On The Potential Value and Limitations of Emphasis Change and Other Exploration-Enhancing Training Methods

Eldad Yechiam, Ido Erev and Daniel Gopher

The Max Wertheimer Minerva Center for Cognitive Research

Faculty of Industrial Engineering and Management

Technion – Isreael Institute of Technology

Mail should be addressed to the first author:

Eldad Yechiam, Faculty of Industrial Engineering and Management, Technion, Haifa 32000, Israel. Email: yeldad@tx.technion.ac.il

Appeared in Journal of Experimental Psychology: Applied, 7 (4), 277-285.

Journal Homepage: http://www.apa.org/journals/xap.html

Abstract:

Two experiments are presented that examine the efficiency of training methods that force trainees to explore the possible strategy space. Both experiments employed a two-dimensional search task. Experiment 1 studies a method that enforces exploration by preventing repetitions within short sequences. It shows that the effect of this "enhanced exploration" method in the abstract setting is similar to the effect of "emphasis change" training (see Gopher, 1993) in high cognitive workload tasks. Specifically, the method led to low initial performance, and improved long-term and transfer performance. These effects appeared only in strategic spaces where intuitive exploration converges to a local optimum. Experiment 2 compares the effect of exploration enhancement with the effect of guided instruction. The results of the experiments, which are captured well by a generalization of Erev and Gopher's (1999) model of the conditions for the success of emphasis change training, shed light on the limitations and potential of exploration-enhancing training methods.

Key words:

Exploration training, guided learning, complex skills, emphasis change, strategy space, reinforcement learning, cognitive strategy, learning model.

On the potential value and limitations of emphasis change and other exploration-enhancing training methods

The basic idea of the Emphasis Change (EC) training method, proposed by Gopher, Weil and Siegel (1989), is that repeated changes in the priorities of task elements during practice can improve performance and lead to higher level skills. The original demonstrations of this method's value focused on performance of two concurrent motor tasks (time sharing experiments). Gopher and his colleagues (Brickner & Gopher, 1981; Gopher, 1993, Gopher, Weil & Siegel, 1989) demonstrated that multiple changes in the perceived value of the tasks' sub-components improved final performance. Moreover, it turns out that in many settings this simple training method outperforms more complex traditional methods. For example, a comparison of several training methods of a very difficult video game (the "space fortress") showed that EC training succeeded well above traditional part-task methods (Gopher et al. 1989; Fabiani, Buckley, Gratton, Coles & Donchin, 1989; Gopher, Weil and Bareket, 1994). The method has been extended and refined in aviation tasks (Gopher, Weil, Siegel & Caspi, 1988; Goettl, Anthoby, Derek, Snooks & Shebilske, 1998; Shebilske, Goettl & Regian, 1999;), and has been successfully applied to improve cognitive and motor performance of senior citizens (Kramer, Larish & Strayer, 1995).

In a recent study, Erev and Gopher (1999) proposed a quantitative model of sufficient conditions for the observed advantage of the EC training method. The basic idea of the model is that EC is successful because (1) it facilitates exploration, and (2) in many settings people do not explore enough. The main goal of the current paper is to clarify three nontrivial practical implications of this simple explanation. The first is that EC is only one example of an effective exploration-enhancing training method. Other (and simpler) exploration-enhancing methods can have similar effects. A second implication involves the set of tasks for which exploration-enhancing methods can be useful. Whereas the original EC research focuses on complex motor tasks with nontrivial attention control components, Erev and Gopher's model implies that exploration-enhancing methods can also be successful in facilitating learning in simpler tasks. The third implication concerns the boundary conditions for the advantage of exploration-enhancing methods. These methods are expected to be optimal only when alternative methods (or free training) lead to a local maximum (this concept is explained below).

To clarify these points we chose to focus on the performance of simple choice tasks and simple exploration-enhancing methods. Appendix A presents the computer screen used in the present study. In each trial of the different conditions, the participant is asked to select one of the 400 cells (ordered in a 20x20 matrix). A selection leads to immediate feedback concerning the points earned in that trial. Table 1 shows the payoff rule used in the two conditions studied in Experiment 1. Notice that the optimal strategy in Condition 1 (right matrix) is to select the cell located in row 19 and column 8 (cell (19,8)). This cell promises a payoff of 483 points. The optimal strategy in Condition 2 (left matrix) is a selection of any of the cells in row 1 (1,X) which promise similar payoff (about 400 points). Notice that both matrices have local maxima: cells with a lower value than the maximum, but higher than that of all their neighbors.

One interesting difference between the two conditions presented in Table 1 involves the likelihood of convergence to a local maximum. Previous research of learning in choice tasks suggests that the adaptation process can be approximated as a hill climbing process (see Busemeyer and Myung, 1987). This result suggests that the likelihood of convergence to local maxima is much larger in Condition 1 than in Condition 2. Condition 1 is an example of a task with strong local maxima (SLM), while Condition 2 is an example of a task with weak local maxima (WLM).

The main prediction of Erev and Gopher’s model is that the value of exploration-enhancing training methods is extremely sensitive to the strength of the possible local maxima. Their analysis suggests that all the successful implementations involved natural tasks with strong local maxima. For example, Seagull and Gopher (1994) used an EC-like method to train helicopter pilots to use head-mounted displays. In justifying their choice of training method, they noted that spontaneous adaptation leads pilots to converge to an inefficient strategy in which they try to avoid head movements. This strategy is inefficient because it impairs planning. To solve this difficulty, Seagull and Gopher trained the pilots under conditions that required exploration of different head movement strategies. This exploration led to a discovery of strategies that are much more effective than the strategy discovered by free training. Obviously, though, Seagull and Gopher’s method is unlikely to be as effective in training people to use eyeglasses while driving. It seems that the task of learning to use eyeglasses does not involve a strong local maximum. People do not appear to converge to inefficient strategies.

Experiment 1: The value of enhanced exploration methods in simple tasks

Experiment 1 was designed to clarify and evaluate the three implications of Erev and Gopher’s model presented above. To achieve this goal it examines the effect of a simple exploration-enhancing training method in a SLM and WLM environments. The SLM environment involves choice tasks of the type presented in the right matrix of Table 1. The WLM environment is outlined like the left matrix in Table 1. The specific Enhanced Exploration (EE) method that was examined facilitates exploration by preventing repeated choices in a short interval. Specifically, after each choice of a particular cell, this cell could not be selected again for k trials (k was uniformly distributed between 1 and 5).

<Insert Table 1 here>

Figure 1 shows the predicted performance in six blocks of 100 simulated trials (average payoff above the expected payoff from random selection) in the EE and the control (no-EE) conditions. The predictions were derived based on the updated variant of Erev and Gopher model (See Barron & Erev, 2000). EE was implemented by excluding the possibility of repeated choices in the first 300 trials.

<Insert Figure 1 here>

Comparison of the simulated learning curves reveals the predicted interaction between training, strategic environment and time. EE improves performance only under the SLM condition by about 10%. This long-term advantage follows an initial low performance. In contrast, EE does not affect performance in the WLM environment. Under the assumption that EE has the same effect as EC, a similar interaction between training and payoff matrix is predicted with regard to learning transfer. EC was found to improve transfer from task to task (e.g. Brickner & Gopher, 1981; Gopher, 1993; Kramer et al., 1995).

Method:

<u>Participants:</u>

Eighty subjects (40 men and 40 women) participated in the experiment. The participants were students from the Technion – Israel Institute of Technology or from the University of Haifa. Their age ranged from 19 to 30. They were paid a sum of 30-60 Shekels (7-$15) for taking part in the experiment, depending on their success in the experimental assignment. Participants were divided into eight experimental groups, keeping to a 50%

gender proportion. Due to a technical problem three of the participants completed only half of the transfer stage (there remained 80 subject in the training stage and 77 in the transfer stage).

Apparatus:

The game of button selection was constructed using Visual Basic (Version 6) on a Pentium II computer with a 17-inch screen (800x600 pixels). The game was presented on screen as a matrix of 20x20 buttons in a square frame (see Appendix A). The size of each button was 1 cubic cm. A counter above the buttons displayed players' accumulated sum of points. Assuming a 50 cm sitting distance from the screen, the matrix's height and width were 26.6 visual degrees. Button clicking was performed using a standard computer mouse. Upon pressing a button with the mouse, the image of the button changed to a "pressed" button for the duration of the mouse click (a standard VB button operation).

Payoffs were given one second after clicking the button. The number of points earned in each click was displayed on the button itself for the duration of one game round. The exact number was determined according to the respective payoff matrix (see Appendix B) with a very small noise factor (an integer value randomly drawn from a normal distribution with average 0, SD of 1, truncated beyond 2, -2).

Design:

The experimental task was similar to the simulations summarized in Figure 1 with the following differences. First, in addition to the 600 fixed trials, participants could play up to 300 additional trials to prevent anticipatory “end of game” effect. The exact number of trials for each subject was selected randomly in the following manner: in every round starting from round 601 there was a 1% probability that the task would end. As a result the average length of the game was about 700 trials. In addition, the game included a transfer task of 400 to 500 rounds with a changed value array. The structure of the environment in the transfer task for each condition was similar to the training environment but with a horizontal shift of 1 or 8 cells for the entire set of values. The transfer task was added to test the qualitative prediction of the model regarding learning transfer.

The experiment was built in a 2x2x2 factorial design, with training condition (EE vs. no EE), payoff matrix condition (WLM vs. SLM), and transfer conditions (similar vs. non-similar) as between subject factors. The dependant variable in the experiment was the accumulated number of points gained over time.

Procedure:

The experiment was run in a single session and comprised two stages, training and transfer. The total duration of the experiment varied between participants but in no case exceeded an hour.

The Training Stage included on average 700 trials of the button selection game. The points gained upon clicking buttons were identical to the ones used in the simulation described above. All players were given the following instructions for this stage: “You will be presented with a matrix with many buttons. When you click a button, a number of points will be displayed. The aim of the game is to accumulate the maximum number of points. The accumulating sum is presented in the box at the top of the screen. The final accumulated sum of points determines your bonus: Each 10,000 points equal 1 NIS (about $.25). The game will end after a certain number of button presses regardless of the speed of response.” Players were not told if the value of each button would remain fixed throughout the game. Instructions were read from the screen and later repeated orally.

Enhanced Exploration training was administered, as in the simulations, by temporarily disabling any (non-disabled) pressed button in the 1-5 trials following its being pressed. It was possible to detect that a button had been disabled only upon pressing it again. A disabled button did not change its image to a "pressed" one, nor did it display payoffs. EE was administrated in the first 300 trials. At this point players received a message stating: “As of this time buttons will not be blocked.” The EE group was instructed, accordingly, that at some points certain buttons would be disabled (although they were not told for how long), and that this was part of the game.

Transfer comprised 400 to 500 trials of the same game with a changed value array. The change constituted a horizontal shift of the entire lattice of either 1 or 8 cells (50% of the participants performed each version). Cells that went over the edge as a result of this shift were wrapped around to the other edge of the display. The instructions for this stage were “in this second part of the game, the instructions are the same as in the first part.”

Results:

Figure 2 presents the learning curves for the experimental groups’ performance in the training and transfer stages. Performance is presented in percentages of the number of points between the lowest and highest row maxima (403.8 to 481 in the SLM environment and 307.4 to 403.8 in the WLM environment). As can be seen, in all tasks (training and transfer) and under all conditions there was learning throughout the game. The difference in scores between the average of the first and last 200 rounds of the training stage as well as of the transfer stage were all significant, except for the transfer in the WLM environment (in SLM training: $t(39)= 10.3$; $p<.001$; in SLM transfer: $t(36)= 3.9$; $p<.001$; in WLM training: $t(39)= 6.0$; $p<.001$). It seems that a ceiling effect was reached at that point.

A comparison of the two matrixes indicates that both EE and no-EE training was better in the WLM condition. This is in line with the general assumption that in this environment intuitive exploration did not converge in a local optimum. In the WLM environment the optimal row-maxima was reached after approximately 200 rounds. In the SLM environment after 600 rounds players still advanced only 60% of the distance between the lowest peak and the optimal peak.

Examination of the two training groups' exploration span shows the predicted effect of EE training on increased search of different cells in the matrix. The average number of cells searched during the EE administration period of the first 300 rounds was 98 cells in the EE group compared to 71 in the no-EE group (about 40% difference). This difference was statistically significant ($T(79)=2.47$, $p<.05$). This indicates that while both groups pressed the same button a number of times, the tendency of the EE group to do so was smaller.

<Insert Figure 2 here>

Training phase. Examination of the effect of EE in training supports the main predictions of the model. EE had an initial negative effect on performance, but a positive long-term effect on both training and transfer performance under the SLM condition.

Specifically, in the SLM environment in the first 300 trials of training (the EE administration phase), players in the EE group were below the lowest maximum in the matrix whereas the average of the no-EE group was already 10% between the lowest and highest maxima. In contrast, the average of the EE group in the final 100 trial block was 66% compared to 50% for the no-EE group. This constitutes about a 30% relative advantage for the EE group. In the WLM environment, the differences in the first half of the task were smaller (3% advantage to the no-EE group), both groups having reached the 90% score. In the final block of training both groups in this condition reached the highest peak, and the differences between the training groups became even smaller (with a 2% advantage to the EE group).

To test the significance of the interaction during training the average percentage increments between the first and last halves of the tasks (300 rounds) in the different matrix classes (SLM vs. WLM), and training groups (EE vs. no-EE) were submitted to a three-way analysis of variance (MANOVA) with gender as a blocking variable. The results indicate a main effect of matrix class ($F(1,79)=70$; $p<.01$; MSE=.06), the SLM performance having a higher increment between the first and second half of the task. As for the training condition, the results indicate a main effect of EE training ($F(1,79)=20.85$; $p<.01$; MSE=.07), and in addition, the predicted interaction of EE and matrix class ($F(1,79)=9.05$; $p<.01$; MSE=.09). The Estimate Analysis of the latter interaction shows a significant advantage to the EE group in the SLM condition (T(1,79)= 3.01; $p<.01$).

In addition, there was a significant interaction between EE training and gender ($F(1,79)=3.95$; $p<.1$; MSE=.09). This interaction is depicted in Figure 3. Analysis of the second half of training shows that in the SLM environment women performed less well than men in the no-EE group. Their performance was 34% compared to the average male performance of 56%. In the EE group, the result of the average female subject rose to about 66% while the results of the average male subject remained closely the same. Thus, EE improved the performance of the average female subject.

A separate analysis of the *first half* of training shows a significant negative effect of EE training ($F(1,79)=8.68$; $p<.01$; MSE=.07) in both environments. This is consistent with the predicted initial negative effect of EE, although here there was no interaction between training and matrix type, indicating that EE reduces performance under both conditions. The same analysis in the second half of training shows no significant effect of EE training. However, a post-hock analysis of the last 200 rounds shows a marginally significant positive effect of EE training ($F(1,79)=3.41$; $p<.1$; MSE=.09).

<Insert Figure 3 here>

Transfer phase. Analysis of the average performance in the transfer test as compared to the initial performance scores reveals the matrix type by training method interaction predicted by the model ($F(1,73)=6.35$; $p<.05$; $MSE=.09$). Namely, under the SLM environment, EE training led to better long-term transfer performance. Specifically, the 30% relative difference between the groups, which was observed at the end of training in this condition, increased to 35%. The extent of the transfer task's similarity to the original task (1 vs. 8 displacements) has no effect on this pattern.

Analysis of the first half (200 rounds) of the transfer session shows no effect of EE training. In contrast, analysis of the second half of the task shows a positive main effect of EE ($F(1,73)=4.05$; $p<.05$; $MSE=.08$). In addition, the interaction of EE and matrix type was marginally significant ($F(1,73)=2.82$; $p<.1$; $MSE=.10$).

In this analysis the main effect for gender was significant ($F(1,73)=10.16$; $p<.01$; $MSE=.07$), with the average male players getting the higher scores. In addition, Figure 3 shows that as in training, EE had a stronger beneficial effect on female players in the transfer session. Specifically, in the no-EE group the percentage difference between genders in the last block of transfer was 35% while in the EE group it was only 15%. However, this training by gender effect was not significant.

Discussion:

The present results show that the value of exploration-enhancing techniques is dependent on the underlying strategic space. The EE manipulation increased the efficiency of subjects' exploration, but only in the strong local maximum environment, where intuitive search converged to local maxima.

These results lead to some interesting observations about the learning pattern under a simple exploration-enhancing method. It appears that in the SLM environment the EE manipulation had very similar effects to the reviewed effects of EC in high cognitive-effort tasks. Specifically, EE-trained performers had, on the average, lower initial performance levels, higher long-term performance, and even higher transfer scores. In addition, female participants had, on the average, lower performance levels under the intuitive exploration condition (which could be attributed either to their lack of interest in the computer task, or to their tendency to prefer non-risky solutions to risky ones, see Powell & Ansic, 1997), but under EE their average performance was improved. Thus, like EC, EE helped individuals with a relatively weak strategic exploration disposition.

Experiment 2: Guided Instruction

Experiment 1 shows that EE can have a positive effect. Yet, in the current setting a similar or larger positive effect can also be obtained by directing the participants toward the global maximum. Indeed many studies demonstrate the value of guided instructions (see e.g., Fredriksen & White, 1989; Fredriksen, Weaver, Warren et al., 1983; Carrol & Kay, 1988; Carrol & Carrithers, 1984; Singley & Carrol, 1996). For example, Carrol and Kay (1988) developed a training program called “Scenario Machine” for a stand-alone word processor, which uses instructions to direct users to specific novice-learnable options.

To facilitate comparison between guided instruction and EE in the current setting, we ran a replication of the SLM condition in Experiment 1 in which participants were instructed to focus on the lower part of the matrix. This part contains the global maximum while avoiding many of the low-value local maxima. Note that this replication of SLM with guided instruction does not encourage exploration; to the contrary, it limits it.

Method:

Participants:

20 subjects (10 men and 10 women) participated in the experiment. The participants were students from the Technion – Israel Institute of Technology or from the University of Haifa. Their age ranged from 19 to 25. As in Experiment 1, they were paid a sum of 30-60 Shekels (7-$15) for their participation, proportional to their success. Participants were randomly assigned to three experimental groups. Each group had an equal number of male and female participants.

Design

The design of Experiment 2 was similar to the design of Experiment 1, but for two differences. First, only the SLM matrix was used. Second, only the far transfer (a horizontal shift of 8 cells) was studied. The dependant variable in the experiment was the accumulated number of points obtained by players over time. The experiment followed a between subject design, with the training condition (instructions, and "control" - no instructions) as the between subject factor.

Procedure:

As in experiment 1, the game had two stages, training and transfer. The objective of players was to accumulate the highest number of points in repeated button presses. *Guided instruction* was administered in the experimental groups by asking players to concentrate on the lower half of the game form (see Appendix A). In contrast to the EE method, no actual blockage of buttons was applied here to assure compliance. Players were simply told that it was recommended that they pressed buttons in the lower half of the form. They were told that it was also possible for them to click the top half. The control group was given no instructions regarding which part of the space they should focus on.

Results:

The results of players' performance under the instruction condition as compared with players who received no special focusing instructions are presented in Figure 4. The figure shows that the instructions had a positive effect on short and long-term performance. Specifically, in the first half of training the instruction group had an average score of 32% (measured as percentages between the lowest and highest row maxima) compared to the control group's average score of 21%. In the second half of training the average of the instruction group was 69% compared to 44% for the control group. This difference did not disappear in the transfer session. The statistical analysis test (performed as in Experiment 1) showed a significant difference only in the second half of the transfer session ($F(1,19)=3.53$; $p<.1$; $MSE=.14$), possibly due to the small number of subjects.

In addition, the gender differences in the instruction group are similar to the EE pattern in Experiment 1. In the control group (as in the identical control group of Experiment 1) male participants had higher performance levels in the final block of training (54% compared to 36%). Vice versa, in the instruction group females had relatively higher performance scores (67% compared to 62% for the average male subject).

<Insert Figure 4 here>

Discussion:

The results reveal that the effect of the guided instruction manipulation was practically identical to the effect of the EE manipulation. These findings suggest that there is nothing "magical" about the success of EE. Other methods to move the trainee away from the local maximum are equally effective. Whereas these findings are not surprising, they appear to be inconsistent with the EC research. Fabiani et al. (1989) found that EC leads to better transfer than guided search (see also Gopher et al., 1994). One explanation of this apparent difference involves the possibility of larger individual differences. In the current setting all of the participants had the exact same payoff matrix. Thus, it was easy to give them instructions that moved them toward maximization. In the complex motor tasks used to evaluate EC, the optimal instructions might have been participant specific. Thus, guided instruction might have led some of the participants in the wrong direction. A second explanation is that the reward structure of the transfer task in Fabiani et al., (1989) changed more significantly compared to the current transfer condition. Accordingly, EC trained individuals were more inclined to explore in the new condition.

Guided instruction requires knowledge of the task, the capabilities of the trainee, and stability of reward structure. Under such conditions (which were maintained in the present experiment) guided instruction and EE led to similar levels of performance.

General discussion

The present results provide an optimistic explanation for the success of exploration-enhancing training methods. The optimistic nature of the explanation is reflected in three of its properties. First, the explanation is extremely general. It can address methods in very different domains including high workload environments, software usage, psychomotor skills, and teaching programs. Moreover, the explanation includes a description of the expected limitations of these methods. Thirdly, as opposed to methods that limit exploration like guided instruction, exploration-enhancing methods are expected to be robust to individual differences, limited knowledge of the trainer on the task structure, and changes in reward structure. Efficient usage of these methods does not depend on deep understanding of the relevant domain. Both EC and EE require only the basic delineation of the strategy space. They can lead individuals to find the best match between their capabilities and task structure by encouraging them to explore the task strategy space.

Exploration-enhancing methods vs. spontaneous acquisition. Experiment 1 demonstrates that even a simple exploration-enhancing method can lead to better performance than spontaneous acquisition. The value of the EE method appears to be a result of the tendency of human trainees to converge to sub-optimal solutions - local maxima (Gopher, Weil & Siegel, 1989). Exploration enhancement methods are useful when they eliminate this source of sub-optimality.

The advantage of the EE manipulation in the transfer conditions indicates that the willingness to explore can be generalized to a new situation in which the reward structure of the strategy space has been changed. This finding is consistent with the increased transfer capabilities of subjects trained with the EC method (Fabiani et al., 1989; Gopher et al., 1994). It thus appears that the exploration experience has a value by itself. It teaches trainees a new approach to the development of expertise in complex situations. This knowledge is very general and relevant above and beyond specific application context.

Necessary conditions of exploration-enhancing methods. Experiment 1 shows that the EE method is effective only in environments in which spontaneous strategic search ends in a local optimum. Here, decision-makers did not explore the environment efficiently, because of lack of knowledge and of willingness to pay the cost of exploration. They were attracted by sub-optimal solutions and neglected the optimal solutions. We assume, however, that this environment represents many of the daily practiced tasks, where performers do not explore the task sufficiently. For instance, in most computer applications the same goals can be achieved by the use of pull down menus or keyboard shortcuts. In a recent article Temple and Schmidt (1999) showed that users adopt one of these styles, and do not explore the value of the other in different tasks and changed conditions. This conservative strategic choice is easy to explain in the framework of the present study. Strategic exploration requires effort and is usually accompanied by an initial period of performance instability. Thus, the currently held style can be considered as a local maximum strategy.

Whereas the current experimental research focused only on the above limitation of enhanced-exploration training methods, the theoretical framework outlined here implies that there are at least three additional limitations. These limitations are more obvious but can be as important. The most obvious limitation involves situations where some of the payoffs in the matrix are extremely poor. It is easy to see that in these situations trainers will not want to use EC or other enhanced-exploration methods. An example of this is on-the-job training in which errors may result in dramatic consequences, such as training pilots in the air. Such costs are, of course, removed, when training is conducted in a simulator (e.g. Gopher et al., 1989; Seagull & Gopher, 1994; Shebilske et al., 1999).[1] Another, and perhaps more common example, involves possible interdependency between strategies. Complex tasks might not include actual losses, but rather negative consequences in terms of the effect of the use of one strategy on the efficiency of using another strategy at a later stage (Wood & Locke, 1990). For instance, Adams, Gopher and Lintern (1977) had subjects practice a motor skill with a visual feedback and then replaced it with a proprioceptive feedback. It was shown that the initial performance in the visual-feedback condition inhibited later performance under the proprioceptive feedback condition.

A second limitation involves the necessary length of training. Both EC and EE require that trainees have the opportunity to explore and examine many strategies. This requirement is relevant to both training methods although it is more critical for EE, because it is based on less knowledge on the underlying structure of the task. Exploration implies that trainees are given an opportunity to examine alternative response strategies and err, which may result in initially lower performance levels. However, the freedom to explore may not exist in all circumstances, thus limiting the use of enhanced-exploration methodologies.

Another set of situations in which enhanced exploration is not expected to be efficient includes situations where people tend to explore too often. Under the current learning model over-exploration is expected when the relevant payoffs are noisy. Thus, for example, an EC method will not be very useful in training people to stop gambling at a Casino (see, Haruvy, Erev & Sonsino, in press).

Exploration-enhancing methods vs. guided instruction. In Experiment 2 we compared the effects of the traditional guided-instruction approach with the exploration-enhancing method used in Experiment 1. The guided-instruction approach has been empirically formulated and evaluated in the behavioral shaping literature (see review in Skinner, 1968). In shaping the trainer reinforces any action that is congruent with the learned strategy (such as movement towards an object which should be manipulated by the trainee). Modern, cognitive-oriented training methods, similarly direct subjects towards strategic solutions, but these solutions need not be overt behavioral responses. For example, in an international collaboration investigating learning strategies in a real-world task (training of the Space Fortress game, a simulation of a complex and dynamic airspace environment, requiring control of a spaceship by using a joystick, while avoiding hostile mines, and shooting at an immobile fortress), Fredriksen and White (1989) used both covert and overt guided instruction. Overtly, they used the advice of experts to advance the choice of certain strategies of task performance. Covertly, they split the task into several scenarios, based upon behavioral task decomposition. Each of the scenarios (part tasks) teaches subjects a particular principle of the entire task.

The approach outlined in the present study suggests that the main difference between these methods and exploration-enhancing methods is related to their assumptions about the prior knowledge of the trainer concerning the detailed structure of the task and the ability of each trainee. In shaping for example, the trainer is assumed to know the optimal strategy in advance. The trainer's goal is to move the learner to that point. Similarly, the part-task approach presented by Fredriksen and White (1989) assumes that the optimal solution is a combination of several strategic responses. These part-tasks are organized in a hierarchy, in order to expedite skill acquisition.

Note also that guided instruction is extremely sensitive to a change in the reward structure, which is typical to many complex dynamic tasks. In the study of Fredriksen and White (1989) the performance of subjects was considerably impaired when they were transferred to a changed version of the original task (this effect was also demonstrated by Fabiani et al., 1989).

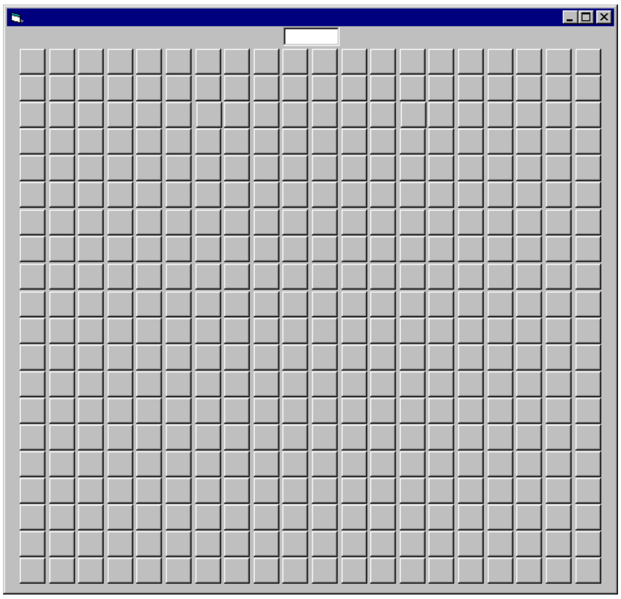
In contrast, exploration-enhancing methods require a much more global delineation of task requirements and strategy space. EE training simply facilitates a self-directed search in any direction, and leaves it up to the trainee to explore possible solutions of the strategic problem. EC demands that the trainee explore a designated set of alternatives. Thus, in EC the strategy space must be organized according to relevant criteria (for example, the ratio of attention invested in two tasks), but no assumptions must be made regarding the exact payoff space. This observation suggests that exploration-enhancing training methods can serve as robust substitutes for shaping in cases where the trainer has limited information about a task. Otherwise, if the trainer has sufficient knowledge about the task and the conditions are stable, the guided exploration afforded by shaping may be relatively more efficient.

For example, in teaching typing it is known that for most people touch-typing is more efficient than visually guided typing. Therefore, guided instruction towards strategies that advance touch-typing appears to be useful. By comparison, Gopher and his colleagues (1989), in training the Space-Fortress game did not know which combination of possible strategies was optimal for each player. It heavily depended on the individual's performance and attention control skills. In this paradigm, with no predetermined single optimal strategy, the EC method had an advantage over other methods that directed subjects towards predetermined behavioral choices based on general assumptions about the requirements of the game (Fabiani et al., 1989).

The variability of training approach (Schmidt & Bjork, 1992; Catalano and Kleiner, 1984; Nitch, 1977) appears to be a hybrid of simple EE and EC. This method asserts that trainees should experience a wide range of conditions, which (unlike EE), are pre-determined by the trainer. However (unlike EC), this method does not base its suggested varied conditions on the structure of the task. Its assumption is that any change of conditions is beneficial for training. In this respect, the method seems to be less useful than EC, as it does not allow trainees to experience a "taste" of the entire variety of possible task conditions. On the other hand, its apparent advantage is that it requires less knowledge than EC. The results of studies using this approach show a pattern that is closely similar to the observations made in Erev and Gopher (1999) and in the present study. Namely, they show that under variable training, trainees' initial performance levels decrease, but long term and/or transfer performance are much improved.

Appendix A

The Game Form



The form used in the game of button selection (Experiments 1 and 2). The form consists of 400 buttons and a label for the accumulating profit. It is presented in about 1:3 of its real size.

Appendix B

Algorithm for the Reward Structures

SLM *environments:*

Each cell offered a payoff (a random number drawn from a normal distribution with the cell mean and SD of 3). The mean of cell *ij*, referred to as *V*(*i*,*j*), was determined by the following algorithm:

1. *V*(*i*,*j*) is assumed to be randomly drawn from a normal distribution with a mean *Mi* and standard deviation *Si*;
2. *Mi* decreases with *i*: Mi = 400 - (5 * i)
3. *Si* increases with *i*: Si = 5 *i.
4. The highest value in row *j* is put in a random position *k*.
5. Each random position *k* is moved so that: k(j+1)=k(j)+b where *b* is randomly determined from the set {-1,0,1}.
6. Other payoffs are arranged in both sides of *k* in a decreasing order.

(The last three rules arrange values both vertically and horizontally. Horizontally, there is an increasing path towards line maxima. Vertically there is a path from one line's maximum to another's, though this one includes local maxima)

WLM *environments:*

Each cell offered a payoff according to rules 1-3 of the SLM algorithm except that rule 3 was changed so that *Si* is fixed: Si = 5.

References

Adams, J.A., Gopher, D., & Lintern, G. (1977). Effects of visual and proprioceptive feedback on motor learning. *Journal of Motor Behavior*, *9*, 11-22.

Barron, G., & Erev, I. (2000). On the relationship between decisions in one-shot and repeated tasks: Experimental results and the possibility of general models. Submitted to Econometrica.

Brickner, M., & Gopher, D. (1981). *Improving time-sharing performance by enhancing voluntary control on processing resources*. Technion-III, Faculty of Industrial engineering and Management, Research Center for Work Safety and Human Engineering, HEIS-81-3, Febr. See also in Gopher, 1993.

Busemeyer, J.R., & Myung, I.J. (1987). Resource allocation decision making in an uncertain environment. *Acta Psychologica*, *66*, 1-19.

Carrol, J.M., & Carrithers, C. (1984). Blocking learner error states in a training-wheels system*, Human Factors*, 26, 377-389.

Carrol, J.M., & Kay, D.S. (1988). Prompting, feedback and error correction in the design of a scenario machine. *International Journal of Man Machine Studies*, 28, 11-27.

Carrol, J.M., & Kellogg, W.A. (1989). Artifact as theory-nexus: hermeneutics meets theory-based design. In K, Bice & C.H. Lewis (Eds.), *wings for the mind, CHI'89 Conference Proceeding* (pp. 7-14). New York: ACM Press.

Catalano, J.F., & Kleiner, B.M. (1984). Distant transfer in coincident timing as a function of variability of practice. *Perceptual and Motor Skills*, 58, 851-856.

Erev, I., & Gopher, D. A. (1999). Cognitive Game-Theoretic Analysis of Attention Strategies, Ability and Incentives. In D. Gopher & A. Koriat (Eds.), *Attention and Performance, XVIII: Cognitive regulation of performance, Theory & Applications*. Cambridge: MIT Press.

Fabiani, M., Buckely, J., Gratton, G., Coles., M., & Donchin, E. (1989). The learning of complex task performance. *Acta Psychologica*, *71*, 259-300.

Fredriksen, J.R., Weaver, P.A., Warren, B.M., Gillotte, H.P., Rosebery, A.S., Freeman, B., & Goodman, L. (1983). *A componential approach to training reading skills* (Final Report 5295). Cambridge, MA: Bolt Beranek and Newman.

Fredriksen, J.R., & White, B.Y. (1989). An approach to training based upon principle-task decomposition. *Acta Psychologica*, 71, 89-146.

Goettl, B.P., Anthony, M., Derek, K., Snooks, S., Shebilske, W.L. (1998). Emphasis, Gender and Interaction Anxiety in Individual and Dyadic Training Protocols. *Human-System Interaction: The Sky's No Limit. Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, Chicago, Illinois, October 5-9, 1998. The Human Factors and Ergonomics Society, Santa Monica, California, Volume 2.

Gopher, D. (1993). The skill of attention control: Acquisition and Execution of Attention Strategies. In D.E. Meyer & S. Kornblum. *Attention and Performance XIV*, London, England: The MIT Press.

Gopher, D., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors*, 36, 387-405.

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to training of complex skills, *Acta Psychologica*, *71*, 147-179.

Gopher, D., Weil, M., & Siegel, D., Caspi, S. (1988). Fidelity of task structure as a guiding principle in the development of skill trainers based upon complex computer games. In *Proceedings of the 32nd Annual Meeting of the Human Factors Society*, 1266-1270. Aneheim, CA.

Haruvy, E., Erev, I., & Sonsino, D. (In press). The medium prizes paradox: Evidence from a simulated casino. *Risk and Uncertainty*.

Kramer, A.F., Larish, F.J., & Strayer, D.L. (1995). Training for attentional control in dual task settings: A comparison of young and old adults. *Journal of Experimental Psychology: Applied*, *1*, 50-76.

Nitsch, K.E. (1977). *Structuring deconstextualized forms of knowledge.* Unpublished doctoral dissertation, Vanderbilt University, Nashville.

Powell, M., & Ansic, D. (1997), Gender differences in risk behavior in financial decision-making: An experimental analysis. *Journal of Economic Psychology*, 18, 605-628.

Seagull, J., & Gopher, D. (1994). Training head movements in visual scanning: An embedded approach to the development of piloting skills with helmet mounted displays. *Journal of Experimental Psychology: Applied*, *3*, 463-580.

Singley, M.K., & Carrol, J.M. (1996). Synthesis by analysis: five modes of reasoning that guide design. In T.P. Moran & J.M. Carrol (Eds.), *Design Rational: Concepts, techniques and use* (pp. 241-265). Mahwah, NJ: Erlbaum.

Schmidt, K., & Bjork, D. (1992). New conceptualizations of practice: Common principles in three pardigms suggest new concepts for training. *Psychological Science*, 207-217.

Shebilske, W., Goettl, B., & Regian, W. (1999). Individual and Team Protocols for Training Complex Skills in Laboratory and Applied Settings. In D. Gopher & A. Koriat (Eds.), *Attention and Performance, XVIII: Cognitive regulation of performance, Theory & Applications*. Cambridge: MIT Press.

Skinner, B.F. (1968). *The technology of teaching*. New York: Appleton-Century- Crofts.

Temple, J.G., & Schmidt, N.A. (1998). Individual differences in word processing strategies. In R. Hoffman, M.F. Sherrick, et al. (1998). *Viewing psychology as a whole: The integrative science of William N. Dember*. (pp. 205-226). Washington, DC, US: American Psychological Association.

Wood, R.E., & Locke, E.A. (1990). Goal setting and strategy effect on complex tasks. In: B. Straw & L. Cummings (Eds.), *Research in organizational Behavior, Vol. 12*. Greenwich, CT: JAI Press.

Zwick, R. Rapoport, A., King Chung Lo, A., & Muthukrishnan, A.V. (2000). Consumers Search for the Best Option Among Relatively Ranked Alternatives with Backward Solicitation: An Experimental Study. A Working Paper, The Hong Kong University of Science and Technology.

Footnotes page:

[1] Zwick et al. (2000) found an example in which subjects tend to underweight the importance of the cost of search.

Figure 1:

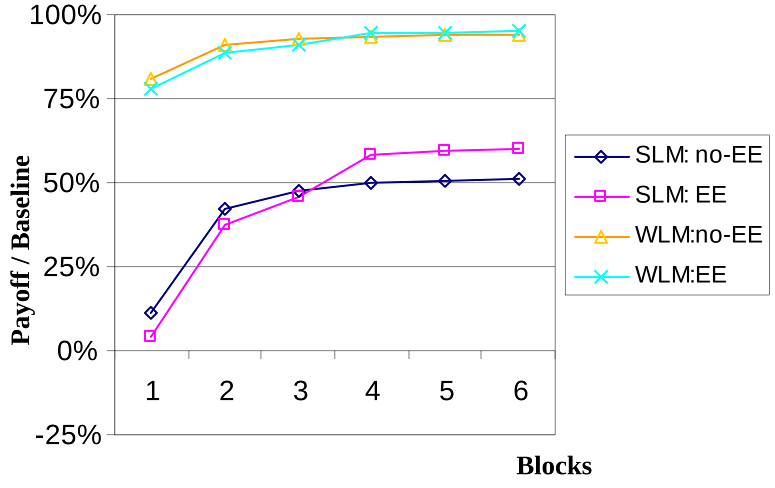
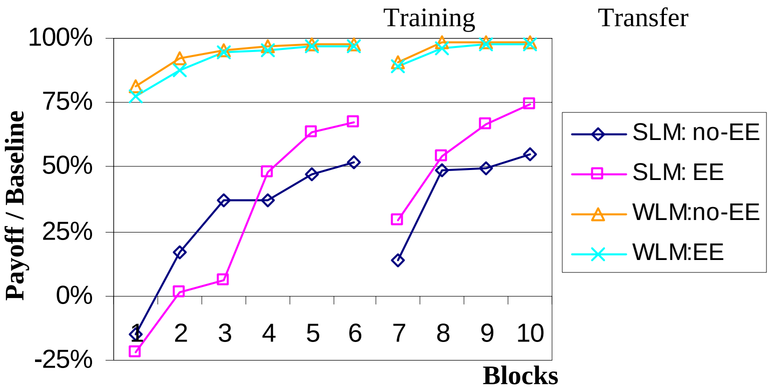

100%
75%
50%
25%
0%
-25%
Payoff / Baseline
1
2
3
4
5
6
Blocks
SLM: no-EE
SLM: EE
WLM:no-EE
WLM:EE

Figure 2:

Performance in the SLM and WLM matrixes



SLM Transfer divided by similarity

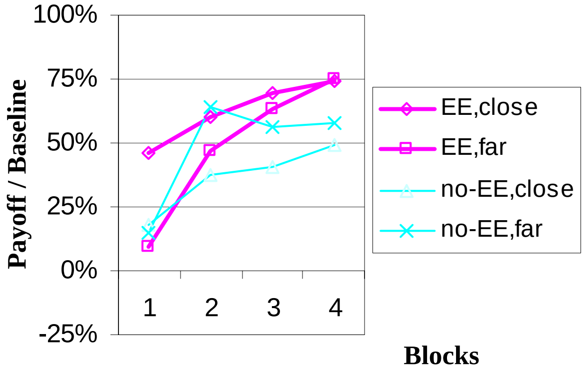
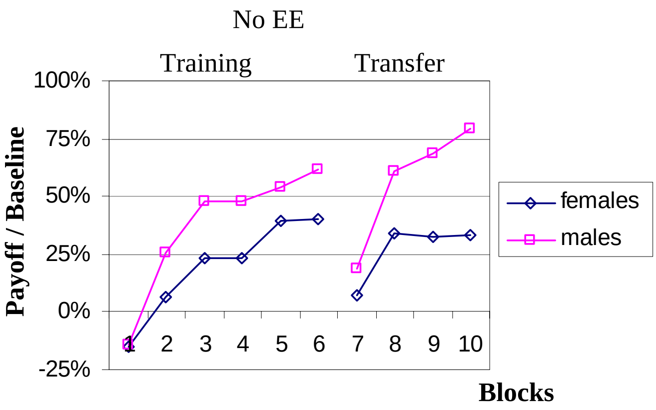


Figure 3:


No EE
Training
Transfer
100%
75%
50%
25%
0%
-25%
Payoff / Baseline
1 2 3 4 5 6 7 8 9 10
Blocks
females
males

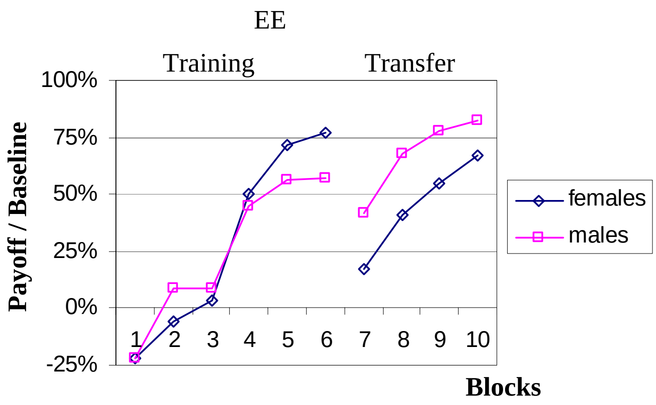
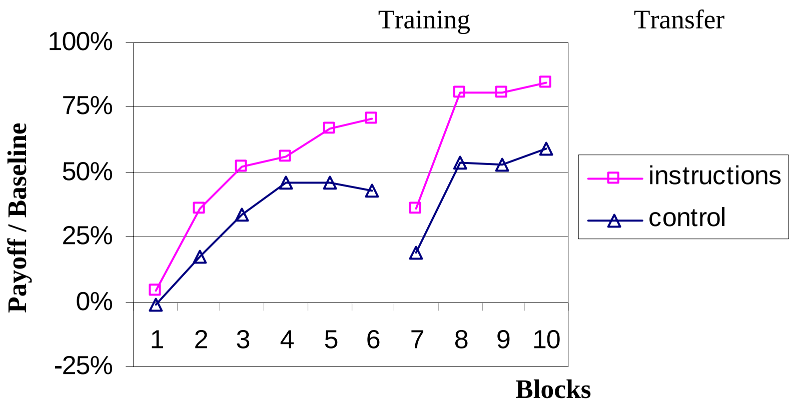

EE
Training
Transfer
100%
75%
50%
25%
0%
-25%
Payoff / Baseline
1 2 3 4 5 6 7 8 9 10
Blocks
females
males

Figure 4:



Training
Transfer
100%
75%
50%
25%
0%
-25%
Payoff / Baseline
1 2 3 4 5 6 7 8 9 10
Blocks
instructions
control

Figure Captions

Figure 1. The prediction of the reinforcement-learning based model for Experiment 1: Average payoff (in payoff relative to baseline) in the different conditions as a function of time (blocks of 100 rounds).

Figure 2. Experiment 1: Average payoff (in payoff relative to baseline) in the different conditions as a function of time (blocks of 100 trials).

Figure 3. Gender effects in the SLM matrix of Experiment 1: Average payoff (in payoff relative to baseline) by time and gender in the EE and no-EE groups.

Figure 4. Experiment 2: Average payoff (in payoff relative to baseline) in the different conditions as a function of time (blocks of 100 trials).

Table 1:

Diagonal Cuts of Matrixes from the Two Payoff Structures: (a) Weak Local Maxima, (b) Strong Local Maxima. In Gray are the Top Fifteen Values of Each Matrix.

(a) *WLM* matrix

No.	Mi	Si	1	2	3	4	5	6	7	8	9	10
1	**395**	**5**	399	398	399	393	393	397	405	388	399	392
2	**390**	**5**	394	391	390	381	388	384	395	388	395	390
3	**380**	**5**	393	393	376	384	380	375	387	385	392	387
4	**375**	**5**	382	388	383	380	382	386	380	376	381	382
5	**370**	**5**	374	375	375	372	376	370	384	376	378	379
6	**365**	**5**	366	366	370	369	369	376	380	368	371	370
7	**360**	**5**	362	366	364	360	367	370	364	367	361	365
8	**355**	**5**	360	359	367	352	362	362	358	356	357	360
9	**350**	**5**	360	349	347	347	361	352	355	353	357	352
10	**345**	**5**	341	355	351	340	350	343	342	349	353	355
11	**340**	**5**	346	340	347	348	343	345	351	343	351	349
12	**335**	**5**	332	334	333	344	335	338	335	332	342	331
13	**330**	**5**	336	335	335	329	327	334	338	341	336	329
14	**325**	**5**	332	338	330	336	331	332	329	329	337	331
15	**320**	**5**	321	323	321	324	322	331	330	321	320	326
16	**315**	**5**	312	316	315	324	318	318	321	311	321	323
17	**310**	**5**	309	315	315	310	312	313	316	313	315	315
18	**305**	**5**	315	316	316	309	311	313	308	306	311	308
19	**300**	**5**	299	311	298	307	311	308	296	304	311	305
20	**295**	**5**	302	301	297	298	298	301	293	302	306	301

(b) *SLM* matrix

Mi	Si	1	2	3	4	5	6	7	8	9	10
395	**5**	392	393	393	395	397	401	403	403	402	399
390	**10**	381	381	387	388	389	392	397	403	403	400
380	**15**	366	374	379	385	388	390	392	397	410	395
375	**20**	375	380	381	385	388	393	397	420	394	389
370	**25**	368	381	381	382	388	389	393	395	407	393
365	**30**	369	373	382	384	394	399	409	413	401	394
360	**35**	348	360	368	371	374	383	423	433	388	381
355	**40**	359	361	370	371	378	397	410	413	436	413
350	**45**	328	344	352	358	382	389	422	435	391	385
345	**45**	305	315	335	341	343	349	375	393	415	377
340	**50**	276	310	314	326	336	345	371	392	396	407
335	**55**	269	281	317	321	333	359	370	454	458	437
330	**60**	258	272	282	306	332	343	385	403	434	451
325	**65**	249	281	299	323	342	375	384	429	435	423
320	**70**	264	306	328	337	339	357	390	440	368	346
315	**75**	301	315	336	340	361	380	433	460	408	365
310	**80**	258	276	324	329	353	364	409	409	364	362
305	**85**	310	333	350	363	376	393	422	391	369	361
300	**90**	282	323	329	402	432	454	463	483	461	440
295	**95**	249	259	307	322	353	391	435	439	426	355