Navy Personnel Research and Development Center
San Diego, California 92152-7250 TR-98-5 July 1998

Evaluation of a Software User Coach for Manpower Planning

Jean MacMillian
Barbara Freeman
B. Charles Tatum
Gary A. Ropp

NPRDC-TR-98-5

July 1998

Evaluation of a Software User Coach for Manpower Planning

Jean MacMillian
Barbara Freeman
BBN Corporation

B. Charles Tatum
Gary A. Ropp
Navy Personnel Research and Development Center

Reviewed and approved by
Dennis Schurmeier

Released by
W. M. Keeney
Commander, U.S. Navy
Commanding Officer

and

Murray W. Rowe
Technical Director

Navy Personnel Research and Development Center
53335 Ryne Road
San Diego, California 92152-7250

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATE COVERED
	July 1998	October 1994-September 1996

4. TITLE AND SUBTITLE
Evaluation of a Software User Coach for Manpower Planning

5. FUNDING NUMBERS
Program Element: 0602233N
Work Unit: RM33M20

6. AUTHOR(S)
Jean MacMillian, Barbara Freeman, B. Charles Tatum, Gary A. Ropp

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Navy Personnel Research and Development Center
53335 Ryne Road
San Diego, California 92152-7250

8. PERFORMING ORGANIZATION AGENCY REPORT NUMBER
NPRDC-TR-98-5

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

10. SPONSORING/MONITORING

11. SUPPLEMENTARY NOTES
Functional Area:
Product Line:
Effort:

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12B. DISTRIBUTION CODE
A

13. ABSTRACT *(Maximum 200 words)*

This research project explored the feasibility of improving user performance on manpower planning tasks by employing a "user coach." A user coach is a software aid, often built directly into software applications, that assists the user at critical stages. The user coach developed for this project closely resembles the "wizards" that are commonly used in commercial software today. The user coach was applied to a software tool known as SKIPPER, a manpower modeling tool that is employed by the Enlisted Community Managers (ECMs) at the Bureau of Navy Personnel (BUPERS). This report describes the user coach and documents a formal evaluation of its effectiveness. The evaluation compared the use and understanding of SKIPPER (1) with the coach, (2) with the coach and a visual metaphor (a hydraulic system representation of the enlisted personnel system), and (3) without the coach or the metaphor. All users (experienced and inexperienced) were able to complete their manpower plans significantly faster using the coach than using SKIPPER unaided. Neither performance on the manpower planning task, nor the user's understanding of the task, was influenced by providing a visual picture of the process. It appears that users found the procedural "what to do" aids useful, but aids designed to help the user understand the purpose and dynamics "why do it" of the task were not effective.

14. SUBJECT TERMS
User coach, associate system, visual metaphor, mental models, display design, user interface, human-computer systems

15. NUMBER OF PAGES
50

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

Foreword

This report is part of an exploratory development project entitled "Information Delivery System Design for Personnel Force Management" (Program Element 0602233N, Project RM33M20). The conceptual framework and research plan was developed by researchers at the Navy Research and Development Center. Data collection and analysis was done by BBN Corporation, Systems and Technologies Division (Contract No. N00244-95-D-0281). The report describes the testing and evaluation of a software "user coach" for a manpower planning model known as SKIPPER, and shows the feasibility of adapting an intelligent interface to improve the effectiveness of manpower modeling tools. The work reported here is also reflected in another report (MacMillian, Getty, Tatum, & Ropp, 1998) that explores a technique for comparing mental models with visual metaphors.

W. M. KEENEY
Commander, U. S. Navy
Commanding Officer

MURRAY W. ROWE
Technical Director

Summary

Problem

Inexperienced or first-time users of complex military manpower planning models often have a difficult time learning and using the available software tools. Advances in user interface design have made it possible to improve the effectiveness of complex software tools. Can these design features help improve manpower modeling tools users' performance?

Objective

This study explores the feasibility of improving user performance on a manpower planning task by employing a "user coach." A user coach is a software aid, often built directly into software applications, that assists the user at critical stages. The user coach developed for this project closely resembles the "wizards" that are commonly used in commercial software today. The user coach was applied to a software tool known as SKIPPER. SKIPPER is a manpower modeling tool that is employed by the Enlisted Community Managers (ECMs) at the Bureau of Navy Personnel (BUPERS) to manage the enlisted personnel in the various skill areas. The ECMs and their assistants sometimes find it difficult to learn and use SKIPPER. This study attempts to address these problems by developing an intelligent user coach to aid the ECMs in their use of SKIPPER.

Approach

The user coach was developed by the Navy Personnel Research and Development Center (NPRDC) and BBN Systems and Technologies as a procedural aid that walks the user through a typical manpower planning exercise ("A" School planning). In addition to procedural assistance, the coach also provides a visual picture of the "A" School planning process based on an hydraulic metaphor (e.g., school planning as a series of holding tanks, faucets, pipes, values, etc.). This report describes the user coach and documents a formal evaluation of its effectiveness. The evaluation compared the use and understanding of SKIPPER (1) with the coach, (2) with the coach and the visual metaphor, and (3) without the coach or the metaphor.

Results

One of the most impressive findings was that inexperienced users of SKIPPER were able to successfully complete an "A" School plan without the need for assistance from another person. All users (experienced and inexperienced) were able to complete their "A" School plans significantly faster using the Coach than using SKIPPER unaided. Neither performance on the "A" School planning task, nor the user's understanding of manpower planning, was influenced by providing a visual picture (the hydraulic metaphor) of the process. It appears that users found the procedural (what to do) aids useful, but aids designed to help the user understand the purpose and dynamics (why do it) of the task were not effective.

Recommendations

1. Some form of a step-by-step procedural instruction (wizard) capability should be provided as part of SKIPPER to help inexperienced users perform common tasks.

2. Users seemed to value being told "what to do" considerably more than they valued being told "why do it" by the coach. This preference for "what" over "why" should be studied in greater depth to discover under which conditions a functional (action oriented) mental model is preferable to a structural (concept oriented) mental model.

Contents

List of Tables

List of Figures

Introduction

This report presents the results of an evaluation of a User Coach that was designed to provide assistance to inexperienced users of the Navy's SKIPPER manpower planning model. The Coach was developed as a prototype, to test ideas about how best to assist first-time and infrequent SKIPPER users. The goal of the evaluation was to test the value of those ideas by comparing the performance of users working with the Coach with the performance of users working directly with SKIPPER. The next section describes motivation for developing the Coach, the goals for the Coach, and the Coach design. Later sections explain how we conducted the evaluation and describes how data were collected to evaluate the Coach. The results of the evaluation are presented, including the performance differences found between Coach users and unaided SKIPPER users, the users' subjective evaluations of the Coach and SKIPPER, and the usability problems that were observed for both SKIPPER and the Coach. The last section summarizes the results, and makes recommendations for the direction of future development of a SKIPPER User Coach.

The SKIPPER User Coach

The SKIPPER User Coach (hereafter referred to as the "coach") was designed to help first-time and inexperienced users of the SKIPPER manpower planning model complete "A" School and advancement planning tasks. This section explains why the Coach was needed, summarizes the goals for the Coach, and describes the Coach design.

Need for the Coach

The SKIPPER model has been developed over a number of years by Navy Personnel Research and Development Center (NPRDC) as a tool for Enlisted Community Managers (ECMs) and their assistants at the Bureau of Navy Personnel (BUPERS). SKIPPER supports several major manpower planning tasks, including the development of "A" School Plans. An "A" School provides entry-level training to enlisted personnel entering a rating (enlisted career field), and the "A" School Plan is an estimate of the number of individuals who should enter the "A" School for a specific rating in each year for the next 8 years in order to maintain target inventory levels.

SKIPPER is a spreadsheet-based manpower forecasting model implemented in MS Excel. SKIPPER is fueled by a large database of continuation rates, based on historical data, that estimates the number of individuals expected to leave the Navy each year. Continuation rates are forecasted by rating (career field), paygrade, length of service, and gender. Based on estimates of the number of people who will leave the Navy in a specific rating, SKIPPER estimates how many individuals should enter the rating through "A" School in order to maintain target-rating inventories. SKIPPER also projects the advancements that will be needed to fill vacancies within the rating.

SKIPPER is a powerful tool with many different capabilities. It can project the number of individuals that should be brought into the "A" School for each rating and the

number that should be advanced within the rating. It provides the capability to make projections by gender and length of service (LOS), the capability to access and modify a large database of continuation rates based on historical data, and the capability to modify projections based on policy changes such as changes in selective re-enlistment bonuses. However, along with SKIPPER's power and versatility comes complexity. The novice community manager, confronting SKIPPER for the first time, usually does not know how to proceed. The documentation and on-line help for SKIPPER concentrate on describing all of SKIPPER's capabilities and features, rather than on providing step-by-step instructions for the novice user.

A novice user's problems are exacerbated by the environment in which SKIPPER is used. ECMs and their assistants must switch their attention among many rewiring responsibilities during the year. For example, "A" School planning is typically done in an intensive period twice a year. Any expertise developed in using SKIPPER to prepare a plan may be forgotten before it is time to prepare the next plan.

Also, the ECMs and Assistant ECMs are Navy officers or enlisted personnel who are serving three year assignments, so that there is a constant influx of individuals unfamiliar with the ECM position and with SKIPPER. Currently, SKIPPER knowledge is shared informally among the ECMs and Assistant ECMs, with support from NPRDC. An outgoing ECM may show the incoming ECM how to use SKIPPER, and one ECM may help another with SKIPPER use. Without an individual to advocate SKIPPER and show how it can be used, however, an ECM is unlikely to learn how to use SKIPPER.

Goals of the User Coach

The primary goal of the Coach is to allow new users of SKIPPER to prepare an "A" School plan unassisted—to allow them to use SKIPPER the first time without the need for explanation and assistance from a fellow ECM. The secondary goals of the Coach are to increase the user's understanding of "A" School planning concepts and terminology (if necessary) and to provide them with experience that increases their ability to use SKIPPER unaided.

The Coach was designed to provide instruction and assistance to novice SKIPPER users as they prepare an "A" School plan—currently the most frequent use of SKIPPER—and as they do advancement planning.[1] For "A" School planning, the Coach provides a description of the "A" School planning process, definitions of the terms used in "A" School planning, a graphic overview of "A" School planning concepts (the "Big Picture"), a "wizard" capability for entering data and producing a plan (using the SKIPPER model), guidance in interpreting the plan produced by SKIPPER, and guidance in changing the constraints used by the SKIPPER model. The Coach walks the SKIPPER user through a step-by-step process resulting in the production of an "A" School plan.

---

[1] The advancement planning portion of the Coach was designed but not fully implemented, and it was not included in the evaluation.

The Coach is intended as a prototype, to test ideas for providing support to inexperienced SKIPPER users. The major goal of Coach development was to produce a dynamic prototype that could be used to evaluate the success of the ideas embodied in the Coach and the usefulness of the features it provides.

There has been considerable debate within the human-computer interface (HCI) design community about how users acquire and apply knowledge about a software system (Carroll & Olson, 1990; Preece, et al., 1994). The knowledge that users acquire from experience in using a system may include direct procedural knowledge (rules that prescribe a sequence of actions), general methods that fit general situations and goals, and "mental models" that represent knowledge about the components of a system and their interconnections. Mental models are often described in terms of images (Johnson-Laird, 1983), and there is considerable evidence that the formation and use of mental models draws on components of the human visual system (Kosslyn, 1989). There is a major controversy within the cognitive science community about whether images are distinct from propositional (language-based) representations of knowledge or whether they are a by-product of propositional reasoning (Pylyshyn, 1988).

This controversy about how knowledge is represented and used carries over to a controversy within the HCI design community about how interfaces should be designed to maximize ease of learning and use, and the extent to which visual-image representations and metaphors can assist in this process. Numerous studies have been done to discover whether, and how, people use mental models in interacting with devices and systems (see Rogers, Rutherford, & Bibby, 1992). It is often suggested that metaphors can be used in the design of user interfaces, and that, to the extent that they are congruent with underlying conceptual models of a system, such metaphors can be helpful in learning to use a complex system (Carroll, Mack, & Kellogg, 1990). Preece, et al. (1994) distinguishes structural models, that describe the internal mechanics of a device or system, from functional models that describe how to use the system. In theory, metaphors may be used in an HCI design to support either structural or functional models. Dent-Read, Klein, and Eggleston (1994) analyze the use of pictorial metaphors to guide action (functional models) and conclude that such metaphors can be effective in directing attention to the information needed to guide skilled action.

Prior work and theory did not offer us a strong basis for choosing an instructional method for the Coach—propositional (text) explanations, procedural (action-oriented) assistance, and visual imagery (metaphors) have all been found to be helpful in different HCI applications. The Coach therefore embodies several different theoretical approaches to assisting the inexperienced SKIPPER user. We offer the user text-based propositional instruction on the terms and concepts used in "A" School planning, accessible via a hypertext-like capability to click on words to see explanations. We also offer an image-based picture of the process of manpower planning, designed to represent the relationships among the concepts used in the SKIPPER model pictorially by means of a hydraulic-flow metaphor. Finally, we offer procedural assistance in the form a "wizard-like" capability that allows users to enter a few data items and compute a projection without having to deal with the complex options present in SKIPPER. One of the goals

of the Coach evaluation was to determine which of these different approaches was most helpful to inexperienced users.

Design of the Coach

The Coach was designed as a separate application that runs in parallel with SKIPPER.[2] It was implemented in Visual Basic, using Dynamic Data Exchange (DDE) to pass data back and forth between the Coach and SKIPPER. The Coach walks the user through a step-by-step "A" School planning process. The Coach is implemented using a tab format, and a sequence of actions is imposed by graying out tabs that are not yet available for selection. Terms are defined in the Coach by underlining and highlighting the terms for which definitions are available—users click on the term to open a box containing the definition.

Figure 1 shows the introductory screen for the Coach, which appears when the Coach is launched. Users choose one of the two tabs at the top to start "A" School planning or advancement planning. Figure 2 shows the initial screen for "A" School planning. Highlighted words may be clicked to bring up definitions. The user clicks on the tabs at the bottom of the screen, in sequence, to develop a plan.

Figure 3 shows the "Big Picture" that was designed to provide users who might be unfamiliar with basic "A" School planning concepts with a graphic presentation of the meaning and interrelationships of the major factors involved in planning. The design of the Big Picture was based on a series of interviews with ECMs, which indicated that they often thought of the movement of enlisted personnel through "A" School and into the inventory as a "flow" process, and that a hydraulic metaphor was a good match for their mental models of "A" School planning and inventory projection.[3] The picture shows that personnel come into the inventory through the "A" School and from non-school sources each year, and that not all of the individuals in the inventory in one year will be there in the following year due to the loss rate (people leaving the Navy) from year to year. It also shows that not everyone who enters the "A" School will enter the inventory due to "A" School attrition.

---

[2] This is SKIPPER Version 2, which uses Excel 4 under Windows 3.1.

[3] These interviews were conducted by B. Charles Tatum of NPRDC.

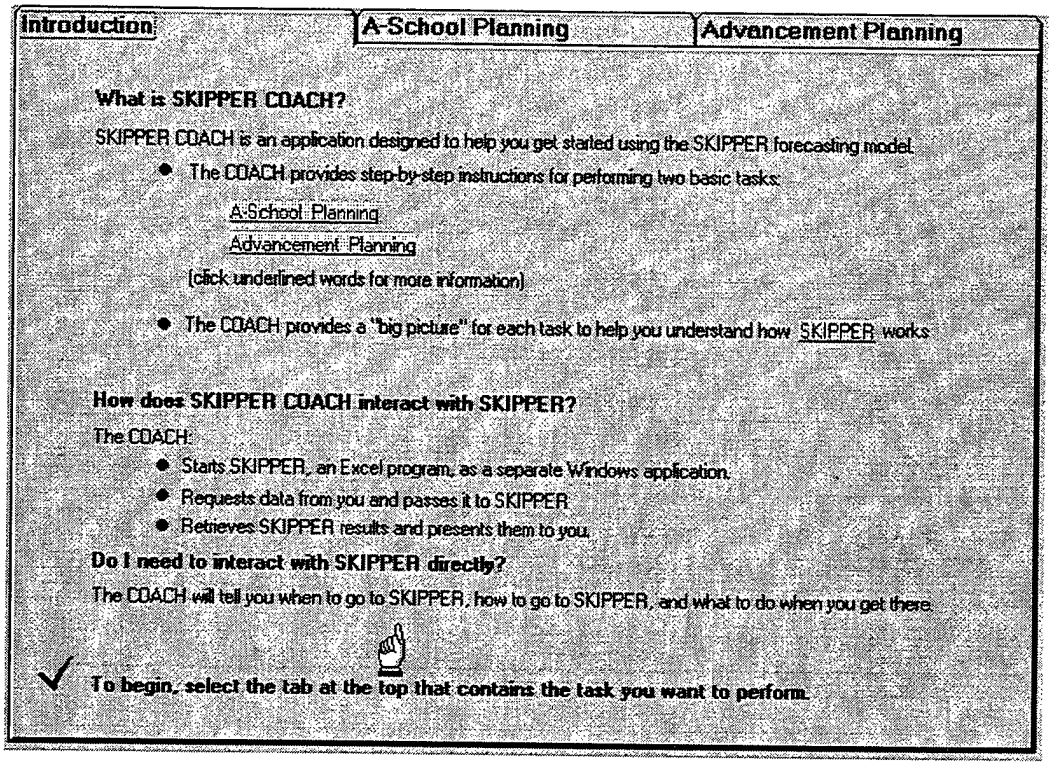


Figure 1. Introductory screen for SKIPPER Coach.

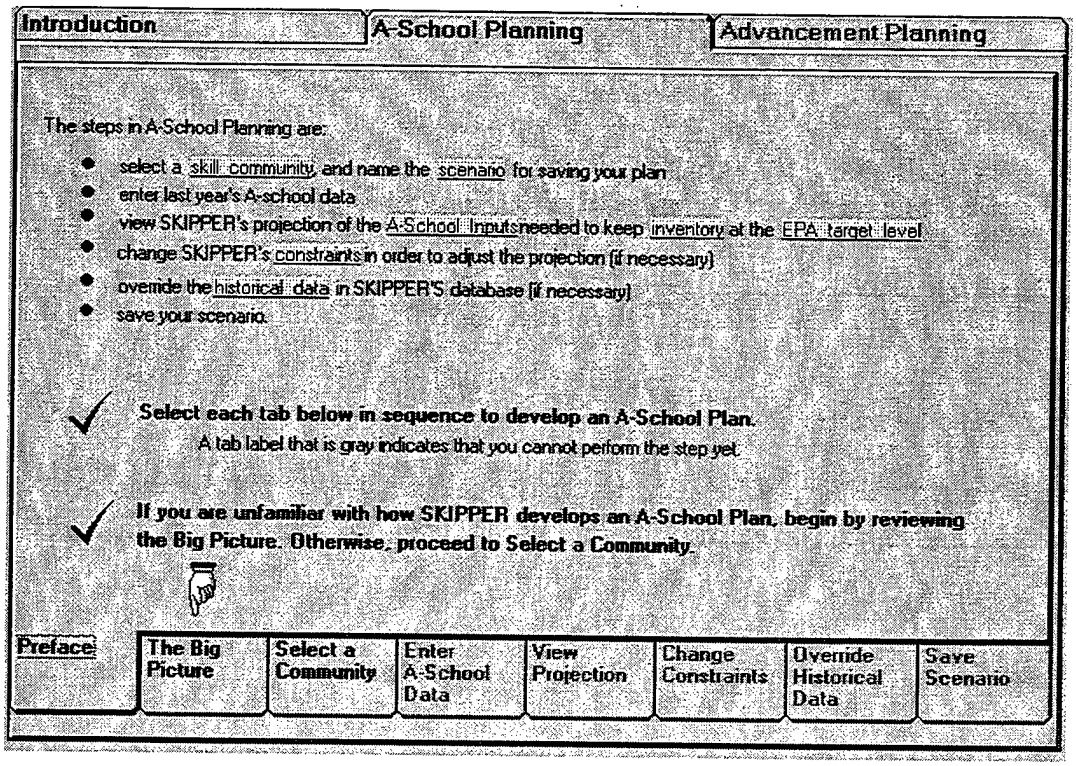


Figure 2. Introduction to "A" School planning.

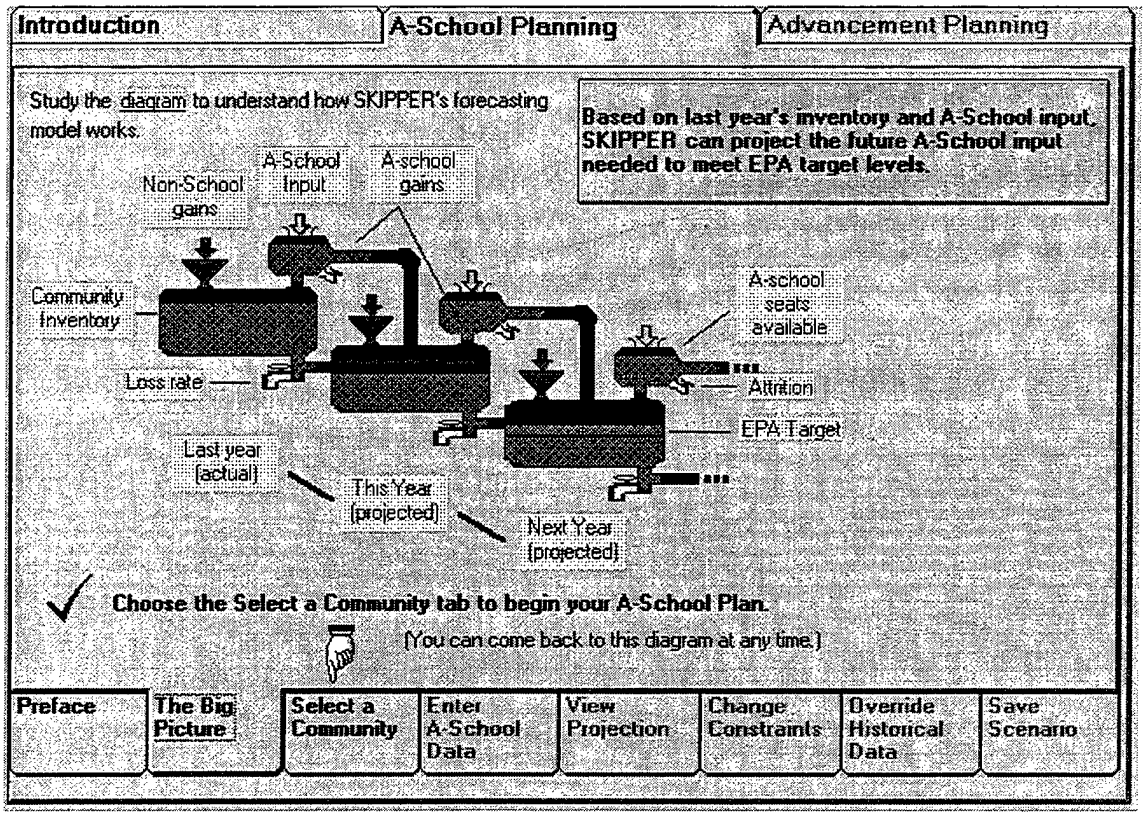


Figure 3. The Big Picture.

Figures 4, 5, 6, and 7 (the Select a Community tab and its subtabs) provide instructions to the user for interacting with the SKIPPER dialog boxes that appear as soon as SKIPPER is launched. These boxes are used to select a skill community for planning and to open a previously saved scenario or start a new scenario. Note the use of a third layer of tabs to the right within the overall Select a Community tab. The dialog boxes shown on the tabs are not "live"—they do not pass data to SKIPPER. On the last of these tabs (scenario) there is a button that launches SKIPPER. SKIPPER is designed so that once the opening dialog process is launched, it is impossible to interrupt it for DDE transfer until the dialog has been completed. This precluded having Coach pass data to SKIPPER at each step, or moving back and forth between Coach and SKIPPER at each step. Either of these options would have been more desirable than the design shown, which forces the user to view and remember the entire dialog process before implementing it in SKIPPER.

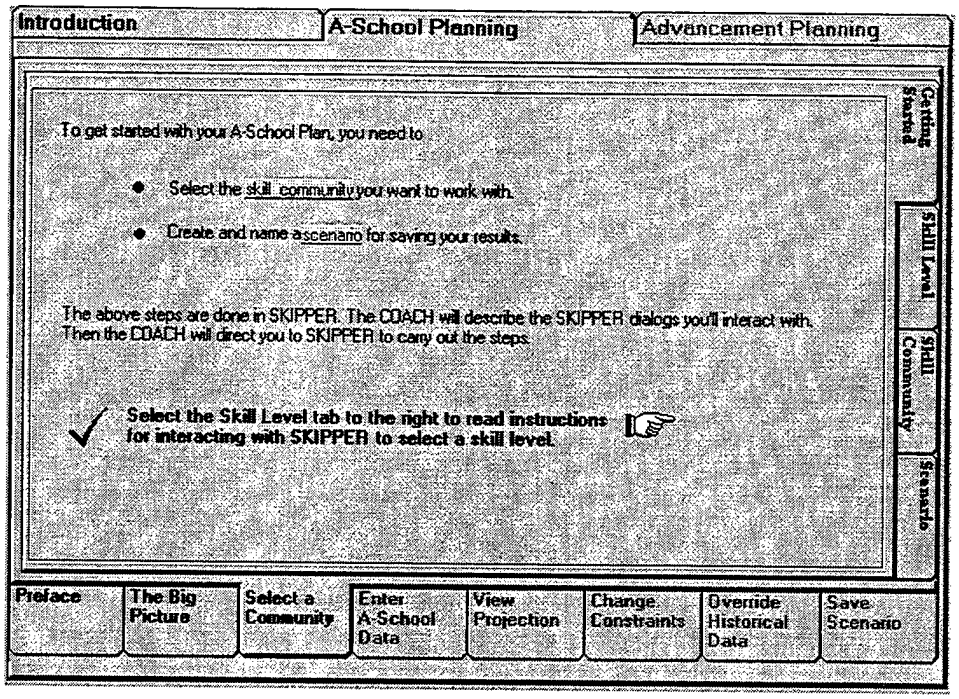


Figure 4. Screen for selecting a community.

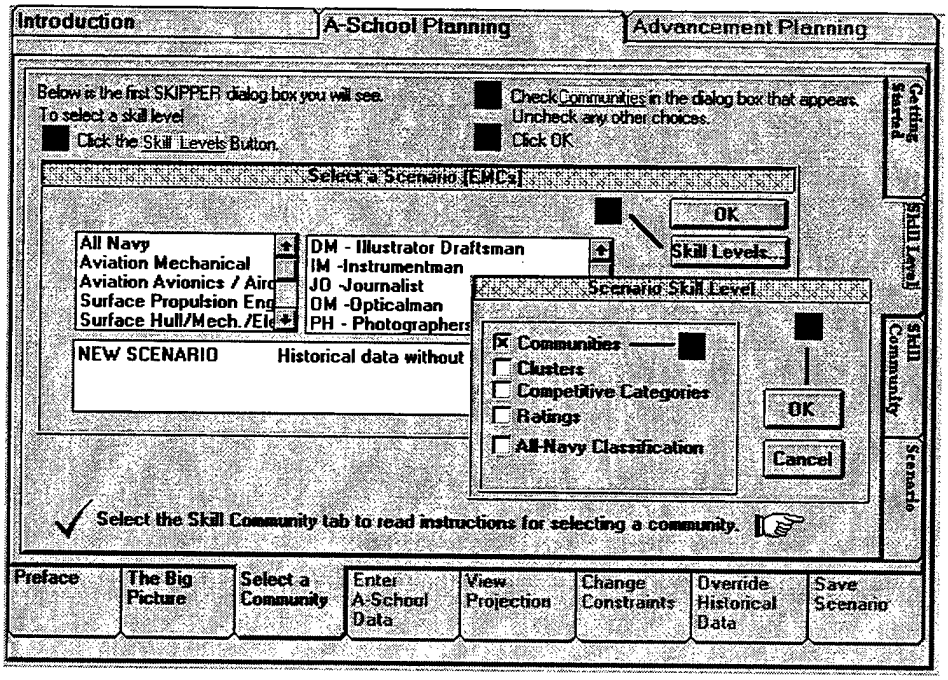


Figure 5. Screen for selecting skill level.

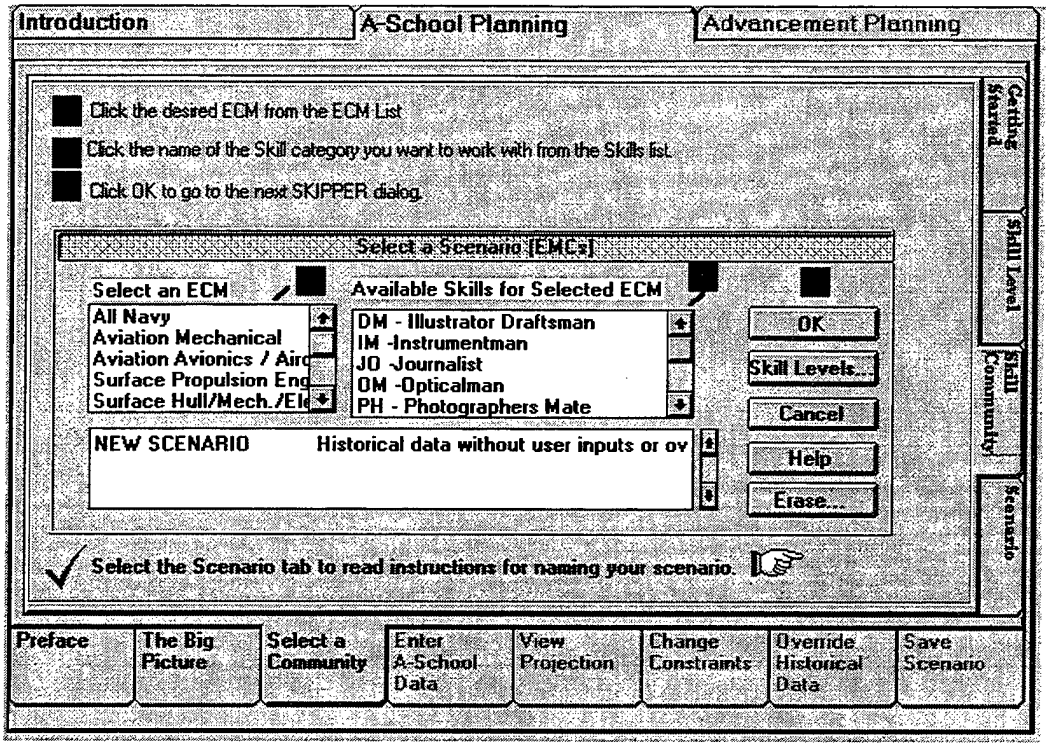


Figure 6. Screen for selecting skill community.

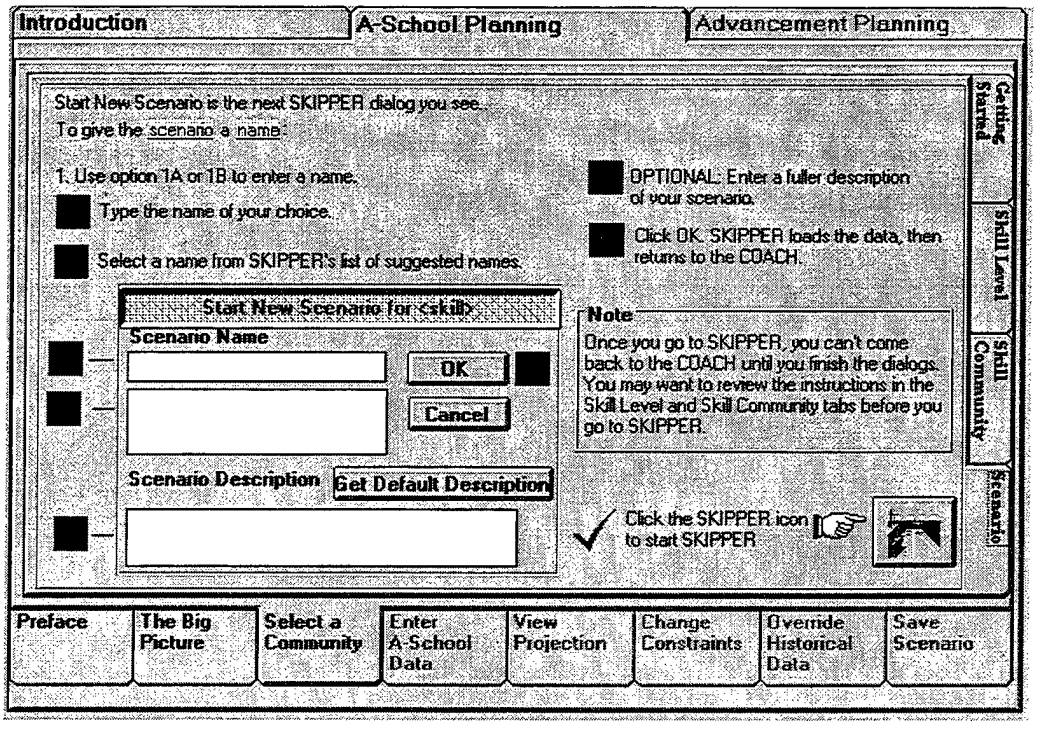


Figure 7. Screen showing scenario options.

Once SKIPPER has been launched, the user is returned to the Enter "A" School data tab, Figure 8. On this tab, the user enters the data that is required in order for SKIPPER to make a recommendation about the number of people to be brought into the "A" School to keep inventory at desired levels. Other data, such as the Enlisted Program Authorization (EPA), which specifies target inventory levels, and historical loss rate data are already available within SKIPPER and need not be entered by the user.

When the user enters data in the three fields on the Enter "A" School data tab, the View Projection tab is highlighted to indicate that the next step is to view SKIPPER's recommendations. When the user clicks on this tab, the data are sent to SKIPPER, recommended "A" School inputs are calculated, inventory is projected, and the user is shifted to the View Projection tab (Figure 9) where the results are displayed. This display focuses attention on the "A" School inputs recommended by SKIPPER, the projected inventory based on these inputs, and the extent to which this inventory meets the EPA target, as indicated by the percentage at the bottom of the data table. The user may click on the See Picture View button at the upper right to see the numerical results superimposed on the Big Picture graphic view (Figure 10).

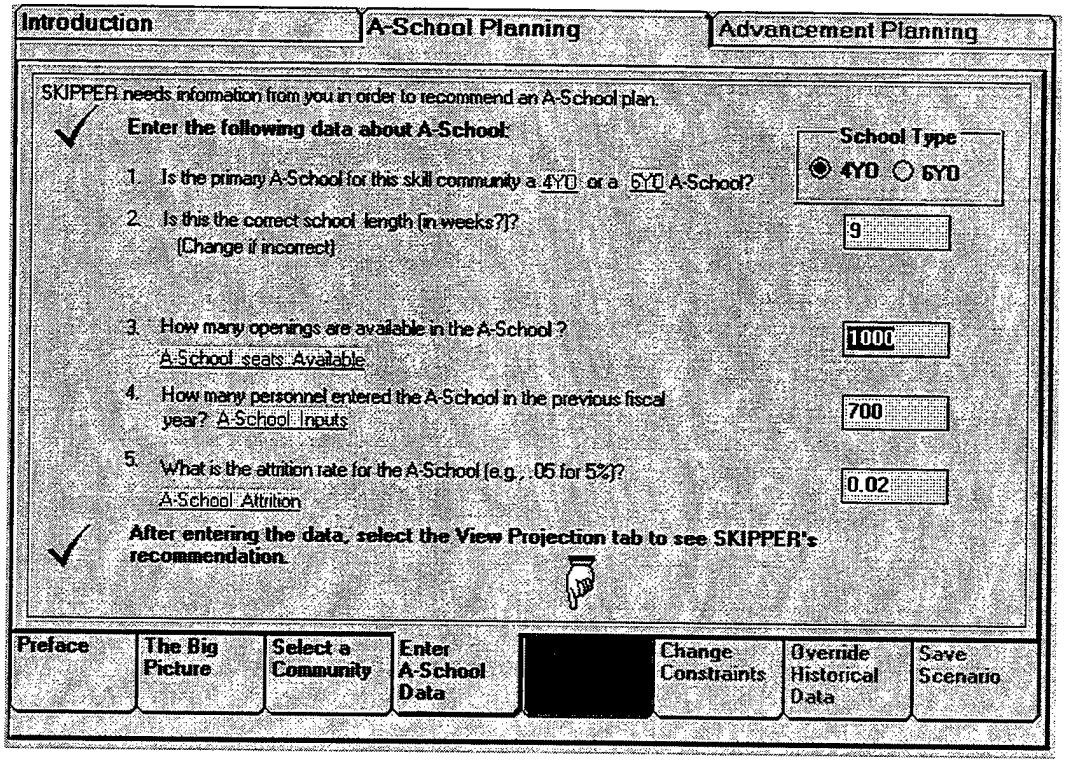


Figure 8. Screen for entering "A" School data.

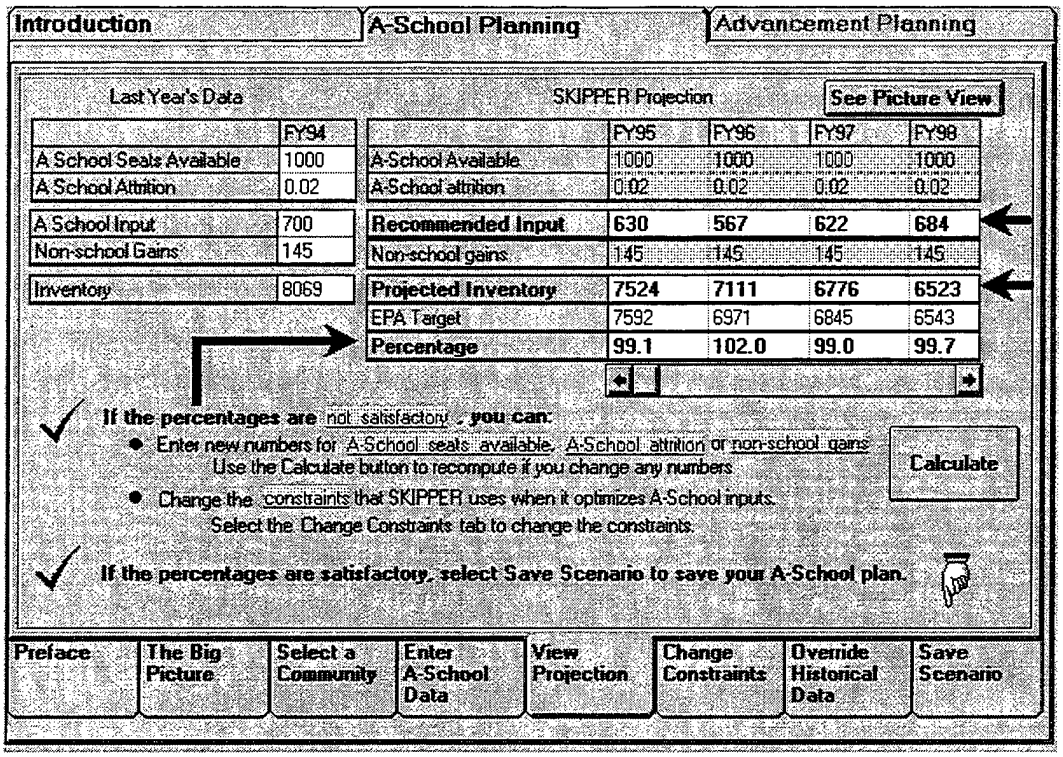


Figure 9. Screen for viewing projections.

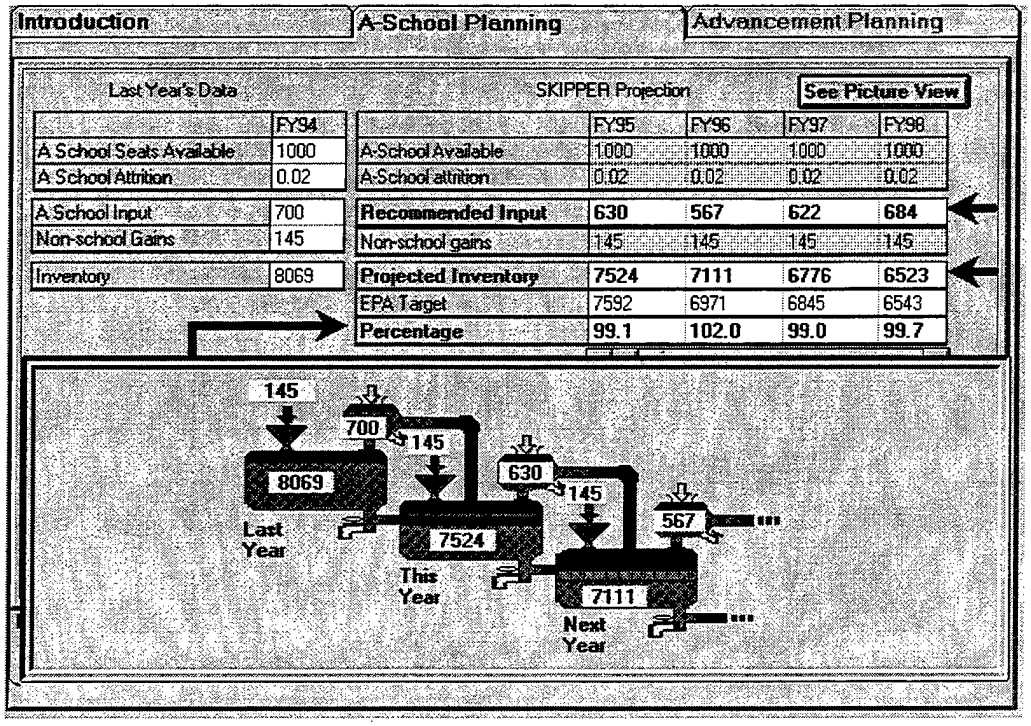


Figure 10. Screen for viewing projections with a view of the Big Picture.

In projecting inventory and recommending "A" School inputs, the SKIPPER model uses a number of constraints that may be modified by the user. The constraints may be modified using the Change Constraints tab (Figure 11). This tab displays a message from SKIPPER indicating which (if any) of the constraints is preventing the projected inventory from reaching EPA target levels. The tab uses a third level of tabs on the right to allow the user to change each of the constraints.

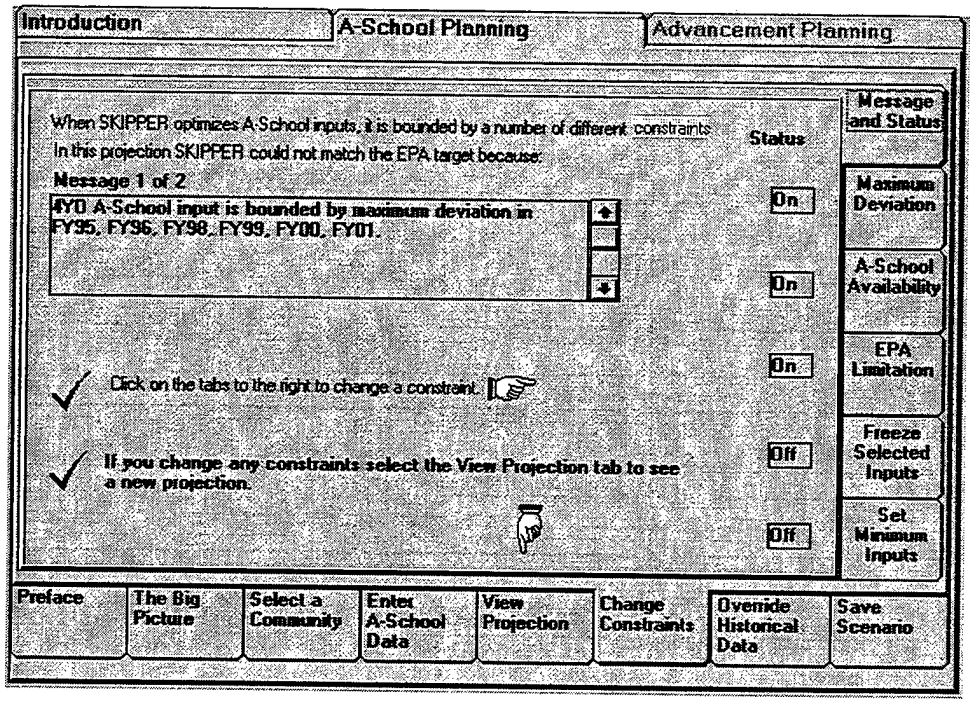


Figure 11. Screen for changing the constraints.

Figure 12 shows the tab used to change the maximum deviation constraint. This constraint limits the fluctuation in "A" School inputs from year to year in order to prevent major disruptions in school size, number of instructors, etc. For example, a maximum deviation of 10 percent means that each year's school inputs cannot be more or less than 10 percent of the inputs in the previous year. The tab allows the user to turn off this constraint or to set its value at any level.

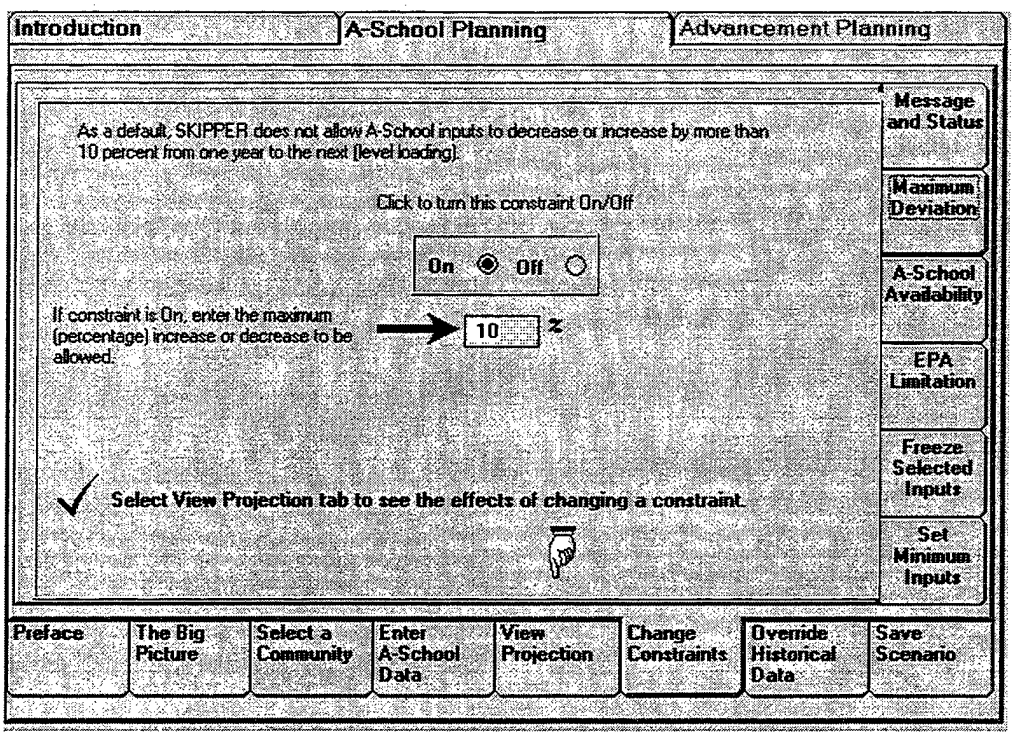


Figure 12. Screen showing display for entering maximum deviation.

When the user is satisfied with the plan, the Save Scenario tab (Figure 13) is used to save the scenario under the name already selected during the SKIPPER start-up dialogs, or to change the name.

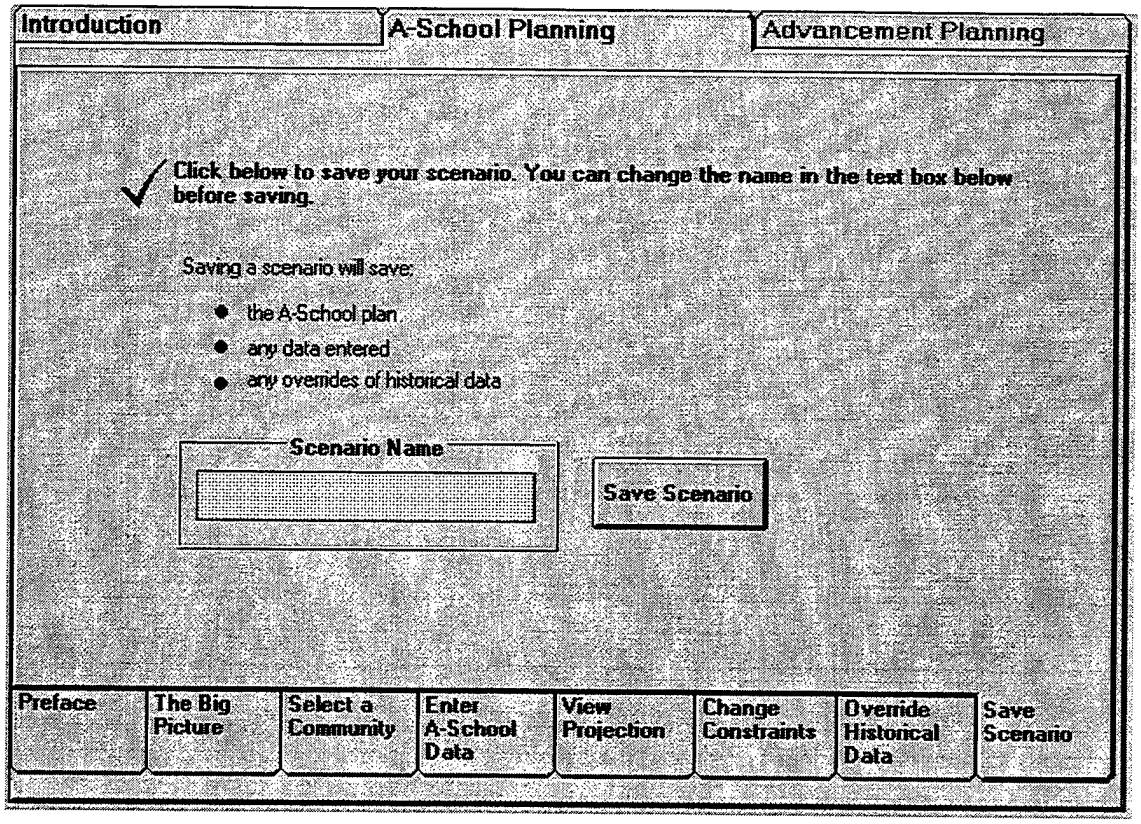
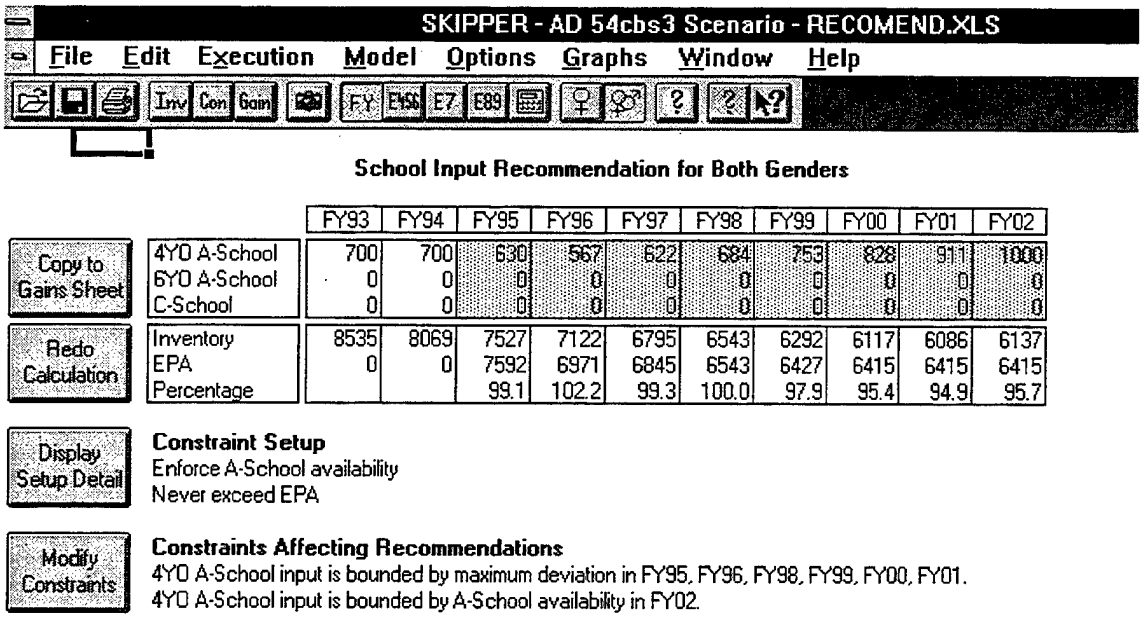


Figure 13. Screen showing display for saving a scenario.

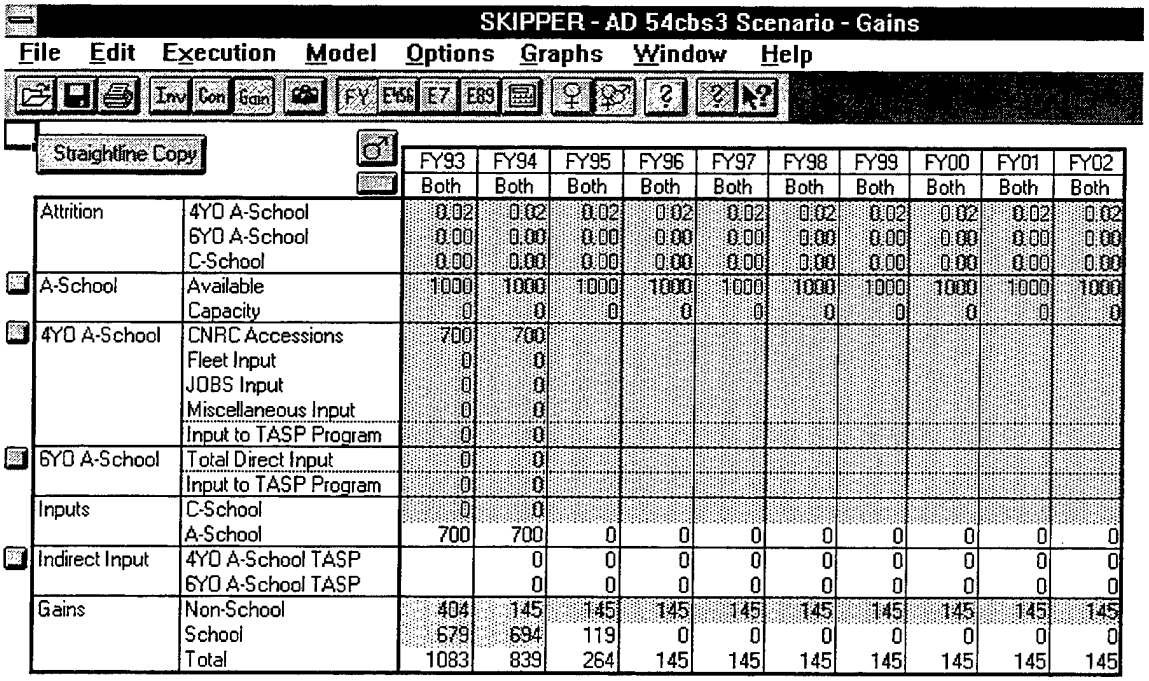
Design of SKIPPER

All of the functions described above for the Coach may also be accessed directly in SKIPPER. When SKIPPER is launched, the user must first complete a series of dialog boxes to select a skill community and select a previously named scenario or start a new scenario. The user is then shown the Inventory sheet (Figure 14). This sheet is not immediately relevant to "A" School planning. To prepare an "A" School plan, the user must transfer to the Gains sheet (Figure 15) in order to enter data. After the appropriate data are entered on this sheet, a menu option is used to have SKIPPER recommend an "A" School plan. The results are then shown on the Recommendations sheet (Figure 16) that also displays the projected inventory based on those recommendations. Constraints may be changed using the button on the bottom left of this sheet.



		FY93	FY94	FY95	FY96	FY97	FY98	FY99	FY00	FY01	FY02
	4YO A-School	700	700	630	567	622	684	753	828	911	1000
	6YO A-School	0	0	0	0	0	0	0	0	0	0
	C-School	0	0	0	0	0	0	0	0	0	0
	Inventory	8535	8069	7527	7122	6795	6543	6292	6117	6086	6137
	EPA	0	0	7592	6971	6845	6543	6427	6415	6415	6415
	Percentage			99.1	102.2	99.3	100.0	97.9	95.4	94.9	95.7

Figure 14. Screen showing inventory sheet from SKIPPER.



			FY93	FY94	FY95	FY96	FY97	FY98	FY99	FY00	FY01	FY02
			Both	Both	Both	Both	Both	Both	Both	Both	Both	Both
	Attrition	4YO A-School	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
		6YO A-School	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		C-School	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	A-School	Available	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
		Capacity	0	0	0	0	0	0	0	0	0	0
	4YO A-School	CNRC Accessions	700	700								
		Fleet Input	0	0								
		JOBS Input	0	0								
		Miscellaneous Input	0	0								
		Input to TASP Program	0	0								
	6YO A-School	Total Direct Input	0	0								
		Input to TASP Program	0	0								
	Inputs	C-School	0	0								
		A-School	700	700	0	0	0	0	0	0	0	0
	Indirect Input	4YO A-School TASP		0	0	0	0	0	0	0	0	0
		6YO A-School TASP		0	0	0	0	0	0	0	0	0
	Gains	Non-School	404	145	145	145	145	145	145	145	145	145
		School	679	694	119	0	0	0	0	0	0	0
		Total	1083	839	264	145	145	145	145	145	145	145

Figure 15. Screen showing the gain sheet from SKIPPER.

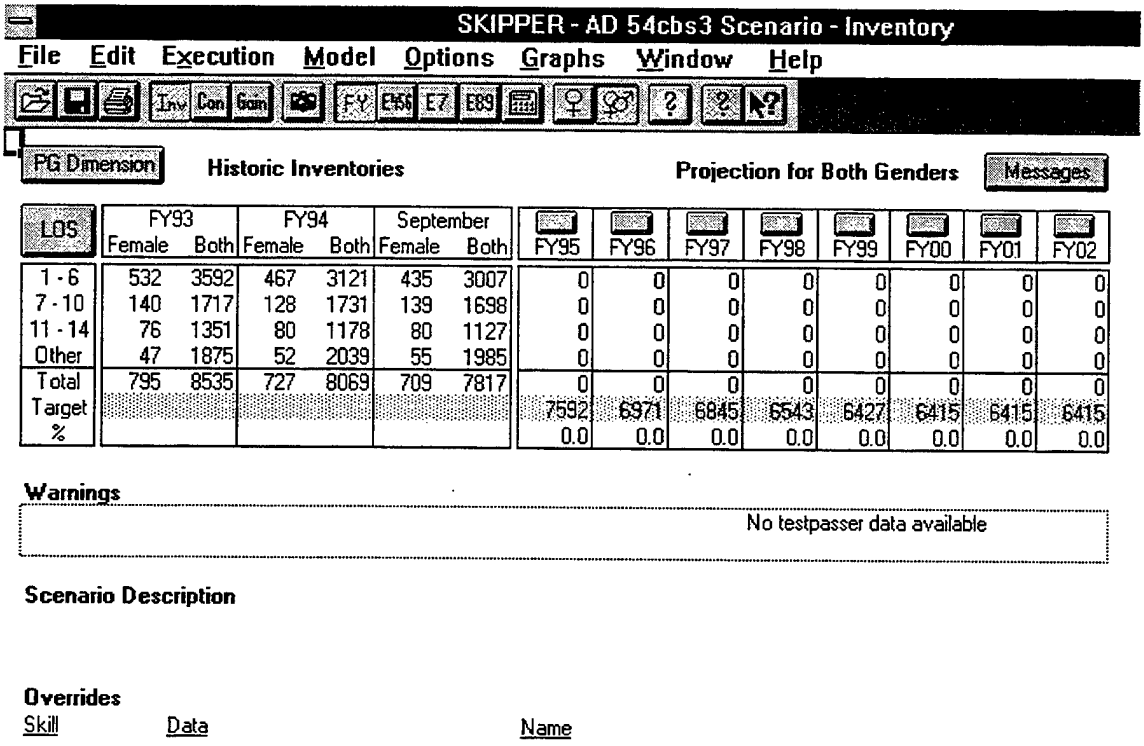


Figure 16. Screen showing recommendation sheet from SKIPPER.

Design of the Evaluation

The Coach was designed to assist new users of SKIPPER in preparing an "A" School plan. We evaluated the Coach to determine whether it met this goal, and to assess its overall usability. We had tested the usability of early designs for the Coach displays using paper prototypes, but a full-scale evaluation using the dynamic prototype provided a much richer opportunity to assess the Coach's usability. By observing users as they worked with the Coach and with SKIPPER, we were able to collect data on the difficulties new users encountered using SKIPPER, the extent to which Coach alleviated those difficulties, the usefulness of Coach features, and the ways that Coach could be improved.

Goals of the Evaluation

The first objective of the evaluation was to assess whether the Coach met its primary goal of allowing new users of SKIPPER to successfully prepare an "A" School plan without assistance. We wanted to determine whether the Coach made it easier (compared with unaided use of SKIPPER) for inexperienced users to produce an "A" School plan. We assessed "easier" in several ways; (1) the amount of time required to complete an "A" School plan, (2) the amount of assistance required to prepare a plan, and (3) users' perceptions of the ease or difficulty of preparing a plan and their overall assessment of the

value of the Coach. The Quality of the plan was not evaluated per se, but the user had to meet a minimum quality standard to complete the task.

We also wanted to evaluate how well the Coach met its secondary goals of increasing users' understanding of "A" School concepts and improving their ability to use SKIPPER unaided. We wanted to measure whether use of the Coach increased inexperienced users' understanding of "A" School planning concepts (compared with unaided use of SKIPPER). We also wanted to measure whether the process of using the Coach to prepare an "A" School plan provided inexperienced users with any knowledge that made it easier for them to prepare an "A" School plan when they used SKIPPER without the Coach.

Finally, we wanted to assess the value of the various Coach features and identify ways in which the Coach could be improved. We wanted to know whether the text-based explanations, the procedural assistance, or the visual image offered by the Big Picture were most helpful to users in understanding "A" School planning and developing a plan. We also wanted to find any usability problems with the Coach—points of confusion or difficulty for users—and develop insights as to how they could be fixed.

As a useful by-product of the evaluation, we wanted to collect data on the usability problems of SKIPPER—what aspects of the SKIPPER interface are most troublesome to inexperienced users? The design of the Coach is based on informal perceptions about the aspects of SKIPPER that are likely to be most troublesome to new users. The evaluation offered an opportunity to systematically collect data from a much larger sample in order to identify the major SKIPPER usability problems and assess the validity of the assumptions on which the Coach was built.

Evaluation Design

A primary goal of the evaluation was to compare the performance of users assisted by the Coach with the performance of unaided SKIPPER users in preparing "A" School plans. We further subdivided the Coach users into users who were provided with the Big Picture metaphor and users who were not provided with the Big Picture in order to determine whether the visual metaphor embodied in the Big Picture had any measurable effects on understanding or performance. This created three conditions for the evaluation: (1) use of Coach with the Big Picture; (2) use of Coach without the Big Picture, and (3) use of SKIPPER unaided. In order to evaluate performance differences between the three conditions, we assigned a different set of users to participate in each condition (between-subjects design). This was necessary because experience using the Coach would be expected to affect an individual's subsequent use of SKIPPER, and vice

versa. We asked users to complete two "A" School planning tasks in the condition to which they had been assigned.[4]

We also wanted to be able to ask users to compare their experiences using the Coach with their experiences using SKIPPER unaided, and to test whether use of the Coach provided users with any knowledge that helped them use SKIPPER unaided. To address these questions, we decided to ask each user to complete the same two "A" School planning tasks in a different condition: Users who had begun by using Coach used SKIPPER unaided, and users who had begun by using SKIPPER unaided used Coach (with the Big Picture).

Table 1 summarizes the design of the evaluation. For purposes of comparing performance with the Coach and SKIPPER, the appropriate data comes from the first two tasks completed. For purposes of measuring any learning effects in using SKIPPER from previously using Coach, or vice versa, the second half of the experiment is of interest.[5] At the end of the session, after users had experience with both Coach and SKIPPER, we asked them to make direct comparisons between their experiences and to assess the overall value of the Coach.

Table 1

Design of the Evaluation

Condition 1 **Coach First**	**Condition 2** **Coach First** **(without Big Picture)**	**Condition 3** **SKIPPER First** **(unaided)**
Task 1 with Coach Task 2 with Coach	Task 1 with Coach (without Big Picture) Task 2 with Coach (without Big Picture)	Task 1 with SKIPPER (unaided) Task 2 with SKIPPER (unaided)
. . . then		
Task 1 with SKIPPER (unaided) Task 2 with SKIPPER (unaided)	Task 1 with SKIPPER (unaided) Task 2 with SKIPPER (unaided)	Task 1 with Coach Task 2 with Coach

[4] As users signed up to participate, we assigned them in sequence to each of the three conditions. This is essentially a random assignment process because there was no systematic pattern in the order in which individuals decided to participate.

[5] Because users were redoing the same tasks in the second half of the evaluation, learning effects would be expected in all three conditions. The effects of previous familiarity with the task should be the same in all three conditions, however.

The Coach is designed for ECMs or Assistant ECMs who are new to their jobs. They may (or may not) have some familiarity with "A" School planning concepts, but have not used SKIPPER to do "A" School planning. For evaluation purposes, it was not possible to draw an adequate sample from this target group—ECMs who were new to their jobs. Instead, we used a mix of individuals who were new to "A" School planning and to SKIPPER, individuals who were somewhat familiar with "A" School planning but not with SKIPPER, and individuals who had used previous versions of SKIPPER for "A" School planning. These users represented the range of individuals who might use the Coach, and we were interested to learn whether their reactions differed.

A total of 33 test users participated. Twelve users were BBN employees who were not familiar with SKIPPER or with "A" School planning. Fifteen users were NPRDC employees who were not familiar with SKIPPER but were somewhat familiar with the concepts of "A" School planning.[6] Six users were ECMs or Assistant ECMs who were very familiar with "A" School planning; five of the six had used a previous version of SKIPPER. These six ECM users represent a sample of the typical advanced user.

Method

All users were given introductory materials that explained "A" School planning concepts and described the "A" School planning task. They were then given two "A" School planning tasks to complete along with the data required for those tasks. The first task was relatively straightforward—SKIPPER produced an acceptable recommendation if the data were entered correctly. The second task was more complex than the first, and required users to reduce the number of seats available in the "A" School in one year, and to develop a plan to compensate for that loss.

Users working with Coach were "on their own"—they were not provided with any verbal assistance by the observers in the room.[7] Users working with SKIPPER unaided (without the Coach) were given a copy of the relevant portions of the SKIPPER manual and were told that they should treat the observer as "the person in the next office" who was available to answer questions if they had problems or were unable to proceed. These evaluation conditions were designed to replicate actual conditions of use. The Coach was designed to be used without assistance, and SKIPPER is typically learned by asking questions and receiving help from another ECM.

All users were asked to think aloud as they worked. The observers kept a record of comments made by users, questions asked, and actions taken, as well as recording the time required to complete each task.

---

[6] Some NPRDC users had seen demonstrations of SKIPPER, but none had used it as part of their jobs.

[7] Pre-planned interventions (not solicited by the user) were necessary at several points because of difficulties in interpreting certain displays. These problems are identified in the discussion of Coach usability issues.

Questionnaire data were collected at several points during a session. Users completed a brief questionnaire after each task about the difficulty of completing the task. After completing two planning tasks, users completed a "concepts quiz" designed to test their knowledge of "A" School Planning concepts. After completing two tasks using the Coach, users completed a questionnaire assessing the usefulness of the various Coach features. Also, we asked users who had viewed the Big Picture to complete a "metaphor" rating matrix in which they rated how well various phrases described the elements of the Big Picture, and how well those same phrases described "A" School planning concepts. At the end of the session, after users had a chance to complete two planning tasks using the Coach and two using SKIPPER unaided, they completed a comparison questionnaire asking them to rate and compare their experiences with both the Coach and SKIPPER. Copies of the introductory materials provided, the tasks, and the questionnaires are included in a report by MacMilliam and Freeman (1996).

Results

The evaluation produced several types of results. This section begins by discussing performance differences among the three conditions, then presents questionnaire data on users' perceptions of the Coach and SKIPPER, and concludes by describing the usability problems that were observed for both the Coach and SKIPPER.

Performance Differences

Performance measures include the time needed to complete "A" School planning tasks, the number of questions asked in order to complete the tasks, and the scores on a "quiz" that measured understanding of "A" School planning concepts.

Time to Complete the First Two Tasks

The primary measure of performance differences among the three conditions in the evaluation was the time required to complete an "A" School plan, supplemented by the number of questions asked in the unaided SKIPPER condition. We expected that inexperienced users would be able to complete their first "A" School plans more quickly using the Coach, without the need for assistance from the observer, and that unaided SKIPPER users would have more difficulty in completing their first plans, causing them to take more time and to ask for assistance one or more times.

Table 2 shows the mean time to complete the first and second tasks in each of the three conditions. A one-way analysis of variance for Task 1 shows that the time to complete the first task was significantly different for the three groups ($F = 9.60$; $df = 2,30$; $p < .001$).[8] Pre-planned contrasts show that each of the Coach conditions is

[8] We also analyzed the data in Table 2 as a two-way ANOVA. In this analysis, there was a significant main effect for condition ($F = 5.26$, $df = 2,60$, $p < .01$) and a significant interaction effect ($F = 3.60$; $df = 2,60$; $p < .05$). The significant interaction results from the pattern that is evident in Table 2—the users in the three conditions differed substantially in the time taken to complete the first task, but not the second.

significantly different from the SKIPPER condition (for Coach versus SKIPPER, F = 8.63, df = 1,30; p <.01; for Coach without Big Picture versus SKIPPER, F = 17.91, df = 1,30; $p < .001$). The times for the two Coach conditions do not differ significantly from each other, however (F = 1.72; df = 1,30). We conclude that the Coach, with or without the Big Picture, significantly reduced the time needed for new users to complete their first "A" School plan.

Table 2

Mean Time to Complete Tasks 1 and 2

Mean Time	**Condition 1 Coach First (n = 11)**	**Condition 2 Coach First (without Big Picture) (n = 11)**	**Condition 3 SKIPPER First (unaided) (n = 11)**
Mean time to complete Task 1 (in minutes)	17.1	13.5	25.1
Mean time to complete Task 2 (in minutes)	16.5	15.8	17.0
Mean time to complete both tasks	33.6	29.4	42.1

This pattern held for all three groups of users. Table 3 shows the mean time to complete the first task for BBN users, NPRDC users, and ECMs. For each group, times were longest for the unaided SKIPPER condition, and fastest for Coach without the Big Picture. Separate ANOVAs for each group[9] show significant differences in the three conditions for the BBN group (F = 6.73; df = 2,9; $p < .05$) and the NPRDC group (F = 4.46; df = 2,12; $p < .05$) but not for the ECM group due to the small sample size.

Unaided SKIPPER users asked a number of questions of "the person in the next office" as they developed their first "A" School plan. The mean number of questions asked by SKIPPER users (Condition 3) was 4.6 for Task 1 and 2.1 for Task 2. Many of the questions asked by users concerned where to enter their data on the SKIPPER spreadsheets. When SKIPPER is started, the first sheet to appear is the Inventory sheet. Users typically did not know the purpose of this sheet (it was not relevant to "A" School planning) and did not know how to move to other sheets. Their first question was often "Where do I enter "A" School data?" Even after users found the Gains sheet, they often did not know where to enter their data. Another frequent question was "Have I entered

[9] The unequal sample sizes in the three groups make a two-way ANOVA difficult.

the data correctly?" Users were often puzzled about how to invoke the command that caused SKIPPER to calculate recommended school inputs, and how to adjust constraints in order to change SKIPPER's recommendations. Although some users were eventually able to find the information they needed in the SKIPPER Manual, most asked questions in order to be able to proceed. Many users commented that they would not have been able to complete their plan in the unaided SKIPPER condition without the help of "the guy in the next office."

Table 3

Mean Time to Complete Task 1 by Condition and Group

Group	Condition 1 Coach First	Condition 2 Coach First (without Big Picture)	Condition 3 SKIPPER First (unaided)
BBN Users ($n = 12$)	20.5 ($n = 4$)	17.5 ($n = 4$)	32.5 ($n = 4$)
NPRDC Users ($n = 15$)	14.4 ($n = 5$)	11.4 ($n = 5$)	20 ($n = 5$)
ECM Users ($n = 6$)	17 ($n = 2$)	11 ($n = 2$)	23 ($n = 2$)

Understanding of "A" School Concepts

We had expected that the Coach, especially the version of the Coach that included the Big Picture, might help users to form a better understanding of the multitude of factors that go into "A" School planning and the relationships among those factors. We designed a manpower planning "concepts quiz" to measure users' understanding of "A" School planning concepts, and administered it after users had completed their first two planning tasks.

Table 4 shows the mean scores (out of a maximum of 15) for the three conditions. Although the means are slightly different, there was considerable variability in the scores and the differences in means are not significant ($F = 1.64$, $df = 2,30$, $p = .21$). We conclude that the Coach, with or without the Big Picture, did not significantly increase first-time users' understanding of "A" School planning concepts as measured by our questionnaire.

We examined the data to determine whether there might have been an effect of the Coach on concept understanding for less experienced users—the BBN and NPRDC users—but not for the ECMs. There is no indication of this, however. There are no significant differences among the three conditions for the BBN or for the NPRDC users.

Table 4

Mean Score on Manpower Planning Concepts Quiz by Condition (maximum of 15)

	Condition 1 Coach First (n = 11)	Condition 2 Coach First (without Big Picture) (n = 11)	Condition 3 SKIPPER First (unaided) (n = 11)
Mean score	12.3	10.9	9.3

The failure to find an effect of the Coach on manpower planning concept understanding may, of course, be due to the design of our measurement instrument, which may not produce a valid measure of concept understanding. It may also be due to the written introductory material that we developed and provided to all users (in all three conditions) before they began work. This introduction to "A" School planning described the factors involved in "A" School planning and defined the terms used. The intent of the instructions was to create a "level playing field" so that completely inexperienced individuals could serve as valid test users in Condition 3. We could not have asked inexperienced individuals to develop "A" School plans using SKIPPER unaided without giving them some introductory training in "A" School planning concepts—it would have been an impossible task. In retrospect, however, we may have done too good a job with this training from the point of view of the evaluation. The written introduction may have provided users with most or all of the information that they could have gained from the Big Picture or from the other instructional material in Coach, wiping out any differences among the three conditions in concept understanding.

Transfer of Training From the Coach to SKIPPER

A secondary goal of the Coach was to improve users' understanding of the "A" School planning task in a way that made it easier for them to use SKIPPER unaided. Users could continue to use Coach to develop their "A" School plans if desired, but direct use of SKIPPER would give them access to a larger range of features and capabilities. There are several ways to assess how much users learned from the Coach that helped them with SKIPPER. First, we can look at how much time was required to complete the second set of tasks, in which Coach users prepared plans using SKIPPER, and SKIPPER users prepared plans using Coach. We would expect that the second set of tasks might be easier and faster for users no matter which system they started with, both because they had gained familiarity with "A" School planning and because they had seen the details of the tasks before. We expected, however, that initial use of the Coach may have reduced the time needed to use SKIPPER for the first time considerably more than initial use of SKIPPER reduced the time needed to use the Coach for the first time.

Table 5 shows the time needed to complete Tasks 1 and 2 using the Coach, and the amount of time needed to complete the same tasks using SKIPPER. The table shows that users who had already completed two tasks using SKIPPER (Condition 3) took about the same amount of time (32 minutes) to complete the same two tasks using the Coach as the users in Conditions 1 and 2 who were completing these tasks for the first time using Coach (33.6 minutes and 29.4 minutes). Users who began by using the Coach, however, were able to complete the same two tasks using SKIPPER more quickly (28.8 minutes and 30.2 minutes) than users completing those tasks for the first time using SKIPPER (42.1 minutes).

Table 5

Mean Time to Complete Two Tasks With SKIPPER and Coach (shaded cells show times for the second set of tasks completed)

Mean Time	Condition 1 Coach First (n = 11)	Condition 2 Coach First (without Big Picture) (n = 11)	Condition 3 SKIPPER First (unaided) (n = 11)
Mean time to complete two tasks using Coach (min.)	33.6	29.4	32.0
Mean time to complete two tasks using SKIPPER unaided (min.)	28.8	30.2	42.1

A two-way mixed analysis of variance (with repeated measures on tool) shows a significant effect for tool (Coach versus SKIPPER) ($F = 4.15$; $df = 1,60$; $p < .05$) and a significant interaction between condition and tool ($F = 8.53$; $df = 2,60$; $p < .001$).[10] Examination of the table shows that the interaction results from the longer time needed in Condition 3 to complete the two plans using SKIPPER by users who had not yet used the Coach. Mean times in all the other cells are similar. Prior use of the Coach apparently provided first-time SKIPPER users with an advantage over totally inexperienced users.

The number of questions asked during unaided SKIPPER use provides another indication of whether users transferred any knowledge from their use of the Coach. Table

[10] The two-way ANOVA has one within-subjects factor (Tool: Coach versus SKIPPER) and one between-subjects factor (Condition). Note that both factors are fixed, not random, i.e., the levels of each factor are not randomly sampled from a larger population. For this design, it is appropriate to use subjects within tool-by-condition combinations as the error term for both factors and their interaction. (See Rosenthal and Rosnow, 1984.)

6 shows the mean number of questions asked during SKIPPER use for subjects in each of the three conditions. Users who had previously used Coach asked about half as many questions when they used SKIPPER as users working with SKIPPER for the first time. Two-way analysis of variance shows a significant effect for Condition (SKIPPER-first users asked more questions than Coach-first users) ($F = 4.05$, $df = 2,60$; $p < .05$) and a significant effect for Task (users asked more questions during Task 1 than Task 2) ($F = 9.26$. $df = 1,60$; $p < .01$).

Table 6

Mean Number of Questions Asked During SKIPPER Use by Condition

Mean number of questions asked during SKIPPER use	**Condition 1 Coach First ($n = 11$)**	**Condition 2 Coach First (without Big Picture) ($n = 11$)**	**Condition 3 SKIPPER First (unaided) ($n = 11$)**
Task 1	2.9	1.7	4.6
Task 2	0.5	1.5	2.1
Total	3.4	3.2	6.7

In the end-of-session questionnaire, we also asked Coach-first users in Conditions 1 and 2 if their experience with Coach had helped them use SKIPPER. These results are shown in Table 7. Responses were made on a scale from 1 (Coach experience not at all helpful) to 5 (Coach experience very helpful). Overall, ratings fall between 3 and 4, indicating that users felt their Coach experience had been somewhat helpful in using SKIPPER, but not "very helpful."

Table 7

Ratings of Helpfulness of Coach Experience in Using SKIPPER

Question Did your experience with Coach help you use SKIPPER in:	**Mean Rating ($n = 22$)**
Developing an "A" School plan	3.5
Answering SKIPPER start-up dialogs	3.7
Understanding SKIPPER data tables	3.6

Note: Scale: 1 = Coach experience not at all helpful; 5 = Coach experience very helpful.

User Perceptions About Coach and SKIPPER

We administered a series of questionnaires to collect data on users' perceptions about the usefulness and usability of both the Coach and SKIPPER. Data were collected after each task was completed, after use of the Coach, and at the end of the session when users had experienced both Coach and SKIPPER and were able to make direct comparisons.

Perceived Ease of Use of the Coach and SKIPPER

Three types of data are available on how easy or difficult users found it to work with the Coach and to work with SKIPPER directly: rating data from the end-of-task questionnaires for tasks in which Coach was used and tasks in which SKIPPER was used, rating data on the overall ease of use of SKIPPER and Coach from the end-of-session questionnaire, and direct comparison ratings of Coach and SKIPPER from the end-of session questionnaire.

Table 8 shows the ratings given by users on the difficulty of completing their first "A" School planning task in each of the three conditions. These results are taken from the end-of task questionnaire administered after the first task was completed. Coach users had not yet worked with SKIPPER unaided, and SKIPPER users had not yet seen the Coach.

User ratings were significantly more positive in Conditions 1 and 2 than in Condition 3 for all of the questions asked on the questionnaire. The differences are especially large for the question that asked if users understood how to carry out each step, and for the question that asked about the overall ease or difficulty of developing a plan. Inexperienced users felt that "A" School planning was considerably easier using the Coach than using SKIPPER without the Coach.

Table 8

Mean end-of-Task Ratings by Condition for the First Task Completed

Question	Condition 1 Coach First (n = 11)	Condition 2 Coach First (without Big Picture) (n = 11)	Condition 3 SKIPPER First (unaided) (n = 11)	Significance of Differences Among Conditions (One-way ANOVA)
Did you understand what your goal was at each step? (1 = never understood; 5 = always understood)	4.3	4.4	3.5	$F = 4.58$ $df = 2,30$ $p < .05$
Did you understand how to carry out each step? (1 = never understood; 5 = always understood)	3.6	4.0	2.4	$F = 16.67$ $df = 2,30$ $p < .001$
When you took an action with the software, did you understand the results? (1 =n ever understood; 5 = always understood	3.9	3.8	3.1	$F = 3.28$ $df = 2,30$ $p < .05$
How easy or difficult was it to develop an "A" School plan using this software? (1 = very difficult; 5 = very easy)	4.2	4.3	2.5	$F = 17.37$ $df = 2,30$ $p < .001$

At the end of the evaluation session, after all of the users had experience with both SKIPPER and the Coach, we asked them to rate the ease or difficulty of developing an "A" School plan with SKIPPER and the ease or difficulty of developing an "A" School plan with the Coach, and to make direct comparisons between their experiences. Table 9 shows these results.

There was a large and significant difference between the mean ease-of-use ratings for the Coach (4.2 on a scale of 5) and the mean for SKIPPER (2.6) (paired comparison t-test $t = 7.78$, $df = 32$, $p < .001$). There was little difference in these ratings in the three conditions.

Table 9

Mean End-of-Session Ratings of Coach and SKIPPER

Question	Mean Rating (n = 33)	
Ease of use questions (1 = very difficult; 5 = very easy)		Paired-comparison t-test for difference
Rate the ease or difficulty of developing an "A" School plan using SKIPPER	2.6	
Rate the ease or difficulty of developing an "A" School plan using Coach	4.2	7.78 df = 32 p < .001
Direct-comparison questions (1 = SKIPPER much easier; 5 = Coach much easier)		.05 confidence interval (neutral = 3)
How easy was it to use the software to develop an "A" School plan?	4.2	3.9—4.5
How well did you understand manpower planning concepts?	3.3	3.1—3.6
How easy was it to follow the steps in developing an "A" School plan?	4.2	3.8—4.5

In the direct-comparison questions in Table 9, users rated the Coach as much easier to use than SKIPPER (4.2 on a scale of 5) and indicated that it was much easier to follow the steps in developing an "A" School plan in Coach (4.2 on a scale of 5). The .05 confidence intervals around these ratings do not include 3 (a neutral rating), indicating that Coach was significantly preferred to SKIPPER. This effect was not so strong for the question that asked if users understood "A" School planning concepts any better using Coach than using SKIPPER. The mean rating was 3.3, almost at the middle of the scale, and the confidence interval just barely excludes 3. This is consistent with the failure to find an effect of Coach on overall understanding of "A" School concepts, as reported above.

Perceived Overall Value of the Coach

Not surprisingly, users perceived the Coach as quite valuable in preparing an "A" School plan. When asked to rate the overall value of the Coach for "A" School planning, users gave it a mean rating of 4.2 (n = 33) on a scale that ranged from 1 (not at all valuable) to 5 (very valuable). There was little difference in these ratings among the three conditions, and the three groups of users also produced very similar ratings of overall value (BBN users, 4.2; NPRDC users, 4.1; ECM users, 4.3).

Perceived Value of Coach Features

After users had completed two tasks with the Coach, we asked them to rate the usefulness of the various Coach features. Users who had seen the Big Picture (Conditions 1 and 3) were also asked about its usefulness. Table 10 shows these results. In general, ratings close to three indicate limited usefulness of the feature, and ratings of four or higher indicate the most useful features. Overall, the Enter "A" School data screen was rated as the most useful feature of the Coach (4.5). This was the screen that allowed users to enter data and access a "wizard" feature that produced recommended "A" School inputs and projected inventory. The screen that displayed these results was also rated as useful (4.1), as was the introductory text that described "A" School planning (4.0). The instructions provided by the Coach on how to respond to SKIPPER dialogs were of only moderate usefulness (3.5). The least useful features of the Coach appear to be the Big Picture (3.0 and 3.2) and the See Picture View option on the View Projections screen (2.9).

We also asked users to complete two open-ended questions in which they listed the most and least useful features of the Coach. These results can be seen in a report by MacMillian and Freeman (1996). Many of the positive comments listed in the report refer to Coach's sequential step-by-step organization (e.g., sequential organization, logical guidance, step by step process). Several users cited the tab layout as a useful feature, and a number mentioned the data input screen. One ECM user commented: "Very user friendly, especially for someone like myself who has used SKIPPER only once."

MacMilliam and Freeman (1996) also list users' responses to the question that elicited the least useful features of the Coach. Users often mentioned the series of tabs that was intended to introduce them to SKIPPER's dialog boxes as not being useful. Users did not like having to go through all of the steps before they could go to SKIPPER to carry them out, and several mentioned mistakenly thinking that the pictures of the dialog boxes in Coach were interactive. Several mentioned being confused as to whether they were in Coach or SKIPPER. Others mentioned the Big Picture or "flow diagram" as not very helpful. Although users were positive in general about the tab-based layout, some had trouble navigating the tab interface and did not like having multiple tabs on several sides of the screen.

Table 10

Mean End-of-Session Ratings of Usefulness of Coach Features

Question	Mean Rating
Ratings of Coach Features	**(n = 33)**
Usefulness of notebook format using tabs. (1 = not at all useful; 5 = very useful)	3.8
Usefulness of introductory text that gave overview of "A" School planning. (1 = not at all useful; 5 = very useful)	4.0
Usefulness of screens that explained how to respond to SKIPPER dialogs. (1 = not at all useful; 5 = very useful)	3.5
Understanding of screen for entering required "A" School data. (1 = not at all understandable; 5 = very understandable)	4.5
Understanding of screen that displayed projections. (1 = not at all understandable; 5 = very understandable)	4.1
Ease or difficulty of using constraints options. (1 = very difficult; 5 = very easy)	3.8
Clarity of Coach instructions. (1 = not at all clear; 5 = very clear)	3.9
Ratings of the Big Picture	**(n = 22)**
How useful was the Big Picture in helping you understand how "A" School planning relates to manpower planning? (1 = not at all useful; 5 = very useful)	3.0
How well did the Big Picture clarify the underlying concepts involved in manpower planning? (1 = not at all useful; 5 = very useful)	3.2
On the View Projections tab, how useful was it to see the projected data transferred onto the Big Picture? (1 = not at all useful; 5 = very useful)	2.9

Extent to Which Coach Addresses Problems Using SKIPPER

In the end-of-session questionnaire, we asked the users in Condition 3, who had used SKIPPER before they used the Coach, to list their problems with SKIPPER and comment on whether they thought Coach addressed those problems. These results are reported by MacMillian and Freeman (1996). Most users mentioned that they did not know where to enter the data in SKIPPER, which data to enter, and where to find the command to project

inventory and recommend "A" School inputs. Most felt that Coach would have helped them with these problems, and indicated that, overall, most of their problems would have been addressed by the Coach. Problems with SKIPPER not addressed by the Coach include confusion between maximum deviation of "A" School input from year to year (expressed as a percentage) and projected inventory as a percentage of EPA, and the time needed to open a scenario.

Perceived Similarity Between Big Picture Elements and "A" School Concepts

The design of the Big Picture was based on interviews with ECMs in which they were asked about their "mental models" of inventory projection and "A" School planning, and whether they used any metaphor in thinking about the concepts involved. A number of the individuals interviewed mentioned "flow" in hydraulic systems as a metaphor for manpower planning in the Navy, and this image was used as the basis for the Big Picture. In the evaluation, we collected data on the extent to which users perceived the features of the elements of the Big Picture (tanks, faucets, etc.) as matching the features of the concepts used in "A" School planning (inventory, attrition, etc.). These data are in the form of a matrix (see MacMillian, Getty, Tatum, & Ropp, 1998), with the elements of a hydraulic system and the major "A" School planning concepts forming the rows, and descriptive phrases forming the columns. The user's task was to fill in the cells of the matrix by rating how well each phrase described each hydraulic element or "A" School concept. All users who saw the Big Picture (Conditions 1 and 3) completed the matrix.[11]

In analyzing the congruence between the elements in the Big Picture and users' mental models of "A" School planning concepts, we were interested in determining the extent to which users' perceptions of similarity matched the assumptions about similarity that are built into the Big Picture. Each "A" School concept is represented graphically in the Big Picture by an element in the hydraulic system, e.g., attrition rates and loss rates are represented by faucets, non-school gains are represented by funnels, and inventory is represented by water in a tank. If these are "good" representations, then users' ratings of how well various phrases describe a hydraulic element and their ratings of how well those same phrases describe the concept it represents should be similar, i.e., highly correlated.

Table 11 shows the correlation between users' ratings for each of the hydraulic elements and each of the "A" School concepts. The descriptive phrases on which each element and concept was rated are listed at the bottom of the table. The shaded cells in the table show the element-concept pairings that are used in the Big Picture representation, e.g., "A" School attrition is represented by water moving through a faucet.

---

[11] One ECM user was unable to complete the matrix at the end of the session because he was called away on an urgent task, so $n = 21$.

Table 11

Correlation of Descriptive Ratings for Hydraulic Elements in the Big Picture and "A" School Planning Concepts (shaded cells indicate Big Picture representation)

Hydraulic Elements	Water Through a Faucet	Tank	Water Through a Funnel
"A" School Planning Concepts			
Loss rates	.32	-.09	.26
"A" School attrition	.58	-.43	.45
Non-school gains	.44	-.42	.56
Inventory	-.17	.53	-.12
"A" School capacity	-.34	.59	-.23
"A" School inputs	.33	-.30	.47

Descriptions rated:

- Involves movement
- Involves one-way movement
- Has a holding capacity
- Has a flow capacity
- Connects parts of systems
- Has a fill level
- Has a flow rate
- Involves input into a system
- Involves exit out of a system

Table 11 shows that, in general, the element of the hydraulic system chosen to represent each concept is perceived by users as the element most similar to that concept, based on the correlation of ratings for each of the nine descriptive phrases listed in the table. Loss rates are the concept that is least well represented in the Big Picture. The correlation of ratings for loss rate and water through a faucet is only .32, although the correlation's of loss rate to other hydraulic elements are even lower. Water through a faucet seems to be a better representation of "A" School attrition than of loss rates: the correlation of those ratings is .58. Non-school gains are fairly well represented by water in a funnel: the correlation of ratings is .56. Inventory and "A" School capacity are best represented by a tank: the correlation's are .53 and .59. "A" School inputs are best represented by water through a funnel: the correlation of ratings is .47.

The results of the correlation analysis suggest that the mapping of hydraulic elements to "A" School concepts in the Big Picture does have some validity for users. At least within the range of elements considered, there is no other mapping that would be more appropriate. These results must be interpreted within the general finding that the Big Picture was not perceived as very helpful by users, however, and the finding that seeing the Big Picture did not seem to increase users' grasp of the relationships among "A" School planning concepts. The Big Picture seems to be "on the right track" in its

representation of planning concepts, but in its current form it is not especially helpful to new users. Several users suggested that animation would make the picture more useful, and one suggested that the ability to move the elements of the picture up or down to see the positive or negative effects on inventory and "A" School inputs would be helpful.

Observation of Usability Problems

One of the most valuable aspects of the evaluation was the opportunity to observe users from a variety of backgrounds as they worked to complete "A" School planning tasks using the Coach and using SKIPPER without the Coach. An observer took detailed notes in all of the sessions, keeping a record of any problems experienced by the users and any confusions that arose during each task. Based on these observations, we identified many usability problems for SKIPPER, and a number of usability problems for the Coach as well. These observations may be used as the basis for improving the usability of both SKIPPER and the Coach

SKIPPER Usability

The usability problems observed for SKIPPER when users performed "A" School planning tasks without the help of the Coach are summarized by MacMillian and Freeman (1996). The most pervasive problems have to do with finding the right place to enter the data, knowing which data must be entered (out of the many cells in the SKIPPER spreadsheets) in order to produce a plan, and commanding SKIPPER to recommend inputs. There was also considerable confusion in interpreting error messages, and almost universal confusion as to the function of the "Redo calculation" button on the Recommendations sheet.

Coach Usability

The most serious usability problem observed for the Coach was users' confusion about the "dummy" dialog boxes shown in the Select a Community tabs. These boxes were intended to show users how to interact with the SKIPPER dialogs once SKIPPER was started. In the current version of SKIPPER, there is no way to interrupt this dialog sequence for DDE transfer of data, so the Coach user was required to remember how to complete the entire sequence without going back to the Coach. Ideally, we would have liked to query the user for information in the Coach and pass this information to SKIPPER during the start-up dialogs, but this was not possible. The inert nature of the dummy boxes shown in Coach was not obvious to users, who tried to interact with them directly. This problem was exacerbated by the implementation of the dummy boxes with Visual Basic controls, which provided them with what appeared to be functionality (the ability to click on items and scroll through lists). Users believed that they were interacting with live dialog boxes in the Coach, and we often had to intervene to explain that the boxes were for instruction only. Usability problems, their frequency, and the usual user response for Coach are summarized in a report by MacMillian and Freeman (1996). Solutions for each problem are also listed in the report.

Experience with the dialog boxes created a false expectation among users that the entire Coach would be instructional only, and would not allow interaction. When confronted with the Enter "A" School data tab, they hesitated because they did not expect to be able to enter data in the Coach. While all users eventually discovered that it was possible to enter data and that the Coach was interactive, the dummy dialog boxes created an early wrong impression about how to interact with the Coach.

The best solution to this problem is probably to remove the part of the Coach that describes the SKIPPER start-up dialogs, and to improve the design of those boxes in SKIPPER so that no coaching is required. Two problems were observed in the SKIPPER start-up dialogs: (1) users did not know they should use the "Skill Level" button to check to make sure that they were working with communities, not competitive categories; and (2) users were often confused by the list of suggested scenario names ("Optimistic," "Pessimistic," and "Best Guess") presented to them after they had indicated that they wanted to start a new scenario. These names are, in fact, just names, with no data associated with them. This was not at all apparent to users, who expected that selecting one of the names would affect the parameters used by the model and the projections that were made. These problems could be remedied in SKIPPER by making the selection of communities versus competitive categories more apparent to the user, and by eliminating the list of suggested names.

Summary and Recommendations

The Coach accomplished its primary goal—allowing inexperienced SKIPPER users to successfully complete an "A" School plan without the need for in-person assistance. Users were able to complete their first "A" School plans significantly faster using the Coach than using SKIPPER unaided. Users needed to ask an average of almost five questions in order to complete their plans with SKIPPER unaided, while the Coach was used without assistance.[12] Users indicated that they found it significantly easier to complete their first task with Coach than with SKIPPER, and rated the Coach as much easier to use than SKIPPER. Finally, the results showed that the users who were allowed to use the Coach first found it easier to use SKIPPER (i.e., took less time to complete the task with SKIPPER) than users who were not exposed to the Coach first. This latter finding suggests that the Coach is an effective learning tool.

The major difficulty encountered by first-time SKIPPER users was knowing which data to enter, where to enter it, and how to command SKIPPER to make a projection. The Coach seems to have alleviated these problems. The most valuable feature of the Coach was its "wizard" capability that elicited the minimal essential data from users and returned a display of recommended "A" School inputs and projected inventory.

---

[12] Observers routinely intervened if users became confused about the dummy start-up dialog boxes.

The more "instructional" features of the Coach—the hypertext definitions of terms and the Big Picture overview of "A" School concepts—do not appear to have been especially helpful to users. We found no differences in "A" School concept understanding, as measured by a questionnaire, among SKIPPER users and Coach users with or without the Big Picture. The written instructions provided to all users at the beginning of the evaluation may have preempted some of the learning that might have taken place with the Coach, however. The least useful feature of the Coach was its instruction on how to interact with SKIPPER's start-up dialogs, which proved more confusing than helpful to users.

We strongly recommend that some form of step-by-step procedural instruction or wizard capability be provided as part of SKIPPER to help inexperienced users perform common tasks. The current Coach implementation—a separate application running in parallel with SKIPPER with DDE transfer of data—is probably not the optimal way to provide that assistance, however. SKIPPER Version 2 running under Windows 3.1 uses a large proportion of available system resources, making it difficult to run another application in parallel. With almost all system resources occupied, the DDE transfer of data between Coach and SKIPPER and vice versa is fragile and slow, with frequent problems and delays. A better implementation might be to provide a Coach-like capability as an integral part of SKIPPER. Inexperienced users could access the Coach, which would give them step-by-step instruction on how to prepare an "A" School plan and provide a wizard capability for entering data and producing a plan.

From a theoretical viewpoint, the findings from the Coach evaluation indicate the value of procedural, versus propositional, assistance to inexperienced users of a system. Users seemed to value being told "what to do" considerably more than they valued being told "why to do it" by the Coach. This preference for action over instruction carried over into the visual metaphor presented in the Big Picture. Although the visual metaphor used seemed to match users' conceptual mental models of manpower planning fairly well (as indicated by a correlation analysis of the different features), users did not find this visual metaphor very helpful. Perhaps this is because the metaphor supports a structural mental model (how the SKIPPER manpower-planning model works) rather than a functional mental model (how to use SKIPPER). A visual metaphor that represented the actions to be taken to develop an "A" School plan, rather than representing "A" School planning concepts, might have been considerably more helpful to users.

References

Carroll, J. M., Mack, R. L., & Kellogg, W. A. (1990). Interface metaphors and user interface design. In M. Helander (Ed.) Handbook of Human-Computer Interaction. Amsterdam: North Holland.

Carroll, J. M. & Olson, J. R. (1990) Mental models in human-computer interaction. In M. Helander (Ed.) *Handbook of Human-Computer Interaction*. Amsterdam: North Holland.

Dent-Read, C., Klein, G., & Eggleston, R. (1994). Metaphor in visual displays designed to guide action. In *Metaphor and Symbolic Activity*. Norwood, NJ; Lawrence Erlbaum Associates, p. 211-232.

Johnson-Laird, P. N. (1983). *Mental Models*. Cambridge, MA: Harvard University Press.

Kosslyn, S. M. (1989). Seeing and imagining in the cerebral hemispheres: A computational approach. In A. Collins, & E. E. Smith (Eds.) *Readings in Cognitive Science*. San Mateo, CA: Morgan Kaufman Publishers, Inc.

MacMillian, J., & Freeman, B. (1996). *SKIPPER coach evaluation: Final technical report*. Document prepared for Navy Personnel Research and Development Center (Contract Number N00244-95-D-0281, BBN Report No. 8152), Cambridge, MA: BBN Corporation.

MacMillian, J., Getty. D. J., Tatum, B. C., & Ropp, G. A. (1998). *A technique for assessing the congruence between visual metaphors and mental models* (NPRDC-TR-98-4). San Diego, CA: Navy Personnel Research and Development Center.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human-Computer Interaction*. Wokingham, UK: Addison-Wesley Publishing Company.

Pylyshyn, Z.W. (1988). The imagery debate: Analogue media versus tacit knowledge. In A. Collins, & E. E. Smith (Eds.) *Readings in Cognitive Science*. San Mateo, CA: Morgan Kaufman Publishers, Inc.

Rogers, Y., Rutherford, A., & Bibby, P. (1992). (Eds.) *Models in the Mind: Theory, Perspective, and Application*. London: Academic Press.

Rosenthal, R. & Rosnow, R. L. (1984). *Essentials of Behavioral Research: Methods and Data Analysis*. New York: McGraw Hill.

Distribution List

AISTA (PERI II), Army Research Institute
Assistant Deputy Chief of Naval Operations (N1B)
Assistant Secretary of the Navy (Manpower and Reserve Affairs)
Assistant Secretary of the Navy (Manpower and Reserve Affairs) (OASN) (M&RA)
Center for Naval Analyses (Dr. John J. Cymrot)
Chief of Naval Personnel (PERS-00H)
Commander in Chief, U.S. Atlantic Fleet (N1)
Commander in Chief, U.S. Pacific Fleet (N1)
Commander, Space and Naval Warfare Systems Command (SPARWAR-00)
Commanding Officer, Naval Air Warfare Center (Technical Library) (Code 72) (5)
Commanding Officer, Naval Air Warfare Center, Training Systems Division, Orlando
Commanding Officer, Space & Naval Warfare Systems Center, San Diego
Commanding Officer Space & Naval Warfare Systems Center, San Diego (Code D0141)
Commanding Officer, Navy Air Warfare Center (NAWC-TSD)
Commanding Officer, Sea-Based Weapons and Advanced Tactics School, Pacific
Deputy Assistant Secretary of the Navy (Manpower and Reserve Affairs)
Deputy Chief of Naval Operations (M&P) (N1)
Deputy Chief of Naval Research (Code 01) (Dr. Fred Saalfield)
Deputy Under Secretary of the Navy, A&T (R&E)
Director of Education (N2)
Director of Research, U.S. Naval Academy
Director of Training and Education Division (N71)
Director, Army Research Institute, Alexandria, VA (PERI-ZT)
Director, Cognitive and Neural Science and Technology Division (Code 342) (Dr. Willard S. Vaughn, Jr)
Director, Defense Activity for Nontraditional Education Support, Pensacola, FL
Director, Defense Personnel Security Research and Education Center
Director, Enlisted Assignment Division (PERS-40)
Director, Manpower Personnel and Training Info (N16)
Director, Navy Reserve Information Systems Office (NRISO) (Robert A. Duley)
Enlisted Plans and Career Management Division (PERS-22)
Head, Human Factors Division (Code 26)
Head, Manpower Personnel Training Branch (N813)
Headquarters Air Education and Training Command (HQ AETC), Randolph Air Force Base, TX
Human Resources Directorate, Technical Library, AL/HR-SDKL Brooks AFB, TX
Naval Postgraduate School
Office of Naval Research (Code 34), (Dr Anna Johnson-Winegar)
Office of the Director, Test & Evaluation and Technology Requirements (N091)
Office of Training Technology (OTT) (N75)

Distribution List (Continued)

Pentagon Library
Program Officer (Code 342) (Dr. Steele)
Program Officer (Code 342) (Dr. Susan Chipman)
Program Officer (Code 342) (Dr. Terry Allard)
Program Officer (Code 342) (Jan Dickieson)
Public Affairs Office (PERS-05)
Defense Technical Information Center (DTIC) (4)