15th ICCRTS

"The Evolution of C2
"
Towards an understanding of the commander's "coup d'oeil

Topic(s)
Experimentation and analysis

Berndt Brehmer and Jan Kuylenstierna

Point of Contact
Berndt Brehmer
Swedish National Defence College
P.O. Box 27805
SE-115 93 Stockholm Sweden
Telephone
+46 8 553 42 837
E-mail Address
berndt.brehmer@fhs.se

Towards an understanding of the commander's "coup d'oeil

Berndt Brehmer and Jan Kuylenstierna[1]
Swedish National Defence College

ABSTRACT

In his chapter on military genius, Clausewitz writes about a commander's *coup d'oeil*, an ability to make quick and correct decisions on the battle field as something that distinguishes a military genius from the rest of us. Clausewitz had little to say about the basis for this for the ability, however. The present study represents a first attempt to understand the commander's coup d'oeil as an expression of military expertise. It uses a standard experimental paradigm from psychology for studying experts by assessing their ability to recall of meaningful and meaningless military scenarios. The results of our experiments agree with those from 19 other domains of expertise in showing that experts (army majors in our case) recall meaningful scenarios significantly better than meaningless military scenarios, while for novices (second year army officer cadets and college students of political science) there was no significant difference in recall of meaningful and meaningless scenarios. Thus military expertise manifests itself in the same manner as expertise in other fields. Two explanations for the results are discussed. Both hypotheses assume that experts perform better than novices for meaningful scenarios because they are able to alleviate cognitive load. The first hypothesis assumes that they do so by learning to perceive and utilize the goal-directed constraints of military scenarios, the second that they learn to recognize patterns of military activity. It is noted that it is not possible to distinguish between these two hypotheses on the basis of static scenario displays, and that they may not be that different, since the patterns that are perceived must be related to the constraints.

Sensemaking has emerged as a central theme in research on command and control (C2), where it refers to the commander's understanding of the situation and what should be done. Although the term may be new, the phenomenon is not. Clausewitz writes about it in his celebrated chapter on military genius, with the terminology of his times. He uses the term *coup d'oeil*, a term popular in among military theorists during the 19th century to denote

> "…the *idea of a rapid an accurate decision* was … based on an evaluation of time and space, and consequently received a name that refers to visual estimates only. …But soon it was also used of any sound decision taken in the midst of action—such as recognizing the right point of attack, etc. *Coup d'oeil* therefore refers not alone to the physical but, more commonly, to the inward eye. The expression, like the quality itself, has certainly always been more applicable to tactics, but must also have its place in strategy, since here as well quick decisions are often needed. Stripped of metaphor and of the restrictions imposed on it by the phrase, the concept merely refers to the quick recognition

[1] We thank Isabell Andersson and Joacim Rydmark for their assistance in running the experiments and Lieutenant Colonel Sten-Olof Olsson and Major Mats Persson for their help with the scenarios used in the experiments.

> of a truth that the mind would ordinarily miss or would perceive only after long study and reflection." (Clausewitz, 1834/1989, p. 102, italics in original).

Any modern day psychologist or cognitive scientist will feel right at home with this description. It implies that a commander would have what we today would call *expertise* that allows him or her to make rapid decisions. Clausewitz restricts this ability to commanders of genius. An alternative to his view would be that it expresses the expertise that all officers, also those who are not military geniuses, acquire, at least to some extent, as a result of their training and experience. This is our view and the point of departure for this project, the aim of which is to first demonstrate that officers have the expertise assumed, and then to find the basis of this expertise. As a first step, we adapt a standard experimental paradigm used to study expertise. It was first used by Chase and Simon (1973a) to study expertise in chess, and it has then has been used to study expertise in at least 19 different fields (Vicente & Wang, 1998), from chess to medical diagnosis. We first explain the method with reference to studies of expertise in chess where it originated. We then report three experiments, the results of which allow us to compare the performance of officers with those from other fields of expertise. Finally, we discuss two possible interpretations of the results.

The Chase and Simon paradigm for the study of expertise in chess

The Chase and Simon (1973a) paradigm is a paradigm designed to assess expertise by studying memory recall. When applied in the domain of chess, it involves comparison of memory recall of the positions of chessmen under two conditions, one where the positions are the result of actual play (meaningful condition) and one where the chessmen have been positioned randomly (meaningless condition) on the chessboard. The participants are allowed to inspect the position for some time. The board is then removed and the participants are given an empty board and a set of chessmen and asked to reproduce the scenario that they have just seen. When this experiment is run with experts (chess masters or grand masters) and novices, there is a remarkable difference in results. Experts recall the positions that result from actual play much better than novices, but for the meaningless positions, there is no difference. Thus, the chess experts do not have better memory generally, not even for chessmen and their positions, but they do have better memory for actual chess scenarios than novices do. This suggests that experts are able to pick up information in the meaningful positions that novices are unable to pick up. As already mentioned, similar results have been obtained in a wide variety of domains, from chess to medical diagnosis (Vicente & Wang, 1998). But so far, there seems to have been no studies of officers. Yet, the generalisation to officer expertise seems straightforward; chess is, after all, a war game.

Below we present three experiments using the Chase and Simon (1973a) paradigm adapted to the study of the expertise of officers. The three experiments differ with respect to the level of expertise of the participants. In Experiment I the participants were Army Majors from the Higher Staff Officer Program at our college and they constitute our expert condition. In Experiment 2 the participants were second year army officer cadets, and in the third they were college students taking courses in political science, our two novice conditions. In the experiment, the participants were asked to recall positions of military units placed on a map under a condition where the positions were meaningful militarily and a condition where the units had been placed as near randomly as possible (with certain restrictions described below). We expect to have a strong effect of meaningfulness for experts and no effect for novices (or at best a very weak one).

Experiments

Method

All three experiments followed the same design and the same procedures (with some minor variations) were used in all three experiments.

Participants

The first experiment employed 16 Army Majors from the Higher Staff Officers program at the Swedish National Defence College, all male and aged around 41 years. The second experiment employed 16 second year male army officer cadets, aged around 25 years. The third experiment employed 16 college students taking courses in political science. This group included 11 male and 5 female participants and their average age was 22 years. The students were asked about their military experience and it turned out that 3 had done basic military training in the Army. Each participant volunteered to take part and received two cinema vouchers as a reward.

Design

The independent variable in each experiment was the level of meaningfulness of a ground warfare scenario presented on a map. The dependent variables were the participants' interpretation of these scenarios and their ability to reproduce the positions in the scenarios after having studied their scenario for five minutes.

Scenario construction

The scenarios for the two conditions were constructed by a Lieutenant Colonel from the Swedish Land Warfare Centre. In the meaningful scenario nine "blue" units were grouped for attack in a terrain occupied by fourteen "red" units grouped for defence. In accordance with the Chase and Simon paradigm, all physical positions in the meaningless scenario were the same as those in the meaningful scenario, but the units in each position was rearranged so that the resulting scenario had no meaning from a military point of view. The rearrangement was not done through randomization however. Preliminary try-outs with randomization suggested that this method was too blunt to create of scenarios without military meaning. Therefore the units were instead rearranged deliberately with the intention of creating an incomprehensible situation. In the meaningful scenario all blue units were grouped together at the left of the map and all red units were grouped too the right. Remembering this simple pattern could be a potential aid to recall, the rearrangement of the units in the meaningless scenario therefore preserved this general pattern. As a result, all positions on the blue side in the meaningful scenario were still occupied by blue units in the meaningless scenario and all red positions were still occupied by red units, but the units had been changed.

Experimental set-up

The participants were seated at a table. In front of them was a white (1.20 m x 1.07 m) sheet of paper. Hidden underneath this sheet of paper was another 1.20 m x 1.07 m sheet of paper with a 0.74 m x 0.82 m scenario map on it. The scenario map had a legend naming the various units. Under the scenario map there was a map with the same terrain as the scenario map, but without any units or legend. On top of this map was a sheet of transparent plastic.

Procedure

The subjects first signed an informed consent form. They were then given their first task orally. It was explained to them that there was a map with a ground warfare scenario with blue and red units fighting each other underneath the sheet of white paper in front of them. They were also informed that their task was to write down a few lines on an answering sheet what they thought that the blue side and the red side were trying to achieve. Finally they were told that they had five minutes to accomplish this task with the clock starting as soon as the scenario map was uncovered. Then the cover sheet was removed and the clock started. After five minutes the participants were asked to put their answering sheet aside and the scenario map sheet was turned over exposing the “empty map” with its plastic sheet. Only then, the participants were informed that there would now be a second task, viz., to reproduce the positions of the units in the scenario map on the plastic sheet. The reason for not informing the participants that there would be a second task beforehand was to prevent them from entering into a “memorizing mode”. There is evidence such a mode may reduce effects of meaningfulness (Norman, Brooks & Allen, 1989). The participants were informed that time for reproduction was free. The participants were then given a blue and a red marker pen and instructed to reproduce the scenario they had just seen, that is; to put the right unit in the right position. During the reproduction phase, the participants in the political science group had access to a grey version of the unit legend on a separate paper. Unlike the Army majors and the Army officer cadets, the political science students could not be expected to be familiar with army units and they were therefore allowed this extra support. The average time required for reproduction was about five minutes and the whole experimental session normally required about 15 minutes.

Results

The participants in all three conditions agreed on the general interpretation of meaningful scenario, viz., that blue was attacking and red was defending. However, the interpretation were not identical. The experts gave a more elaborate interpretation. For the meaningless condition, there was no agreement among participants in either experiment on the interpretation.

A template with 10 concentric circles was used to evaluate the results. The innermost circle was 1 cm. and, the next 2 cm. the third 3 cm, and so on. For each unit, the template was placed on the map with the correct position for the unit in the center. If the participant had drawn the symbol for the correct unit and placed it in the innermost circle, he, or she, received a score 10 points, if in was in the next innermost circle, the participant received a score of 9, and so on. If the participant had placed the unit outside the 10^{th} circle, or if he or she had drawn the wrong unit, the score was 0. The maximum possible score was thus 230 and the minimum 0. Measurements were made by two persons independently, and the correlation between the scores they generated was .93. Disagreements were resolved through discussion before further analysis. The difference in mean scores between the two conditions in each experiment was subjected to a t-test, and in addition, a Cohen’s d-score was computed as a measure of effect size. The results are given in Table 1.

Table 1. Mean recall performance scores in the meaningful and meaningless condition in the experiments, and t-test results and effect size (Cohen's d)

Experiment	Mean meaningful condition	Mean meaningless condition	t	df	p
1	73.63	42.00	2.18	14	0.047
2	75.88	60.75	0.91	14	0.38
3	26.38	15.25	1.57	14	0.13

Note. The maximum possible score in each condition is 230.

As can be seen from the table, the results of our experiments 1 and 3 replicate those from studies of expertise in other domains. For the experts, there is a significant difference in performance between the meaningful and the meaningless condition, but for the novices there is not.

The results for the second year cadets (Experiment 2) may seem puzzling, but similar results have been found in other studies especially in the medical field (e.g., Hassebrock, Johnson, Bullemer, Fox & Moller, 1993, Patel, Groen & Fredriksen, 1986). The important result here, however, is that although the cadets perform as well as the experts with the meaningful scenario, there is no significant difference between their performance for the meaningful and the meaningless scenarios. We will return to this result below.

Analysis of the scores for different units, using 2 (conditions) by 23 (units) analysis of variance for each experiment yielded some significant differences in scores but they could not be given any meaningful interpretation.

Discussion

The results of the present experiments show the standard outcome for our experts: better recall of a meaningful military situation than for a meaningless one. For the novice groups, there is no significant difference between the two conditions. The performance for the novice group comprising political science students is much below that of the experts.

For Experiment 2, that with the second year cadets, the level of performance is as high as that for the experts in the meaningful scenarios and considerably above that for the experts in the random scenarios. As noted above, similar results have been obtained in studies of medical expertise, where medial students are sometimes found to outperform experts in their recall of the information (Hassebrock, et al. 1993, Patel, et al., 1986). The explanation for this result seems to be that the experts and the medical students remember different things, presumably related to what they extracted from the cases presented. The students remember surface characteristic (descriptions of symptoms) while the experts remember the inferences that they have drawn from the case presentations. That is, the experts seem to remember the meaning, but the students simply remember the cases. This suggests that just testing for recall of surface characteristics (such a positions and units) may not be sufficient to distinguish between levels of expertise. It may also be necessary to test for the recall of more directly for meaning related aspects, such as inferences. The medical studies do not involve any comparisons between novices without any medical training and medical students comparable to our comparison between students of political science and second year officer cadets, so we can only speculate the explanation for the difference in results between Experiments 2 and 3. A likely explanation would be familiarity with the materials presented, officer cadets being much more familiar with military material of the kind being presented here than students of political science (for a review of research on familiarity and recall, see Younelinas, 2002). This would suggest that we need different explanations for the level of recall and for the difference in recall between meaningful and random scenarios. It is, however, the latter of these two problems that is the focus for the present paper, and we now turn to the results relating to that problem.

As already noted, the difference in recall for meaningful and meaningless scenarios is a standard result from expertise research in many different domains. The literature on expertise offers two different explanations for these results. The first, we call the *constraints hypothesis*, the second the *chunking*, or *pattern recognition hypothesis*. According to the former hypothesis, which was proposed by Vicente and Wang (1998), experts have learned to pick up information about the goal-related constraints in the domain and use that information to cut down on the number of possibilities when reproducing the situation. In contrast, the chunking, or pattern recognition hypothesis, first proposed by Chase and Simon (1997a, b) suggests that experts have learned to categorise information in terms of recurring patterns in their domain. They then remember the patterns and use them in recall. Researchers in the field of C2 will recognize this as a close relative to, and indeed the foundation for, Klein's , popular recognition primed decision model (e.g., Klein, 1999).

Although the two hypotheses sound very different, it is in fact, difficult to distinguish between them empirically for they make the same predictions and they do so on a very similar basis. Thus, both hypotheses assume that the advantage enjoyed by the experts has its basis in their having to store less information than novices do to perform well at recall. According to the constraints hypothesis, experts have to store less information because they know the constraints, i.e., in this case, what can and cannot be done with military units to achieve a given goal (in this case for the blue side to attack the red side effectively) and this cuts down on the number of possibilities for actually placing the units in the recall conditions. To put matters differently, they make fewer errors than novices because, basically, they reproduce what they know must have been the case, given their understanding of the goals and the constraints. According to the pattern recognition hypothesis, the experts have learned to recognize patterns of positions, and when inspecting a scenario, they recognize and store the pattern as a unit (*chunk*, hence the name "the chunking hypothesis") rather than the positions of the individual units. They then reproduce the chunk from memory without having to recall the positions of each individual unit as such. Both hypotheses thus predict an advantage for experts, based on their having to store less information to reproduce what they have seen, and both hypotheses explain the experts' ability to do this in terms of what they have acquired on the basis of experience. They differ in what is assumed to have been learned, however: constraints or patterns. The explanation for the inferior performance in the meaningless condition is, in the case of the constraints hypothesis that knowing the constraints will not help when the positions of units to be remembered violate the constraints. In this case, knowing the constraints will not help at reproduction and it becomes necessary to remember the individual units, just as a novice would have to do. In addition, for a meaningless condition, where not even an expert can detect the goals, it will be difficult to decide what the constraints are, since it is the goal that introduces the constraints. According to the chunking, or pattern matching hypothesis recall will be worse for meaningless material because the expert cannot rely on the patterns that he or she has learned to reproduce the positions of the units. This is because the scenario will not match any of the patterns that the expert would have stored. For novices, who have neither earned the constraints nor the patterns, both hypotheses, predict that meaningful material will not give any advantage at recall. Knowing neither the constraints and nor any patterns, novices have to remember the positions of the individual units for both meaningful and meaningless material, making both conditions equally difficult. The novices will not be helped by the fact that they have understood what is going on in the scenarios, for even though they know the goals, they do not know the goal-related constraints of the domain. As sown by the results of Experiment 2, familiarity with the domain seems to help recall, but it does not give any advantage when having to deal with meaningful scenarios, only for meaningless ones, suggesting that the cadets recalled the surface information rather than the meaning of the scenarios in the form of constraints or patterns.

It is interesting to note, that the two hypotheses are related to Clausewitz's distinction between understanding what is to be done by visual inspection on the one hand and looking inwards on the other. The constraints hypothesis assumes that the expert is able to *see* different things than a novice does, viz., the constraints. The chunking hypothesis, in contrast, assumes that the experts learn and store patterns, and by "looking inwards" at these stored patterns when asked to recall, they are able to reproduce the positions of the units. Chase and Simon emphasize this view the title one of their papers describing the chunking theory: "The mind's eye in chess" (Chase & Simon, 1973b).

For static displays, such as a situation reproduced on a map, it will obviously not be possible to distinguish between the two hypotheses. On second thought, however, the hypotheses may not really be all that different. The patterns that an expert remembers, if that is indeed what he or she does, must, of course, have some basis in reality to be of any use to them, i.e., they must have a basis in patterns in military scenarios, in the present case. That basis will, of course, be generated by the goal related constraints. If there were no constraints, there simply would no be any patterns to learn, for military scenarios could take any form. That is, the reason why there are patterns in the first place is that there are constraints: If the red or the blue side are engaged in some goal-related military activity, the units on the map just cannot occur in any position, just as chess men in a real chess game cannot occur in just any position.

It will thus be very difficult to distinguish between the two hypotheses from static displays that show only a frozen scenario on a map or chess board. However, using dynamic displays, it may be possible to distinguish between them, for this would offer the possibility to present not only the final positions, but also how the scenario evolved. Thus, it would be possible to present final positions that evolved either naturally or by violating constraints. Here the constraints hypothesis predicts an advantage in recall for the scenarios that have developed naturally compared to scenarios that have evolved by violating constraints. A chunking hypothesis would not make this prediction, since it only takes the pattern of positions of the units into account. Experiments testing this hypothesis are currently being planned in our laboratory.

On a more theoretical level, a chunking, or pattern recognition hypothesis does not seem to be a very likely explanation for military expertise. Learning patterns is a slow and laborious process, requiring years of experience. For example, expertise in chess is based on ten or more years of playing many games daily (Chase & Simon, 19731, b). Very few officers will have the kind of experience where they have seen a given scenario very many times, or indeed very many scenarios at all. The basis for learning patterns may well be too limited, even at the lower tactical level, and certainly at the higher levels of military activity. The constraints hypothesis, in contrast, suggests a more solid base for learning; the constraints are quite general. They are present in every military scenario, thus providing frequent opportunities for learning.

It seems more likely, therefore, that military expertise is a matter of learning constraints than patterns. This also casts some doubt on the Klein's Recognition Primed Decision Model as a viable theory of military expertise, especially at higher levels (see also Thunholm, in press).

Even if this first foray into the area of military expertise has not produced conclusive evidence concerning the basis of the expertise of officers, it has nevertheless given clear evidence that such expertise exists and that it has the same character as expertise in other fields. It is interesting to note that this expertise, at least for these simple scenarios, does not seem to help experts interpret the situation any better, or differently, than novices do. Both experts and novices agree on the same general interpretation, just as Clausewitz said that they would, given enough time. Given shorter time for analysis, however, experts may well be superior, and arrive at more similar interpretations, than novices, giving evidence of an officer's ability grasp the situations by a *coup d'oeil*. A natural next step, therefore, would be to vary the time for analyzing the situation. Experiments of this kind are also being planned in our laboratory.

References

Chase, W. G. & Simon, H. A. (1973a). Perception in chess. *Cognitive Psychology, 4*, 55-81.

Chase, W. G. & Simon, H. A. (1973b). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. (pp. 215-281). New York: Academic Press.

Clausewitz, C. von (1834/1989) *On war*. Edited and translated by M. E. Howard & P. Paret. Princeton: Princeton University Press.

Haselbrock, F., Johnson, P. E., Bullemer, P., Fox, P. W. & Moller, J. H. (1993). When less is more: Representation and selective memory in expert problem solving. *American Journal of Psychology, 106*, 155-189.

Klein, G. A. (1999). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Norman, G. R., Books, L. R. & Allen, S. W. (1989). Recall by expert medical practitioners as a record of processing attention. *Journal of Experimental Psychology: learning, Memory and Attention, 15,* 1166-1174.

Patel, V.L., Groen, G.J. & Fredriksen, C. H. (1986). Differences between medical students and dosctors in memory for clinical cases. *Medical Education, 20,* 3-9.

Thunholm, P. (In press). Managing complex military planning: Processes and techniques used by experiences and less experienced commanders. In K. Mosier & U. Fisher (Eds.), *Informed by knowledge: Expert performance in complex*. London: Taylor & Francis.

Vicente, K. J. & Wang, J. H. (1998). An ecological theory of expertise effects in memory recall. *Psychological Review*, *105*, 33-57.

Younelinas. A.P. (2002). The nature of recollection and familiarity: A review of 30 yaars of research. *Journal of Memory and Language, 46*, 441.517..