Title
The Précis of Project Nemo, Phase 2: Levels of Expertise

Authors
Kirschenbaum, Susan S.
Gray, Wayne D.

Peer reviewed

The Précis of Project Nemo, Phase 2: Levels of Expertise

Susan S. Kirschenbaum (kirschenbaumss@csd.npt.nuwc.navy.mil)
Naval Undersea Warfare Center Division
Code 2214, Building 1171/1
Newport, RI 02841 USA

Wayne D. Gray (gray@gmu.edu)
Human Factors & Applied Cognition
George Mason University; m/s 3f5
Fairfax, VA 22030 USA

Abstract

Project Nemo examines the cognitive processes and representational structures used by submarine Commanders while attempting to locate an enemy submarine hiding in deep water. In phase 2 we collected performance and protocol data from junior, mid-career, and senior submarine officers. The data support the conclusions from phase 1 (Gray, Kirschenbaum, & Ehret, 1997) that most AO actions can be characterized as a sequence of small, steps in a shallow goal hierarchy (rather than as following a detailed master plan). The nature of these successive choices vary as a function of the officer's expertise. The results are congruent with an interpretation in which the process of schema instantiation provides the control of cognition.

Introduction

In phase 1 of Project Nemo (Gray et al., 1997) we analyzed six hours of verbal and action protocols from expert submarine Approach Officers (AOs) as they detected and localized (i.e., determined the course, speed, and range) a hostile submarine hiding in deep water.

The results of phase 1 support a description of the cognitive control structure that orchestrates the AOs' behavior as *schema-directed problem solving with shallow and adaptive subgoaling* (Ehret, Gray, & Kirschenbaum, in press). The schema is the task-relevant knowledge accumulated in over 20 years of experience as a submariner (half of it at sea). It is a set of declarative as well as procedural knowledge structures. An implication of shallow subgoaling is that the knowledge available to AOs is so rich that steps to supplement this knowledge can be shallow.

A second implication is that the AO solves a series of problems, one every 30 to 300 s. The problem is always the same; namely, "what is the state of the world – NOW." The AO is trying to find a quiet target hiding in a noisy environment while remaining undetected himself. The protocol analysis revealed that he takes a series of short steps that either (a) assess the noise from the environment or signal from the target – NOW, or (b) attempt to reduce the noise or increase the signal from the target by maneuvering ownship. As shown in Figure 1, these short steps result in shallow subgoaling. When a

subgoal pops, the schema is reassessed. The result of this reassessment directs the next step (i.e., selects the next subgoal). This step is accomplished, it returns information to the schema, the schema is reassessed, and so on.
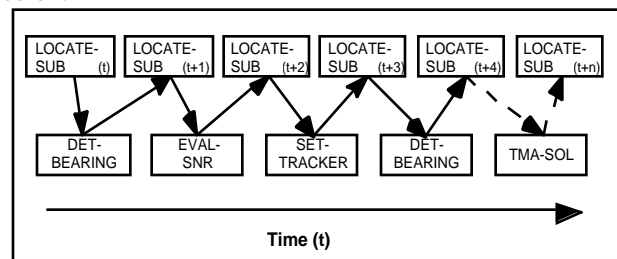

Figure 1: Schema-directed problem solving with shallow and adaptive subgoaling

The process of subgoaling is adaptive in two senses. First, the subgoal that is chosen next reflects the current reassessment of the schema. Second, this choice is sensitive to both the long-term importance of the subgoal as well as its recent history of success or failure. Regardless of a goal's long-term importance, AOs will not continue to attempt a goal if successive tries fail. Instead, they will choose another goal and return to the more important goal later.

The dynamic aspect of the AO's task plays an important role in this view of schema-directed problem solving. First, the state of the AO's world is continually changing – both ownship and target are moving at a given depth, direction, and speed. For ownship the value of these attributes can be changed, but the problem cannot be stopped. Consequently, time is an important part of the picture. Second, subgoals are not accomplished once and then discarded. In the AO's world, subgoals bring in certain types of information or accomplish certain changes to ownship. As the world changes, any given subgoal may be revisited not only to acquire the current value, but also to acquire information about the rate and direction of change (e.g., DET-BEARING in Figure 1).

Phase 1 ran 10 senior officers on a high fidelity simulation located at the Naval Undersea Warfare Center

in Newport, RI. For phase 2, we built the Ned[1] scaled world in Macintosh Common Lisp to run on a portable computer. (A description of the simulation can be found in Ehret et al., in press.) This portability enabled us to take Ned to U. S. Navy submarine bases in Bangor, WA and Pearl Harbor, HI. Consequently, we were able to collect data from 36 active-duty submarine officers.

In this paper we present a brief overview of the phase 1 empirical data. (More details can be found in Gray et al., 1997; and Kirschenbaum, Gray, & Ehret, 1997.) Our focus is on the data collected using the Ned scaled world, its similarities to the phase 1 data, and the variations among levels of expertise.

The Submariner's Task and Tools

The job of the Approach Officer is to respond to hostile targets. He[2] heads the team that must detect, track, classify, localize, and if ordered, attack the target. He performs this task with the support of many special-purpose systems run by skilled operators, but is ultimately responsible for the success of the encounter.

Two features of the problem make it an especially challenging one. First, this is a dynamic problem. Both ownship and the contacts are moving, and, perhaps, changing course, etc. during the encounter. Second, there are only sparse and highly uncertain data about the contacts. The AO's expertise lies in using his knowledge of the relationships among the cues to guide information gathering over the course of the scenario and instantiate a generic "contact" schema for each contact.

Special tools are used for controlling own ship, listening to the contact, and localizing. As sound transmission is distorted, reflected and bent in the ocean by temperature, salinity, pressure, detecting, tracking, and locating the source of a passive sonar contact is highly very difficult and impacted by uncertainty. Because passive sonar only provides bearing (direction) data, target-motion-analysis (TMA) tools for localizing the targets employ statistical methods to estimate the target's course, speed, and range. As this is a mathematically under-constrained problem, submariners call this process "finding a solution."

Review of Phase 1

Method

All of the participants in phase 1 were highly experienced submarine officers who had served as Commanding Officers (COs) or Executive Officers (XOs) aboard U. S. Navy Submarines. The officers were presented with scenarios that required localizing an enemy submarine hiding in deep water. The scenarios were presented on the CSEAL (Combat Systems Engineering and Analysis Laboratory) high fidelity simulation. CSEAL is a submarine command-center-in-a-box. It has generic versions of all of the essential submarine tools. As CSEAL is a developer's tool, it was run by an operator. The AOs requested information and ordered actions to be carried out by the operator. Videotapes and verbal protocols were the primary data. These were supplemented by computer-logged action protocols.

In both phases we investigated the situation assessment part of the AO's mission. Situation assessment begins with detection of the contact and ends when the AO is sufficiently confident of the solution to declare that he is ready to move to the engagement phase. Each scenario began with a status report such as an AO might receive when first taking his turn on watch. The status report provided scenario specific information including ownship course, speed, and depth as well as information on any contacts. All scenarios began with a single contact, classified as a merchant.

Review of Phase 1: Results

During phase 1 we developed an encoding scheme (Gray & Kirschenbaum, in press) that included nine operators and a three-level goal structure (for detailed information, see Kirschenbaum et al., 1997). Most of the AOs' time and effort was spent in service of two top-level goals: LOCATE-MERCHANT (LOC-MERC) and LOCATE-SUBMARINE (LOC-SUB). Given that localizing the sub is clearly the higher priority, we were puzzled to find that the two goals were used with approximately equal frequency (see the left side of Figure 2). However, as the middle of Figure 2 shows, this equal frequency of use masked a large difference in the number of subgoals per LOC-MERC versus LOC-SUB.

More interesting, this disparity in number of subgoals per goal was not reflected in the number of operators per subgoal. As shown by the right-side of Figure 2, the mean number of operators per subgoal was constant. The same number of operators were used in a subgoal regardless of whether its supergoal was LOC-MERC or LOC-SUB.

Along with other analyses that we conducted, this analysis suggested that the basic unit of action was the subgoal. Formalized plans or established methods with fixed number of steps, exist at the subgoal level. At the level of LOC-MERC or LOC-SUB, each subgoal returns a discrete piece of knowledge that is added to the schema. The schema is reevaluated to determine what piece of knowledge to select next. When there is little new information to be gained by continuing working on the current goal, the goal is popped and a new top-level goal is pushed.

The question pursued below is whether the phase 2 data support the phase 1 interpretation of expert performance and in which ways intermediate and novice behavior conforms or differs from the experts.

---

[1] Ned Land was an able seaman and trusted assistant to Prof. Aronnax aboard the Nautilus.

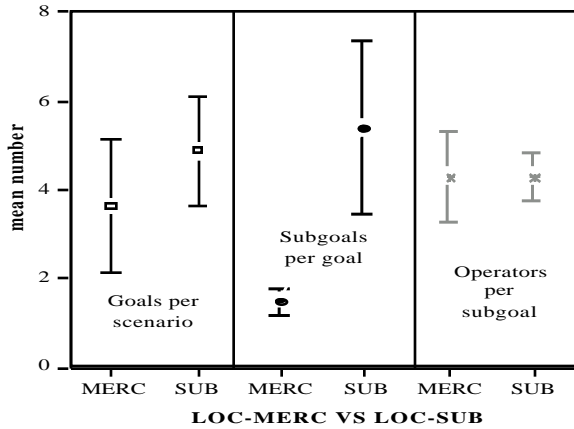[2] In the current US Navy all submariners are men.

Figure 2: Phase 1. Data for the two main top-level goals, LOCALIZE-MERC and LOCALIZE-SUB. Left -- mean number of level-1 goals per AO-Trial. Middle -- mean subgoals per goal. Right – mean number of operators per subgoal. [Error bars show the 95-percent confidence intervals for the standard error of the mean (SEM).]

The Ned Experiment: Replication and Expansion

Table 1: Demographic data on participants.

Means for	CO/XO	DH	JO
n	15	11	10
Years at sea	8.7	6.4	3.2
Years in Navy	17.8	13.4	7.3
Age	38.9	34.4	28.3

Participants

Three groups of current submarine officers participated in the study: Junior Officers (JOs), Department Heads (DHs), and Commanding Officers or Executive Officers (CO/XOs). The average number of years spent at sea, years in the navy, and ages can be found in Table 1. The expert participants in this study had slightly less experience than those in phase 1. This was most likely because, unlike the phase 1 AOs, all were active duty and none were post-command. (The phase 1 COs had a mean of 10.8 years at sea and 20.3 years in the Navy.)

Ned Simulation

The Ned simulation was designed to overcome problems encountered in collecting and encoding data from CSEAL. (These problems and their solution are detailed in Ehret et al., in press.) For the current discussion, the two most relevant improvements in Ned over CSEAL were the elimination of redundant information and the control that Ned provided over access to information.

With minor exceptions, Ned's displays were specialized so that each type of information was available from one display only. For example, when an AO went to the display for the broadband spherical sonar sensor, we could be sure that he wanted one of the 10 types of information

that was available only from that display. Once an AO selected a display, all information fields for the display were covered by black boxes (as in the bottom display of Figure 3). Clicking on the field removed the black box and revealed the data until the mouse was moved from the field. (Ned consists of 10 specialized displays.)

Ned captured all AO interactions, including display navigation and viewing information (enter and exit times and information content). It also recorded truth every 20 seconds. In addition, the AOs were encouraged to think aloud and all sessions were video recorded.
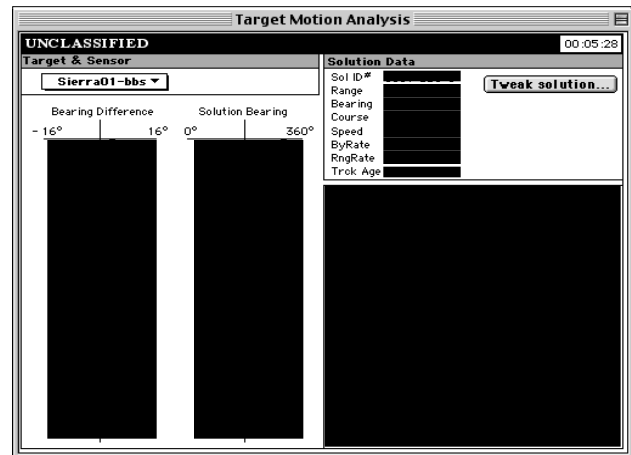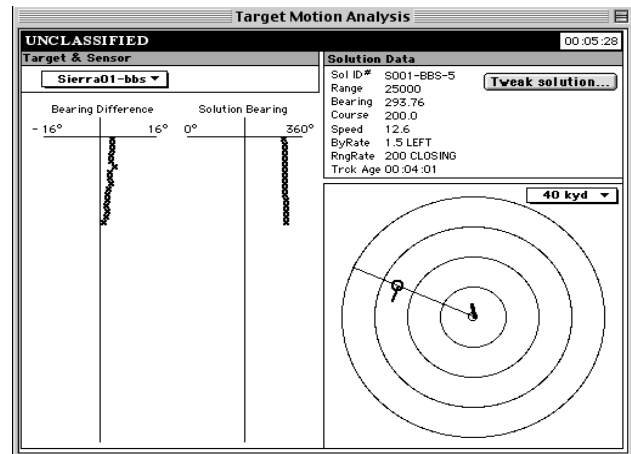

Figure 3: The Ned Target Motion Analysis screen without (above) and with (below) black boxes covering data fields.

Scenarios

Four scenarios were used. Two were identical to those used in phase 1 and two were slightly modified versions of the phase 1 scenarios. At the beginning of each scenario the AO had ownship position (course, speed, and depth) and confirmed contact and bearing (direction from ownship) for a merchant. He also had intelligence that a "hostile" submarine was in the region.

Procedures

Each session began with training on Ned and training in talking aloud while problem solving. Each AO solved two

scenarios. Sessions lasted approximately 2 hours.

Protocol Encodings

Five CO/XO's were unable to complete two scenarios due to interruptions for other responsibilities. From the remaining 31 AOs, data from six AOs at each experience level (18 in all) were selected for detailed encoding. In each case, the second scenario was encoded. Protocols were selected on the basis of completeness, the lack of technical glitches, and the clarity of the recorded protocols.

Semi-Automatic Protocol Encoding

Each click of the mouse on a menu item or a black box was saved to a log file. This enabled us to write code that encoded each action protocol and segmented groups of related actions. For example, if the AO clicked on the black box covering the range information in Figure 3 (see upper right part of the display), he was credited with two operators one for querying and one for receiving range information from the target motion analysis system.

Operators

Operators are the lowest level encoding and represent the AO's direct interaction with the Ned simulation. Unlike phase 1, the majority of operators (approximately 99%) were encoded automatically from the computer outputted action protocols. All encodings were confirmed and/or modified by comparison with the video-taped verbal protocols. Across the three groups a total of 9,073 operators were encoded as belonging to one of nine operator categories.

Table 2: Example of goal and operator encodings.

GOAL	OP	INFO-SOURCE	SHIP	ATTR	VALUE	DUR
DETERMINE-BEARING						
	QUERY	NBT-BY-FIELD	SUB	BY		
	RECEIVE	NBT-BY-FIELD	SUB	BY	316 or 244	1.15
POP						.

An example of the result of the automatic encoding of operators is provided in Table 2. Prior to this point in the scenario, the AO has called up the narrowband towed display (NBT). Here he queries the bearing (BY) information for the SUB by clicking on the black box that covers the field. The black box disappears, enabling the AO to receive the information that the narrowband towed sensor gives the ambiguous information that the hostile submarine's bearing from ownship is either 316 or 244 degrees. The bearing information is uncovered for 1.15 sec before the AO moves the cursor out of the bearing field.

Details of the operator types and categories used in phase 1 are available from Kirschenbaum, et al. (1997).. The phase 2 operator types and categories differed minimally from those used in phase 1; however, their similarity and differences from the phase 1 operators are beyond the scope of the current report.

Goals and Subgoals[3]

The AO's mission, as given in the instructions, is to destroy the hostile submarine. Therefore his primary goal is to detect and localize the sub. However, these are not his only goals. He must also keep track of the merchant, avoid collision, evaluate the environment, and keep track of ownship.

Under Ned we have a precise record of the AO's information access. This record, linked by time to the verbal protocol, permitted a more detailed encoding of goals than was possible for phase 1. Hence, the goal and subgoals used in phase 2 differed from those discussed in Kirschenbaum et. al, (1997). However, the discussion of these differences will have to await a fuller report.

Of the 18 scenarios studied in phase 2, six were used to train the three encoders. These are referred to as "consensus encodings." The operators for each of the remaining 12 scenarios were encoded into goals by two independent encoders. Cohen's Kappa for interrater reliability averaged 0.84 and ranged from a low of 0.54 to a high of 0.96. All correlations are significant ($p < .001$). The discrepancies between encodings were resolved by the third encoder.

Table 3: Typical goal, subgoal, operator sequence. (This is a truncated sequence and is for illustrative purposes only).

L-1	L-2	L-3	OPERATOR
LOC-SUB			
	EVALUATE-TRACE		
			Query
			Receive
			Derive
			DisplayNav
	TMA-SOLUTION		
		EVAL-SOLUTION-BEARING	
			Query
			Receive
			Derive
		TWEAK-SOLUTION-BEARING	
			Tweak
		EVALUATE-SOLUTION-RANGE	
			Query
			Receive
			Derive
		TWEAK-SOLUTION-RANGE	
			Tweak
		DETERMINE-SOLUTION-QUALITY	
			Query
			Receive

Goals and Schema

The schema instantiation process that we hypothesize controls cognition during situation assessment proceeds by pushing and popping a series of largely independent subgoals (see Figure 1). When a goal pops, information is returned to the schema being instantiated. The amended instantiation selects the next goal to push. For example, a typical sequence of goals, subgoals, and operators might read like that in Table 3

---

[3] For ease of exposition, level-1 goals will be referred to simply as *goals*; level-2 and level-3 subgoals collectively as *subgoals*.

In this sequence the AO first evaluates the sonar trace. This may return information to his schema regarding the target's bearing and bearing rate. He then switches to the display shown in Figure 3 and examines the TMA solution, alternately evaluating and tweaking the values of different parameters. As the values are returned to his schema he can compare them with his knowledge of how targets and the TMA algorithms work to derive better values to test. At the end of the sequence, he determines the solution quality by examining how closely the dots in the bottom-left section of the TMA screen (Figure 3a) stack on the central line.

Summary

Ned records AO actions with greater specificity and accuracy than permitted by CSEAL. Consequently, we revised the goal types and categories to take advantage of this greater detail. However, the phase 2 revisions are elaborations on the goal categories and types used in phase 1. Thus, the phase 1 goal structure, with minor modifications, can support the detailed analysis of Ned data.

Data Analysis and Results

The 9,073 operators collected in phase 2 can be aggregated and examined for many different purposes. In the current paper we limit our purposes to three. First we generally compare the goal and subgoal structure used in phase 2 with that of phase 1. For our second and third purpose, we limit ourselves to the three measures used in Figure 2: LOC-MERC and LOC-SUB goals per scenario, number of subgoals per LOC-MERC and LOC-SUB, and number of operators per subgoal. We begin by using these measures to compare the experts in phase 2 (i.e., the CO/XO's) to those in phase 1. We then use these same measures to look across levels of expertise for phase 2.
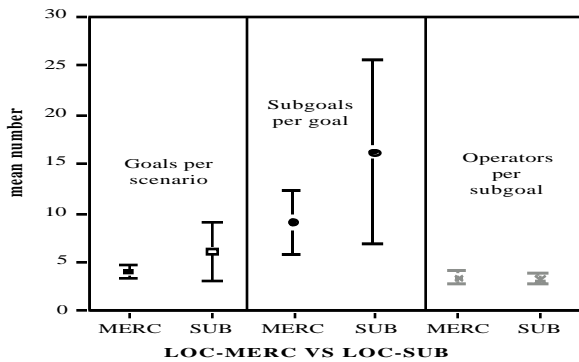

Figure 4: Phase 2 : Data for the two main top-level goals, LOCALIZE-MERC and LOCALIZE-SUB. Left -- mean number of level-1 goals per AO-Trial. Middle -- mean subgoals per goal. Right – mean number of operators per subgoal. (Error bars show the 95% CI for SEM.) Compare with Figure 2.

Comparison with CSEAL Data

The Ned data replicated all of the major findings reported in phase 1. In phase 1 we reported a relatively flat goal

hierarchy of 2-3 levels. This is confirmed by the more precise Ned data. Level-3 goals were confined to three level-2 goals, and the large majority 62.1% of all level 3 goals, were subgoals of a single, high-frequency level-2 goal, TMA-SOLUTION.

Secondly, in phase 1 we were able to encode the protocols using only 9 operators. Nine operators worked well for phase 2. The only notable difference in operator sets was exchanging the N/A category from the verbal protocol encodings of phase 1 for a display-manipulation category (i.e., clicking on menu item or black-box) in phase 2. Also, as in phase 1, we found relatively few operators per goal with a mean of 6.0 operators per Level 2 subgoal and 3.4 operators per Level 3 subgoal (though this varied by subgoal).

CO/XO: Phase 1 versus Phase 2 Comparisons of Expert Level Performers

Comparing the three frames of Figure 2 with those of Figure 4 yields a qualitatively similar picture. In both phases, although the differences in numbers are small, the CO/XOs push more LOC-SUB than LOC-MERC goals. However, these small differences at the goal level are countered by large differences at the subgoals level (middle frame of Figure 4). As in phase 1, for phase 2 the number of operators per terminal subgoal (right frame of Figure 4) does not differ as a function of the top-level goal.

These comparisons are consistent with our phase 1 conclusions that the subgoal level captures a basic unit of AO expertise. The goal level, LOC-MERC and LOC-SUB, divides the world into episodes. An episode requires a varying number of subgoals. The exact number depends on features of the current scenario. Merchants are noisy and easy to localize. Hence, most LOC-MERC episodes occur between attempts to detect the submarine and most end with the AO obtaining a good solution on the merchant.
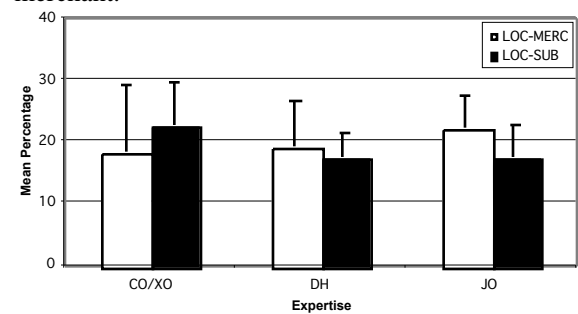

Figure 5: Mean frequency of LOC-MERC and LOC-SUB goals as a percentage of total Level-1 goal usage for three experience levels.

In contrast, enemy submarines are quiet and trying to avoid detection. Hence they are difficult to localize. Most LOC-SUB episodes end after the AO concludes that the current data are not very good and will not get better unless he can take some action to reduce noise or to collect data that will disambiguate data already collected. This decision to halt the current attempt to localize the submarine is never cut-and-dried.

These characterizations of the differences between LOC-MERC and LOC-SUB provide an explanation for the large differences in variance (see the error bars in the middle frame of Figure 2 and Figure 4) in number of subgoals per level 1 goal. For LOC-MERC, localizing is routine. In contrast, LOC-SUB requires flexibility to determine either that the current data are inadequate to enable the target to be localized or that the current best solution is such-and-such.
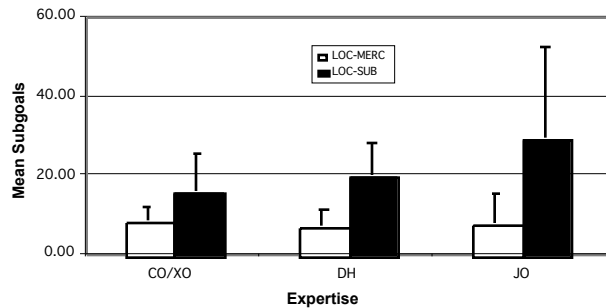

Figure 6: Mean total time spent in LOC-MERC and LOC-SUB goals for the three levels of expertise.

Expertise Effects

All expertise groups pushed `LOC-MERC` and `LOC-SUB` goals with approximately the same frequency (see Figure 5). For all groups, within-group variability overshadows the apparent difference between the goal frequencies. Despite the approximately equal number of LOC-MERC and LOC-SUB goals, across expertise levels there were large differences in the amount of time spent trying to localize the merchant as opposed to the submarine (see Figure 6). The inequality in total time spent pursuing the two goals increases with inexperience.
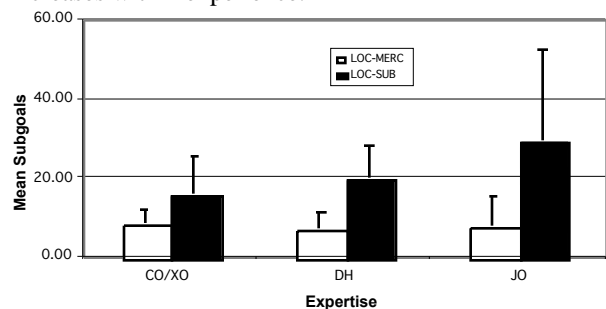

Figure 7: Mean number of subgoals for LOC-MERC and LOC-SUB goals.

As suggested by Figure 7, this difference in time as a function of expertise can be largely accounted for by differences in the number of subgoals. The Junior Officers use almost twice as many subgoals as the most experienced officers. Analyses not reported show that the number of operators per subgoal does not vary with expertise.

Summary and Conclusions

The similarity between the CO/XO's in the two phases of Project Nemo support our characterization of performance as schema-directed problem solving with shallow and adaptive subgoaling. The top-level goals, LOC-MERC and LOC-SUB, do not involve a fixed number of steps; rather, progress on a goal continues until a reevaluation of the schema determines that further effort would be wasted. What is fixed are the number of steps (operators) required for the terminal subgoals.

The phase 2 differences in expertise enrich our hypotheses. The most junior officers use the same building blocks as the most senior officers; that is, the same terminal subgoals are used with the same number of operators per subgoal. In contrast to the typical study of expertise, our "novices" were experienced (see Table 1). Very few officers switch branches of the Navy. Hence, our novices had spent 7.3 years in the Navy with 3.2 years at sea. All of their sea time was spent in submarines.

Where our novices differ from our experts is in their facility at schema-directed problem-solving. Simply put, the less experienced officers pursue bad data longer than the more experienced ones. The experienced ones know when it is time to give up on the current data set and do something to obtain more or better data.

References

Ehret, B. D., Gray, W. D., & Kirschenbaum, S. S. (in press). Contending with complexity: Developing and using a scaled world in applied cognitive research. *Human Factors*.

Gray, W. D., & Kirschenbaum, S. S. (in press). Analyzing a novel expertise: An unmarked road. In J. M. C. Schraagen, S. F. Chipman, & V. L. Shalin (Eds.), *Cognitive task analysis*. Mahwah, NJ: Erlbaum.

Gray, W. D., Kirschenbaum, S. S., & Ehret, B. D. (1997). The précis of Project Nemo, phase 1: Subgoaling and subschemas for submariners. Proceedings of the *Nineteenth Annual Conference of the Cognitive Science Society* (pp. 283-288). Hillsdale, NJ: Erlbaum.

Kirschenbaum, S. S., Gray, W. D., & Ehret, B. D. (1997). *Subgoaling and subschemas for submariners: Cognitive models of situation assessment* (Technical Report 10,764-1). Newport, RI: NUWC-NPT.