Department of Psychology   Dietrich College of Humanities and Social Sciences

November 2008

Components of Dynamic Skill Acquisition

Frank J. Lee
Carnegie Mellon University

John R. Anderson
Carnegie Mellon University

Michael P. Matessa
Carnegie Mellon University

Components of Dynamic Skill Acquisition

Frank J. Lee
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213
fjl@cmu.edu

John R. Anderson
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213
ja@cmu.edu

Michael P. Matessa
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213
matessa@cmu.edu

Abstract

In this paper, we examine the components of dynamic skill acquisition using a data set collected by Ackerman (1988) with the Kanfer-Ackerman Air-Traffic Controller Task©. Our analysis indicates that subjects are improving in both the strategies they use to solve the task and the speed with which they execute the task. One strategy that subjects develop reduces the number of overt actions required to land a plane. Another strategy that subjects develop enables them to land more planes simultaneously. A satisfactory model of this task must include both an improved strategic component and an improved speed component. The ACT-R theory (Anderson, 1993) is well suited to model these components as it is able to separately learn over trials which strategies are better and how to execute each more efficiently.

Keywords: Dynamic skill acquisition, Cognitive models of problem solving, Strategy learning

Introduction

Thought and action in the real world are embedded in dynamic environments, but past research in problem solving and skill acquisition has focused primarily on tasks in static environments. In static tasks, changes in the problem state can only occur through people's actions, while in dynamic tasks, changes can occur independent of people's actions. An example of a static task is the Tower of Hanoi puzzle, where disks of different sizes must be moved from one peg to another according to simple rules. In this task, the disks move only if a person moves them. Driving, on the other hand, is a dynamic task: between any two consecutive actions in driving (e.g. turning the steering wheel and stepping on the accelerator), changes in the problem state, such as a decrease in fuel or a dog running in front of the car, occur independently of the driver's actions.

We have gained important insights from previous research in static task domains. The legacy of past problem solving research in static task domains includes, the identification of search heuristics in problem solving (Newell & Simon, 1972), the discovery of the differences of novices and experts in problem solving in physics (Chi, Glaser, & Rees, 1982; Larkin, McDermott, Simon, & Simon, 1980) and in programming (Anderson, Corbett, & Conrad, 1984), and the isolation and the quantification of the elements of skill transfer (Singley & Anderson, 1989). However, to completely understand and appreciate the domain of problem solving and skill acquisition, we must extend investigations to *dynamic* tasks.

The Kanfer-Ackerman Air Traffic Controller© [1] (ATC) task is an ideal vehicle for studying dynamic skill acquisition. It simulates dynamic aspects of real air traffic control (e.g., planes lose fuel and weather conditions change), yet is simple enough to be tractable for study. In addition, Ackerman (1994) has collected data from over 350 subjects on the ATC task and has made them available on CD-ROM (Ackerman & Kanfer, 1994) to the Office of Naval Research (ONR). ONR intends to use this task as a testbed to compare a number of cognitive architectures including our own, ACT-R (Anderson, 1993).

Ackerman has extensively analyzed the predictive measures of performance in the ATC task using a battery of psychological tests that measure cognitive, perceptual-motor, and psychomotor ability (Ackerman, 1988, 1990). Ackerman found that while cognitive ability best predicts performance in early trials, psychomotor ability best predicts performance in later trials. [2] We take a different approach to the study of dynamic skill acquisition. Instead of looking in from the outside - that is, instead of using task-external tests to predict individual performance in the ATC task - we propose to go inside and see what subjects are actually doing in order to illuminate the components of dynamic skill acquisition in the ATC task.

In this paper, we use a data set from Ackerman's study (study #6 in the Kanfer-Ackerman CD-ROM, as published in Ackerman, 1988) to examine the cognitive components of dynamic skill acquisition in the ATC task. We will briefly review the ATC task. We will then analyze the role of strategies and speed in the ATC task through correlational and regression analyses of different variables. We argue that, even after taking subjects' increase in motor speed into consideration, their strategy use contributes significantly to performance.

The Air Traffic Controller Task

The ATC task is composed of the following elements displayed on the screen (see Figure 1): (a) 12 hold pattern

---

```
FLT#    TYPE    FUEL    POS.
----    ----    ----    ----
342     DC10    5       3 n          Score  :   380
148     727     6       3 s          Landing Pts:  350  Penalty Pts:   -20
                        3 e          Runways : DRY
692     747     4       3 w          Wind   :   0 - 20 knots from SOUTH
                        2 n          ┌─────────────────────────────────┐
                        2 s          │ Flts in Queue: ......            │
428     prop    * 3     2 e          │         <F1> to accept           │
                        2 w          ├─────────────────────────────────┤
-> 259  727     4       1 n          │                                  │
                        1 s          │       Winds from South           │
840     prop    4       1 e          │                                  │
190     DC10    5       1 w          └─────────────────────────────────┘

n ===========================  s  #1

n ====================  s  #2

w ||||||||||||||||||||||||||||  e  #3

w |||||||||||||||||  e  #4
```

Figure 1. The Air Traffic Controller task. (Note: This figure is a reconstructed representation of the Kanfer-Ackerman ATC task.)

positions, (b) 4 runways, numbered 1 through 4, (c) feedback information on subject's current score and penalty, conditions of the runways, wind direction and speed, (d) a queue stack with planes waiting to enter the hold pattern, and (e) 2 message windows, one for notifying of weather changes (shown) and one for providing feedback on errors (not shown). The 12 hold pattern positions are divided into 3 levels corresponding to altitude, with hold level 3 being the highest and hold level 1 being the lowest.

Six rules govern this task: (1) Planes must land into the wind, (2) Planes can only land from hold level 1, (3) Planes can only move one hold level at a time, but to any open position in that level, (4) Ground conditions and wind speed determine the runway length required by different plane types (747's always require long runways, DC10's can use short runways only when runways are dry or wet, and wind speed is less than 40 knots, 727's can use short runways only when the runways are dry or wind speed is 0-20 knots, and PROP's can always use short runways), (5) Planes with less than 3 minutes of fuel left must be landed immediately, and (6) Only one plane at a time can occupy a runway. A weather change occurs approximately every 30 seconds; planes enter into the queue approximately every 7 seconds. Three principal actions are used in this task: (1) accept planes from the queue into a hold pattern, (2) move planes within the three hold levels, and (3) land planes on a runway. All three actions can be accomplished by using the keys: ↑, ↓, F1, and ↵. The ↑ and ↓ keys move the plane up and down between the different hold positions and levels, and the F1 key accepts the planes from the queue into a holding pattern. The ↵ key can select a plane in the queue, place a selected plane (either from the queue or from a hold position) into an empty hold position, or land a plane on the runway. A subject's cumulative score is computed as follows: a) 50 points for landing a plane, b) minus 100 points for crashing a plane, c) minus 10 points for violating one of the six rules that govern the task.

Method and Analysis

We used data from 58 subjects in Ackerman's (1988) study. We excluded data from 5 subjects who had not completed at least eighteen (10-minute) trials and data from 2 subjects due to an error during their decompression from the Kanfer-Ackerman CD-ROM (Ackerman & Kanfer, 1994).

We used cumulative score (Score) as dependent measure of performance. Figure 2 plots mean score and standard deviation of the 58 subjects across trials.
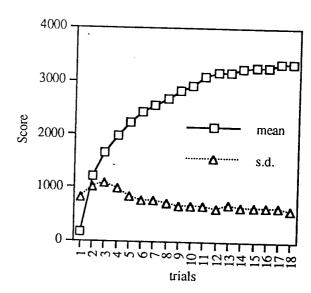

Figure 2: Mean Score and standard deviation.

As can be seen in Figure 2, subjects' cumulative scores grew from almost nothing in the first 10-minute block to an average of over 3000 points in the last block, while the

standard deviation tended to decrease, indicating a reduction in inter-subject variability. The increase in Score closely follows a power function: $f(x) = -2024.04 + 2791.40 * x^{.241}$, with $R^2 = .954$.

To help understand the basis for this improvement, we looked at strategy change over trials. One strategy that many subjects developed, which we call the *hold 1 strategy*, involved bringing planes directly into hold level 1, thereby skipping hold levels 2 and 3. On average, 6 keystrokes (1 ⌐ key to select a plane, 4 ↓ keys to move down to the next level, and 1 ⌐ key to place the plane) are needed to move a plane down one hold level. If we assume that the average number of keystrokes to land a plane from hold level 1 is equal to C, then the average number of keystrokes needed to land a plane from hold levels 1, 2, and 3 are C, C + 6, and C + 12, respectively. By bringing planes directly into the hold level 1 from the queue, subjects eliminate the need to use 6 to 12 additional keystrokes per plane. Using this strategy therefore increased subjects' keystroke efficiency by reducing the number of keystrokes needed to land a plane.

We measured hold 1 strategy as the percentage of planes brought directly from the queue into hold level 1. Figure 3 plots mean hold 1 strategy use and standard deviation for the 58 subjects across trials. As can be seen, hold 1 strategy increases over the first half of the experiment and then asymptotes. However, variability in the hold 1 strategy remains high, indicating that hold 1 strategy is an important source of individual differences.
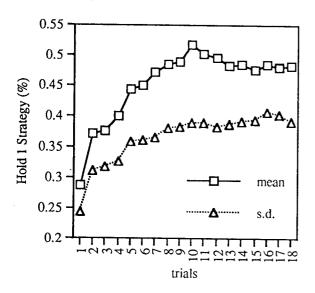

Figure 3: Mean hold 1 strategy and standard deviation.

Another strategy that many subjects used involved maximizing the number of planes landing simultaneously. A special opportunity for this occurs when the wind direction changes. This allows subjects to use the runways in a new direction, while planes are still taxiing on runways in the former direction. For instance, while landing planes on the north-south runways, a subject can respond to a change in wind direction to east or west by landing planes

on the east-west runways. These "crossover landings" when a subject lands a plane while at least one other plane occupying an orthogonal runway. Crossover landings possible only if subjects quickly respond to a wind direction change during a brief period after the change. We define *runway efficiency* as the percentage of crossover landings achieved by the subjects out of the maximum crossover landings possible within a trial. Figure 4 plots mean runway efficiency and standard deviation for the 58 subjects across trials. As can be seen, runway efficiency increases throughout the experiment and maintains a fair degree of inter-subject variability.
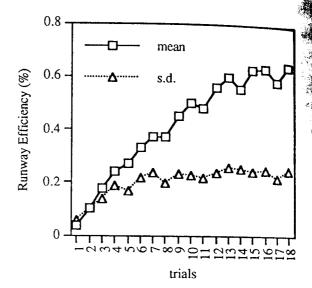

Figure 4: Mean runway efficiency and standard deviation.

We wanted to assess the contribution of these strategy variables to Score, controlling for psychomotor factors. We defined two such psychomotor variables. One is *total keystrokes*, the total number of the relevant keys (↑, ↓, F1, and ⌐) used per trial. The other measure is *mean reaction time* to orthogonal wind direction change.[3] Figure 5 plots mean total keystrokes and standard deviation, and Figure 6 plots mean RT to wind change and standard deviation for the 58 subjects across trials.

As can be seen, total keystrokes increased across trials, while maintaining a fair degree of inter-subject variability. This is to be expected, given high inter-subject variability in the use of hold 1 strategy. Subjects who use hold 1 strategy require fewer keystrokes to achieve performance comparable to that of subjects who do not use hold 1 strategy .

Mean RT to wind change decreased steadily, with a corresponding reduction in inter-subject variability. Nevertheless, the ratio between the standard deviation and the mean RT to wind change remains relatively constant.

---

[3] Ackerman (1988) used mean reaction time to wind change, as the dependent measure of performance. However, we chose to use cumulative score as the dependent measure of performance, because that is what subjects are trying to optimize in this task.
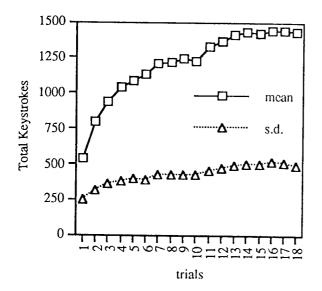
508

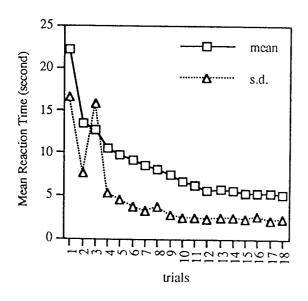Figure 5: Mean keystrokes and standard deviation.

Figure 6: Mean reaction time to wind change and standard deviation.

Correlations between strategy variables, psychomotor variables, and Score

We examined the Pearson r correlations between the four predictor variables and Score. We used the 18 trials for each of the 58 subjects as data, for a total of 1044 observations. Table 1 lists the correlations between the five variables: hold 1 strategy (HS), runway efficiency (RE), total keystrokes (KS), mean RT to wind change (MT), and Score. Runway efficiency correlated most strongly with Score, $r = .839$, followed by mean RT to wind change, $r = -.668$. In addition, runway efficiency is fairly strongly but negatively correlated with mean RT to wind change, $r = -.552$, and hold 1 strategy is strongly but negatively correlated with total

keystrokes, $r = -.604$. This supports our assertion that hold 1 strategy reduces the number of keys necessary to land a plane. By breaking down the eighteen trials into halves (1-9 and 10-18), we see a striking decrease in the correlation between total keystrokes and Score: r dropped from .575 in the first half of the trials to .107 in the last half of the trials.

Table 1: Correlations between Score, hold 1 strategy (HS), runway efficiency (RE), total keystrokes (KS), and mean reaction time to wind change (MT).

	Score	HS	RE	KS
HS	.291	-	-	-
RE	.839	.349	-	-
KS	.492	-.604	.316	-
MT	-.668	-.178	-.552	-.347

These results indicate that while using more keys raised Score in the early trials, it had little impact on Score in the later trials. If increase in Score was due to an increase in motor speed, one would expect the correlation between total keystrokes and Score to increase with repeated trials. The opposite is true, however, which indicates that increase in motor speed, as reflected by total keystrokes and mean RT to wind change, was neither the sole nor the most important determinant of Score. Indeed, what best predicted Score was a combination of strategy and speed, as reflected by runway efficiency. We explore this issue in more detail in the following regression analyses.

Regressions with strategy and psychomotor variables as predictors of Score

A multiple regression using all four variables to predict Score accounts for 87.7% of the variance. The regression equation with the four variables is:

Score = - 307 + 1566 HS + 1524 RE + 1.43 KS - 22.1 MT

with the following t-ratios[4]: HS = 25.31, RE = 22.57, KS = 30.72, and MT = -10.68. Mean RT to wind change contributed the least to this four-variable model. A model using only hold 1 strategy, total keystrokes, and runway efficiency still accounts for 86.3% of the variance. The regression equation with the three variables is:

Score = - 843 + 1745 HS + 1676 RE + 1.60 KS

with t-ratios of 27.80, 24.11, and 34.73, respectively. Deleting any of the other three predictor variables leads to much bigger reductions in the prediction of Score. This indicates that runway efficiency is a better predictor of Score than mean RT to wind change.

Our analysis indicates that subjects' Score increased with adoption of either the hold 1 strategy or the multiple-landings strategy, as measured by runway efficiency. However, overall speed also contributes to Score, especially

---

[4]   All t-ratios in the regression analyses we present have 1 $d.f.$ and are significant at $p < .001$.

509

in the early part of the experiment. In the first half of the experiment, the regression equation is:

$$Score = -1237 + 1862\ HS + 1619\ RE + 1.95\ KS$$

with t-ratios of 20.45, 12.10, and 28.09, respectively. In the second half of the experiment, however, it changes to:

$$Score = 632 + 879\ HS + 1873\ RE + 0.764\ KS$$

with t-ratios of 10.54, 28.07, and 12.04, respectively. This change indicates the increasing importance of the runway efficiency across trials relative to other factors, such as total keystrokes. The multiple runway strategy, as measured by runway efficiency, allowed subjects to have more runways in service at the same time. As subjects become more skilled, the length of time required by planes to taxi down runways (15 seconds) became the performance-limiting factor.

Taxiing time limits the utility of the hold 1 strategy. When both runways are in use, the additional resources afforded by hold 1 strategy are wasted. Increasing one's key efficiency is irrelevant when both of the runways are occupied and no additional planes can be landed. This is indicated both by the reduction in the importance of hold 1 strategy in the second half of the trials compare to the first in the regression analysis, and by the fact that subjects' use of hold 1 strategy reached an asymptote fairly quickly at about the fifth trial (Figure 7 plots mean hold 1 strategy use by low-third and high-third performers).
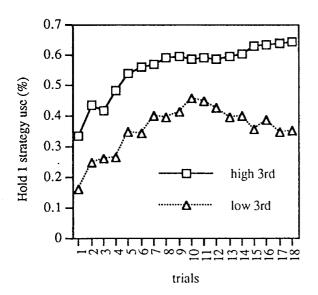
Figure 7: Hold 1 strategy use of low third and high third subjects as measured by Score.

Runway efficiency, which measures opportunistic use of the opposing runways during wind change, does not suffer from this performance-limiting factor. This is indicated by the continuing increase in subjects' runway efficiency across trials (Figure 8 plots mean runway efficiency of low-third and high-third performers).
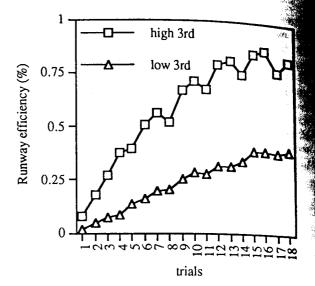
Figure 8: Runway efficiency of low third and high third subjects as measured by Score.

As previously noted, Ackerman found that task-external measures of individual differences in both cognitive ability and psychomotor speed predicted performance in the ATC task. He also found that cognitive ability had a stronger effect early in the experiment, while psychomotor skills had larger impact later in the experiment. This might seem to contradict the results of our analysis of task-internal variables (hold 1 strategy, runway efficiency, and total keystrokes), which indicates that keying speed becomes less significant in later trials. It is difficult to compare our results with Ackerman's, because Ackerman used mean RT to wind change as the dependent measure for subject performance, while we use Score. One might argue that mean RT to wind change came to reflect psychomotor factors towards the end of the experiment, which explains Ackerman's conclusion that psychomotor factors best predicted performance. However, psychomotor factors themselves cannot be such a significant contributor to Score given the relatively slow rate of planes taxiing on the runway and the invention of the hold 1 strategy, which relieves the need for keystrokes. Instead, runway efficiency emerges as an important factor, allowing subjects to turn psychomotor skill into useful actions for increasing Score.

Discussion

Our analysis of the ATC data indicates that skill improved in a complex way across trials in the experiment. Subjects improved both in the strategies used to solve the task and in the speed with which they executed the task. This indicates that any satisfactory model of this task will have to include both a strategic improvement component and a speed improvement component. There are currently many different theories of skill acquisition. It is useful to examine some of them in light of the findings presented here. We examine two in particular: Logan's (1988) instance theory of

automatization and the Soar (Newell & Rosenbloom, 1981, Newell, 1990) chunking theory of learning.

Logan (1988) proposed a theory of learning based on the idea of retrieving "instances" from memory. An instance is composed of the stimulus, subjects' goal during the encounter with the stimulus, their interpretation of the stimulus in relation to the goal, and their response to the stimulus. Logan argued that subjects automatically encode and store each encounter with a stimulus as an instance, and retrieve the instance when the stimulus is encountered again. In this theory, although skills are initially acquired through algorithmic processing, performance depends increasingly on retrieving past instances of solutions from memory.

The Soar (Newell & Rosenbloom, 1981; Newell, 1990) theory of learning is based on chunking. In this theory, skills are acquired by chunking together productions that successfully solve a problem. For example, if subjects encounter an impasse during a problem-solving episode, they can set a subgoal to solve that impasse. If they are successful at finding a solution to the impasse, the individual solution steps taken are then chunked together as a single production. When subjects encounters the same situation again, the newly chunked production can be used automatically, eliminating the need to repeat the laborious process of searching for a solution.

Both theories are capable of explaining why keystroke rate was related to improvement in Score. Logan's instance retrieval and Soar's chunking result in less "cognitive time" between actions. However, it is unclear how either of these theories could predict the contribution of hold 1 strategy, which requires a reorganization of behaviors and a shift to a strategy that requires a different sequence of keystrokes. In addition, subjects do not switch hold strategies (i.e. hold 3, hold 2, and hold 1) in a step-like transition, as Logan's theory and Soar's model would predict. Instead, the hold strategies overlap by a fair amount during transition, and in most cases subjects do not completely abandon the use of hold 2 and 3. Also, how these theories could explain the important contribution of runway efficiency (multiple runway use through crossover landings) to Score is unclear, since this would depend on the particulars of the initial algorithms used.

While we have yet to undertake the simulation effort, we think that the ACT-R theory is well suited to model these components, since it separately learns over trials which strategies are better and how to execute each strategy more successfully. For example, Lovett and Anderson (in press) have shown, in an artificial problem solving task, that subjects learn in both dimensions of strategy and speed. In particular, their subjects come to execute the less successful strategy less frequently and yet more rapidly with experience. They have modeled these phenomena successfully in ACT-R.

• Skill acquisition in the ATC task involves a complex interaction between improvements in strategies and improvements in speed. We are currently working on an ACT-R model to explain these phenomena, as a first step toward extending the ACT-R theory to dynamic skill acquisition.

References

Ackerman, P.L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*, 288-318.

Ackerman, P.L. (1990). A correlational analysis of skill specificity: Learning, abilities, and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 883-901.

Ackerman, P.L. & Kanfer, R. (1994). Kanfer-Ackerman Air Traffic Controller Task© CD-ROM Database, Data Collection Program, and Playback Program. Office of Naval Research, Cognitive Science Program.

Anderson, J.R. (1993). *Rules of the Mind*. Hillsdale, NJ: Erlbaum.

Anderson, J.R., Corbett, A.T., & Conrad, F. (1984). Learning to program in LISP. *Cognitive Science, 8*, 87-129.

Chi, M.T.H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence*. Hillsdale, NJ: Erlbaum.

Larkin, J., McDermott, J., Simon, D.P., and Simon, H.A. (1980). Expert and novice performance in solving physics problems. *Science, 208*, 1335-42.

Logan, G.D. (1988). Towards an instance theory of automatization. *Psychological Review, 95*, 492-527.

Lovett M.C. & Anderson, J.R. (in press). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology*.

Newell, A. & Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the law of practice. In J.R. Anderson (Ed.). *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.

Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, M.A.: Harvard University Press.

Newell, A. & Simon H. (1972). *Human Problem Solving*. Englewood Cliffs, N.J.: Prentice Hall.

Singley, M.K. & Anderson, J.R. (1989). *The Transfer of Cognitive Skill*. Cambridge, M.A.: Harvard University Press.