the interviews, a literature study was carried out on models of team performance, performance measurements and performance indicators. The selection of models is largely based on Militello, Kyne, Klein, Getchell & Thordsen (1999), as this article provided an up-to-date overview and integration of current models of team performance in Command & Control settings. The following models have been reviewed: Team Evolution and Maturation Model (Morgan, Glickman, Woodward, Blaiwes, & Salas, 1986), Teamwork Model (McIntyre & Dickenson, 1992), Team Performance Model (Fleishman & Zaccaro, 1992), Model of Organizational Competence (Olmstead, 1992), Flightcrew Performance Model (Helmreich & Foushee, 1993), Advanced Team Decision Making (ATDM) (Zsombok, Klein, Kyne, & Klinger, 1993), Comprehensive Model of Team Performance (CMTP) (Militello, Kyne, Klein, Getchell, & Thordsen, 1999), Anti-Air Teamwork Observation Measure (ATOM) (Smith-Jentsch, Zeisig, Acton, & McPherson, 1998). Detailed discussions of these models are provided by the respective authors, or can be found elsewhere (Van Berlo & Schraagen, 2000a; Swezey & Salas, 1992; Kokorian, 1995; Brannick, Salas & Prince, 1997; Van Berlo, 1998).

The Comprehensive Model of Team Performance (CMTP) has been chosen because of the comprehensiveness of this model. This model needs to be attuned to the specific situation at the Airbase Volkel, however, because on the one hand certain dimensions need not be relevant and can be left out, while on the other hand particular dimensions may need to be added because of the local relevance. Focusing on the situation of the Airbase has been done by relating the characteristics of good team performance, as indicated by the key commanding officers in the interviews, to the CMTP. Furthermore, the characteristics of

other kinds of teams (management, sports and project teams) were related to the CMTP. This could result in the identification of dimensions that are not covered by the CMTP, but are relevant to the Airbase Volkel. The characteristics of these teams are derived from Larson and LaFasto (1989) and from Cohen and Bailey (1997). An overview of the results is depicted in Table 1. On the highest level, eight components are distinguished, each consisting of several dimensions. Depending on the particular domain and team, these dimensions need to be further elaborated in terms of clearly observable behavior.

Components of the Comprehensive Model of Team Performance (Militello <i>et al.</i> , 1999)	Interviews with key commanding officers of Airbase Volkel	ATOM (Smith-Jentsch <i>et al.</i> , 1999)	Larson & LaFasto (1989)	Cohen & Bailey (1997)
1. Team competencies:				
- Member-leadership competence	✓	✓	✓	
- Shared practices (SOP proficiency)	✓	✓		
2. Team identity:				
- Defining roles and functions, resources	✓	✓	✓	✓
- Engaging all members	✓		✓	✓
- Compensating and coaching	✓	✓	✓	
- Interpersonal aspects			✓	✓
3. Team planning and decision making:				
- Envisioning goals	✓		✓	✓
- Maintaining dynamic focus	✓	✓		
- Situation assessment	✓	✓		
- Articulating expectations				
- Envisioning and evaluating courses of action (synchronization)	✓			
4. Team self-management:				
- Monitoring	✓	✓	✓	
- Adjusting	✓	✓	✓	
- Detecting gaps and inconsistencies	✓	✓	✓	
- Time management	✓	✓		
Elements falling outside of the CMTP:				
5. Organizational context (external support,	✓		✓	✓

rewards, training, resources)				
6. Environmental factors (e.g. physical space, turbulence)	✓			✓
7. Task design (autonomy, interdependency)				✓
8. Group structure (e.g. size, demography, diversity of team members)	✓			✓

Table 1: Comparison of the interviews and models; a checkmark indicates that the component is covered by the interviewees or the models.

Based on the results of Table 1, it can be concluded that two dimensions of the CMTP, namely 'interpersonal aspects' and 'making explicit the expectations considering the upcoming actions', were not mentioned during the interviews. However, because of their relevance as stated in the literature, we decided to include these dimensions in the design of the C2 assessment tool. Aspects not being included by the CMTP but mentioned during the interviews refer to 'environmental factors', 'organizational context' and 'group structure'. Another aspect falling outside both the CMTP and the interviews refers to 'task design'. These categories do have an impact on C2, but merely on the long term (e.g., group structure and task design were considered as 'givens' when evaluating a particular C2-exercise); it was therefore decided to leave these outside the scope of this research project.

Based on the identified characteristics of effective team performance and likely performance measures and performance indicators, concrete form has been given to the STEM, resulting in designing and prototyping the Command & Control Process Measurement Tool (C2PMT). This will be discussed in the next section.

DESIGNING AND PROTOTYPING THE C2PMT

The commander of Airbase Volkel required that the C2PMT be based on the original STEM, this being the official standard used by NATO evaluators. Analogous to the

STEM, in the C2PMT two categories are distinguished: 'resources' and 'performance', which will be discussed in the following two sections. Next, the try-out of the C2PMT is described. Finally, the generalization of the tool to non-MDF (Main Defense Force) operations will be briefly explained.

Resources

The category 'resources' concerns the personnel, materials, infrastructure, and relevant documents. Preliminary to the conduct of an exercise, the availability and quality of the resources can already be assessed. Table 2 depicts the format of the resources assessment.

Depending on the specific command team and its specific command post, certain aspects could be stressed or left out.

Performance

The category 'performance' concerns the base specific C2 organization, checklists and standard operating procedures on the one hand, and conducting, monitoring and checking the task performance of the C2 team on the other. Based on the literature study and the interviews, every amplifying note as formulated in the C2 chapters of the STEM (for both OPS, LOG and STO) has been given concrete form by specific performance indicators. These performance indicators are concrete and clearly observable behaviors enabling the evaluators to assess, in a more objective way than is currently being done, to what

extent the C2 team performance meets the standards. In addition, in order to gain a deeper understanding of the C2 process,

some additional amplifying notes have been formulated.

Criteria	Performance indicators
Physical layout	<ul style="list-style-type: none"> • enough space/room • audibility of team members • visibility of team members • readability of information
Information systems	<ul style="list-style-type: none"> • available • operational
Communication devices	<ul style="list-style-type: none"> • available • operational
Totes, maps, displays	<ul style="list-style-type: none"> • available • readable • up-to-date
Books, documents	<ul style="list-style-type: none"> • available • up-to-date
Personnel	<ul style="list-style-type: none"> • sufficient number available • level of training (of both individuals and teams)

Table 2: Format of the resources assessment

Examples of performance indicators are: adequately using the communication devices, requesting crucial information if not provided in time, brief and relevant communication, exchanging correct and consistent information, exchanging information without unnecessary delay, informing the right persons, providing unsolicited information, expressing an adequate security awareness, information within the command team being coordinated, stating the right priorities, monitoring the time available, every team member is actively engaged, supporting each other without neglecting one's own task, and being susceptible to team members' comments. The format of the performance assessment will be discussed next (see Table 3).

The heading of the table contains the amplifying note as depicted in the STEM, or an additional amplifying note. Every amplifying note is briefly clarified and explained in order to ensure a uniform interpretation by the evaluators: it describes the contents and coverage of the amplifying note and, if applicable, the relation with other amplifying notes.

The left column comprises the performance indicators that give concrete form to the

particular amplifying note enabling the evaluators to observe and interpret the C2 process. These performance indicators are formulated concisely, and are easily scored in terms of whether the behavior was observed or not; this can be indicated in the respective column (Yes or No). In the right hand column, the evaluator explains and illustrates the rating: this contains both positive and negative examples being observed. Inclusion of these example behaviors is important for providing feedback in the final written report and for enhancing learning opportunities. Completeness and usability of the performance indicators are balanced as well as possible. Because of a partial overlap, various amplifying notes could be given concrete form by rather similar performance indicators. If possible, these redundancies are avoided in order to prevent the same behavior being rated in several categories. After having filled out a table covering one amplifying note, a provisional grading can be determined (EX: excellent, ST: satisfactory, MA: marginal, UN: unsatisfactory, NG: not graded). Every provisional score is transported to an overview of results, providing for a concise summary of the C2 team's performance. In this way, the

evaluator can determine the final overall grading reliably and objectively.

1. Amplifying note (either original or additional) <i>Brief explanation of amplifying note to ensure a uniform interpretation.</i>			
Performance indicators:	Yes	No	Explanation/illustration:
a) Performance indicator 1			
b) Performance indicator 2			
i) Performance indicator n			
<i>Provisional grading of this amplifying note:</i>			EX ST MA UN NG <i>(Transport to overview of results)</i>

Table 3: Format of the performance assessment

Try out

A prototype of the C2PMT was tested during a three-day exercise. Fifteen C2 evaluators from three different Dutch bases have used the prototype tool. In the morning of the first day, the evaluators were briefed, and a handout covering the usage of the tool was provided. During both the conduct of the exercise and the writing of the final evaluation report afterwards, the project team members of TNO-HF and Airbase Volkel provided support. The evaluators were enthusiastic about the prototype C2PMT. Especially with respect to the performance indicators, giving concrete form to the rather general amplifying notes, the tool was regarded as helpful. This was reported not only by beginning evaluators, but by more experienced evaluators as well. Based on the evaluators' comments and the project team members' observations, some performance indicators were slightly reformulated.

Generalization

In the first instance, the C2PMT has been designed specifically for evaluating exercises in a Main Defense Force (MDF) context (Van Berlo & Schraagen, 2000b), because this will be the focus of the upcoming NATO's OPEVAL. In an MDF-

context, the base is involved in a large-scale military conflict, meaning that F-16's should get airborne and conduct their missions, and that the base should be defended against enemy fighters, (nuclear, chemical and conventional) missiles and ground-based troops. Besides the specific MDF-version, however, the commander of Airbase Volkel would like to have a similar tool for assessing C2 in other contexts (both war and peacetime) as well. Therefore, a generic version of the C2PMT has been developed (Van Berlo & Schraagen, 2000c), which can be adjusted to the specific C2 team and the specific context.

Although the generic version of the C2PMT is based on the STEM, it is not restricted by its format. Consequently, each of the three areas (OPS, LOG and STO) has an identical formulation and classification of the amplifying notes. First, general amplifying notes are described being valid for all C2 teams. Next, area-specific amplifying notes are presented. Finally, C2 related amplifying notes as described in other chapters of the STEM (but also given concrete form by performance indicators) are presented in order to have a complete overview on the C2 team's performance. Consequently, the C2PMT has a more consistent structure, enabling training results to be compared more easily.

CONCLUSION

This paper discussed the design and development of the Command & Control Process Measurement Tool that enables the assessment of the quality of the performance of C2 teams. Two versions of the C2PMT have been developed, namely an MDF and a generic version. Based on the try-out and further exercises, the C2PMT has proven to have the following advantages:

- the amplifying notes of the STEM are explained to ensure a uniform interpretation;
- additional amplifying notes have been formulated to gain a more complete insight into the C2 process;
- concrete and clearly observable performance indicators have been formulated in order to assess the C2 team's performance in an objective way;
- especially inexperienced evaluators are supported during the assessment of the C2 process and the development of the assessment report;
- the results of an exercise are easier to interpret;
- the results of various exercises are mutually comparable;
- the lessons learned can be determined more easily;
- follow-on actions can be determined in a more structured way.

By emphasizing the process of C2 rather than the outcomes, alternative exercises can be (and have already been) organized, like office simulations or command post exercises. In this way, C2-teams can be trained separately from the lower-echelon personnel, thereby enhancing the efficiency of the exercises.

The C2PMT provides the means for gathering facts and data concerning the quality of C2 processes, and therefore has the potential to determine follow-on actions not exclusively related to training issues, but also to the real-life (quality) management of organizational processes. Possible follow-on actions could relate to, for instance, the quality and availability of checklists, the physical layout of a workspace, and the security awareness of the personnel.

Based on the results, the commander of Airbase Volkel has decided to implement the

C2PMT in its organization in order to continuously monitor and improve the quality of the C2 process.

FURTHER RESEARCH

Further research could be aimed at developing an electronic version of the C2PMT. Currently, the tool is paper-based and requires evaluators to riffle through the pages. Besides, they have to monitor the progress of the exercise scenario as described in the script. An electronic version of the tool, like for instance SHIPMATE (Pruitt, Burns, Wetteland, & Dumestre, 1997) could further enhance the evaluators' performance.

It would be interesting to apply the C2PMT in the training of civil emergency management teams. A major difference with military C2 teams is that the members of an emergency management team normally (i.e., in non-emergency situations) conduct different tasks in different teams and organizations. Learning to operate in a rather new team, and to coordinate with other teams at various hierarchical levels, poses particular training issues, especially concerning the team performance measurement. To what extent the C2PMT can be applied to the training of emergency management teams needs to be examined.

Acknowledgment

This research project was sponsored by Airbase Volkel of the Royal Netherlands Airforce; the authors would like to thank colonel Eikelboom, major Op den Kamp and captain Soetenga for their kind support.

REFERENCES

- Brannick, M.T., Salas, E. & Prince, C. (Eds.) (1997). Team Performance Assessment and Measurement: Theory, Methods, and Applications. Mahwah, NJ: Lawrence Erlbaum.
- Cohen, S.G., & Bailey, D.E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. Journal of Management, 23, 239-290.
- Fleishman, E.A., & Zaccaro, S.J. (1992). Toward a taxonomy of team performance functions. In R.W. Swezey

- & E. Salas (Eds.), Teams: their training and performance (pp. 31-56). Norwood, NJ: Ablex.
- Helmreich, R.L., & Foushee, H.C. (1993). Why crew resource management? Empirical and theoretical bases of human factors training in aviation. In E.L. Wiener, B.G. Kanki, & R.L. Helmreich (Eds.), Cockpit resource management (pp. 3-45). New York, NY: Academic.
- Kokorian, A. (1995). Military Crew Competence Model (MCCM): a review of the crew performance literature and a model for diagnosing and measuring military crew competence (6/DSTI/CFH/SF/96). Hampshire, UK: Délégation Générale pour l'Armement (DGA) & Defense Research Agency (DRA).
- Larson, C.E., & LaFasto, F.M.J (1989). TeamWork: What must go right / What can go wrong. Newbury Park, CA: Sage.
- McIntyre, R.M., & Dickenson, T.L. (1992). Systematic assessment of teamwork processes in tactical environments (Report submitted to Naval Training Systems Center under Contract No. N61339-91-C-0145). Norfolk, VA: Old Dominion University.
- Militello, L.G., Kyne, M.M., Klein, G., Getchell, K., & Thordsen, M. (1999). A synthesized model of team performance. International Journal of Cognitive Ergonomics, 3(2), 131-158.
- Morgan, B.B., Glickman, A.S., Woodward, E.A., Blaives, A.S., & Salas, E. (1986). Measurement of Team Behaviours in a Navy Environment (Technical Report 86-014). Orlando, FL: Naval Training Systems Center, Human Factors Division.
- Olmstead, J.A. (1992). Battle Staff Integration (IDA Paper P-2560). Alexandria, VA: Institute for Defense Analysis.
- Pruitt, J.S., Burns, J.J., Wetteland, C.R., & Dumestre, T.L. (1997). SHIPMATE: Shipboard mobile aid for training and evaluation. Proceedings of the Human Factors and Ergonomics Society 41st annual meeting, 1997, pp. 1113-1117.
- Smith-Jentsch, K.A., Zeisig, R.L., Acton, B., & McPherson, J.A. (1998). Team Dimensional Training: A strategy for guided team self correction. In J.A. Cannon-Bowers, & E. Salas (Eds.), Making decisions under stress: implications for individual and team training (pp. 271-298). Washington DC: APA.
- Swezey, R.W., & Salas, E. (Eds.), Teams: their Training and Performance. Norwood, NJ: Ablex.
- Van Berlo, M.P.W. (1998). Team training for team performance: models of team task performance (Report TM-98-B014). Soesterberg: TNO Human Factors.
- Van Berlo, M.P.W., & Schraagen, J.M.C. (2000a). The development of a generic assessment tool for evaluating command & control exercises (Report TM-00-A010) (Confidential) (in Dutch). Soesterberg: TNO Human Factors.
- Van Berlo, M.P.W., & Schraagen, J.M.C. (2000b). Evaluating command & control during an MDF-exercise using the 'Command & Control Process Measurement Tool' (C2PMT) (Report TM-00-A011) (Confidential) (in Dutch). Soesterberg: TNO Human Factors.
- Van Berlo, M.P.W., & Schraagen, J.M.C. (2000c). A generic tool for evaluating command & control: the 'Command & Control Process Measurement Tool' (C2PMT) (Report TM-00-A012) (Confidential) (in Dutch). Soesterberg: TNO Human Factors.
- Zsombok, C.E., Klein, G., Kyne, M., & Klinger, D.W. (1993). How teams excel: A model of advanced team decision making. Performance Technology-1993. Selected Proceedings of the 31st NSPI Conference (pp. 19-27). Chicago, IL: NSPI.