In Press: To Appear in D. Gopher & A. Koriat (Eds.),
Cognitive regulation of performance: Interaction of theory and application. Attention and Performance XVII.

11 Bringing Together the Psychometric and Strategy Worlds: Predicting Adaptivity in a Dynamic Task

Lynne M. Reder and Christian D. Schunn

ABSTRACT There are two traditional approaches to the study of individual differences in cognitive skill. One assumes that people differ in the strategies that they use. The other assumes that all people use the same strategies or processes but differ in one or more performance parameters affecting how these processes are executed (e.g., speed, memory capacity). This chapter explores another possibility, that people differ in how well they adaptively shift strategies in response to changing features of the task environment. To test this, we examined the performance of 57 participants in a variant of the Kanfer-Ackerman Air Traffic Control Task (Kanfer and Ackerman 1989), a dynamic task in which features of the environment frequently change. We found that, while most participants adapted their strategy selections in response to our manipulations of the task environment, not all participants were equally adaptive. Furthermore, using the CAM 4 (Kyllonen 1993), a cognitive assessment battery, we were able to determine what cognitive subskills were associated with adaptiveness. In this context, we found that inductive reasoning skill (and not working memory, declarative learning, procedural learning, or processing speed) was associated with adaptiveness to our specific manipulations, and to the general dynamic character of the air traffic control (ATC) task.

Historically, there have been two lines of inquiry related to individual differences in cognitive skill. The first assumes that people have fixed processes that vary on the settings of some process parameters, such as speed of processing (Kail 1988; Salthouse 1994) or the capacity of working memory (Just and Carpenter 1992); it attributes differences in performance primarily to differences in capacity or aptitude on a fixed set of dimensions.

The second line of inquiry assumes that each individual uses just one strategy, or at most evolves to using another, and that there is no switching back and forth among a variety of strategies. This alternative view attributes differences in performance to the different strategies adopted by different individuals to solve the same task.

In this chapter, we wish to explore a third possibility. We suspect that a large fraction of performance differences among individuals may be explained by something other than differences in unitary parameters (such as working memory capacity) affecting how a fixed set of strategies are executed or differences in the explicit strategies that are adopted. We propose that performance differences among participants may result from their differing ability to shift strategies as the task demands change. In the sections that follow, we shall describe these three alternative views in further detail.

11.1 INDIVIDUAL DIFFERENCES UNDERSTOOD AS PERFORMANCE PARAMETER DIFFERENCES

There have been numerous investigations into the nature of individual differences in task performance and acquisition of skills from the perspective of performance parameter differences. These investigations have included psychometric approaches (e.g., Ackerman 1989; Snow, Kyllonen; and Marshalek 1984) as well as information-processing approaches (e.g., Just and Carpenter 1992; Lovett, Reder, and Lebière 1996; Sternberg 1977). Some have focused on finding factors general to many tasks and domains, and have found a single general factor, *g* (e.g., Spearman 1904), or a set of process-specific factors—working memory (e.g., Case 1985; Just and Carpenter 1992), processing speed (e.g., Kail 1988; Salthouse 1994), decision-making skill (e.g., Hunt, Joslyn, and Sanquist 1996), and so on. Others have focused on skills specific to particular tasks or domains (e.g., Gardner 1983; Thurstone 1938). Finally, some have built hierarchial models combining task-specific and task-general factors (e.g., Ackerman 1988; Kyllonen 1993).

In all these cases, there has been strong evidence that performance differences among individuals can be predicted by differences on a battery of ability tests. These results have been used to argue that performance differences are due to structural differences or capacity differences in some cognitive hardware. While not all researchers cited above would agree, the thrust of the arguments represented in those papers is that individual differences stem from inherent differences in their aptitude or ability to execute the *same* set of strategies, and not from differences in the strategies themselves.

11.2 INDIVIDUAL DIFFERENCES UNDERSTOOD AS DIFFERENCES IN STRATEGY USE

The alternative perspective taken in the literature is that the differences among people can be understood as differences in the *strategies* used rather than in aptitude or ability to use a given strategy or set of strategies. Many researchers putting forth this view (e.g., Cooper and Regan 1982; Farah and Kosslyn 1982; MacLeod, Hunt, and Matthews 1978) were in fact questioning the simpler view that all people achieve a particular goal using the same strategy or method. Their research demonstrated that different groups of participants could produce different patterns of results in the same experiment, and that these patterns could most easily be understood by assuming different strategies for accomplishing the same task. Moreover the error patterns and the participants' retrospective reports after the experiment supported these conjectures of different strategies for different participants.

Schunt

11.3 INDIVIDUAL DIFFERENCES UNDERSTOOD AS DIFFERENCES IN ADAPTIVITY OF STRATEGY USE

In contrast to both the view that a given individual uses only one strategy (or a fixed set of strategies) to accomplish any task and the view that individual differences are due to different strategies being selected, we propose that individual differences may be accounted for by differences in *adaptivity* of strategy selection, that performance differences among individuals arise from the optimality of the particular strategy used at any given time and the speed with which individuals shift their strategy use in response to changes in the environment. Which is to say that even though people may have the same strategies or procedures in their repertoire and may be equally proficient in executing them, performance differences arise from differences in knowing when to apply a particular procedure or strategy. This position has had few if any proponents.

Evidence for Strategy Alternation

The view first proposed by Reder (1982, 1987, 1988) that individuals will alternate among strategies within a given task has only recently gained wider acceptance (e.g., Lovett and Anderson 1996; Siegler 1988). Earlier theories of skill acquisition (e.g., Anderson 1982, 1983; Fitts 1964; Schneider and Shiffrin 1977) tended to postulate an improvement in performance as increased speed in the execution of a single strategy.

Reder (1979, 1982) found that participants did not always execute the strategy that nominally corresponded to the task instructions, frequently electing to use a different strategy to accomplish the task. In Reder 1982, participants were to answer questions based on short stories they had read: one group was asked to make recognition judgments, that is, to judge whether a particular sentence had been presented as part of the story, while the other was asked to judge whether a particular sentence was plausible, given the story. Participants in the recognition group did not exclusively use a retrieval strategy to search for a verbatim trace, nor did the plausibility group exclusively use a plausibility strategy; frequently, participants in one group tried the strategy that corresponded to the other task. Preference for a particular strategy was influenced not only by official task demands, but also by other variables such as delay between reading the story and answering the questions.

In another series of experiments, Reder (1987) found that participants' strategy preferences were also influenced by other situational characteristics of the task, such as the proportion of trials for which a given strategy had been working in the recent past. Participants adjusted their tendency to adopt the plausibility strategy over the direct retrieval strategy as a function of the proportion of trials where the statement to be judged or its exact

lc s

contradiction had been explicitly stated as part of the story. This proportion was varied across blocks of the experiment, and participants adjusted their use of the two strategies accordingly.

Participants have been seen to adapt to shifting proportions of the features of an experiment in a number of other contexts, such as arithmetic verification (Lemaire and Reder forthcoming) and in a problem-solving task (Lovett and Anderson 1996).

Factors that Affect Strategy Choice

Given that strategy selection has been found to vary within an individual within a task, one can ask what determines how the strategy is selected at any given time. Reder (1987, 1988) theorized that the strategy selection process involves two mechanisms, one sensitive to *extrinsic* factors and the other sensitive to *intrinsic* factors. The mechanism sensitive to extrinsic factors responds to cues in the situation rather than the question or problem, whereas the mechanism sensitive to intrinsic factors responds to cues within the question or problem itself, which can produce a quick feeling of knowing used to guide a direct search of memory rather than the execution of a backup reasoning or computational procedure. Cues in the question also can be used to guide the specific type of reasoning strategy evoked.

One type of evidence for the role of intrinsic features comes from the story paradigm described above. Reder (1982, 1988) demonstrated that participants prefer the direct retrieval strategy for questions about stories just read and the plausibility strategy for questions pertinent to the older stories. In Reder 1988, participants did not know before seeing a question whether it referred to a story just read or to a story from two days earlier, because questions from the two types of stories were intermingled. In this manner it was possible to determine if a rapid inspection of the question (i.e., intrinsic variables—the words themselves) affected response strategy selection. In fact, participants did use different strategies depending on the age of the story to which the questions referred: participants used inferential reasoning when the questions referred to old stories and a direct retrieval strategy for questions concerning new stories.

11.4 EXAMINING STRATEGY CHOICE IN A DYNAMIC TASK WITH CHANGING TASK CHARACTERISTICS

From the work described in the previous sections, we know that people are sensitive to environmental characteristics in that their preference for particular strategies is affected by these characteristics. On the other hand, there is still much we do not know about strategy selection. Of particular relevance to the issue of individual differences, we do not know whether there are differences among people in their adaptivity to environmental characteristics. Do some people show consistently more adaptive use of strategies in given

| Client: MIT-U | J-8878 Gopher | Neuro 45 | Palten_M | (.01) .29 .02 .05 |

situations? And do people vary in terms of how quickly they shift their strategy use when aspects of the environment change?

The kinds of domains where this issue of individual differences in adaptivity is likely to be especially important are ones that involve dynamic tasks, namely, those where the features of the task or problem at hand change in real time, independent of any actions of the participant, for example, driving a car, flying a plane, directing air traffic. Because the environment is constantly changing, in dynamic tasks it becomes especially important to be adaptive in one's strategy selections.

However, most research on strategy selection has been conducted using static tasks, for example, answering questions about a story, solving a math problem or attempting a simple problem-solving task. Clearly, dynamic tasks pose a greater cognitive load on the performer than static tasks. Conceivably, when a task has a high cognitive load and changes dynamically, an individual may not be able to attend (implicitly or explicitly) to noncritical features of the task that are also changing, such as base rate information about features of the situation. Thus they may not be able to shift strategy use in response to changes in base rates.

For these reasons, we have chosen to explore individual differences in a dynamic task, specifically an air traffic controller's task. In particular, the study presented in the remainder of this chapter addresses the following questions:

1. Are participants adaptive in their strategy use in a dynamic task such as air traffic control?

2. If participants are adaptive, are there individual differences in amount of adaptation? That is, are some participants adaptive, others not, and some intermediate?

3. If participants vary in their adaptive use of strategies in response to changing environmental features, does this level of adaptation correlate with overall performance?

4. If participants vary in their adaptiveness, can we predict what makes participants adaptive (i.e., determine which cognitive subskills are associated with adaptiveness)?

11.5 OVERVIEW OF AIR TRAFFIC CONTROLLER TASK

The Kanfer-Ackerman the Air Traffic Controller Task© (KA-ATC; Ackerman and Kanfer 1994; Kanfer and Ackerman 1989) simulates dynamic aspects of real air traffic control (e.g., planes lose fuel in real time, weather conditions change, certain types of planes require longer runways than others). Using the KA-ATC task, Ackerman (1988, 1989) has found that what predicts individual differences in performance varies with time during training: early in training, reasoning ability best predicts performance; later in training, perceptual speed ability best predicts performance, and by the end of training, simple reaction time ability best predicts performance. In contrast, coming from the strategy difference perspective, Lee, Anderson, and Matessa (1995) have found that strategy differences account for a large percentage of the variance in individual performance in the KA-ATC task at all points of training. Thus the KA-ATC task provides a good candidate for potentially unifying the strategy and performance parameter perspectives. In our study, we use a slightly modified version of the KA-ATC task designed to diagnose individual differences in strategy adaptiveness. A description of this version is presented below.

The KA-ATC task has as its primary goal to accumulate as many points as possible. Points are earned by landing planes and are subtracted by errors or rule violations, the most serious being allowing a plane to run out of fuel before landing it, thereby allowing it to crash. The KA-ATC task requires that the participant monitor a variety of elements that are displayed on the screen (see figure 11.1): twelve hold pattern positions that are divided into three altitude levels; four runways of two different lengths, one of each running north-south and the other running east-west; a queue of planes waiting to enter the hold pattern; two message windows (not shown), one noting changes in weather, runway conditions, wind direction and speed, and the other providing error feedback; and feedback on current score and penalty points.[1] A weather change occurs approximately every 26 seconds; planes enter into the queue every 7 seconds.

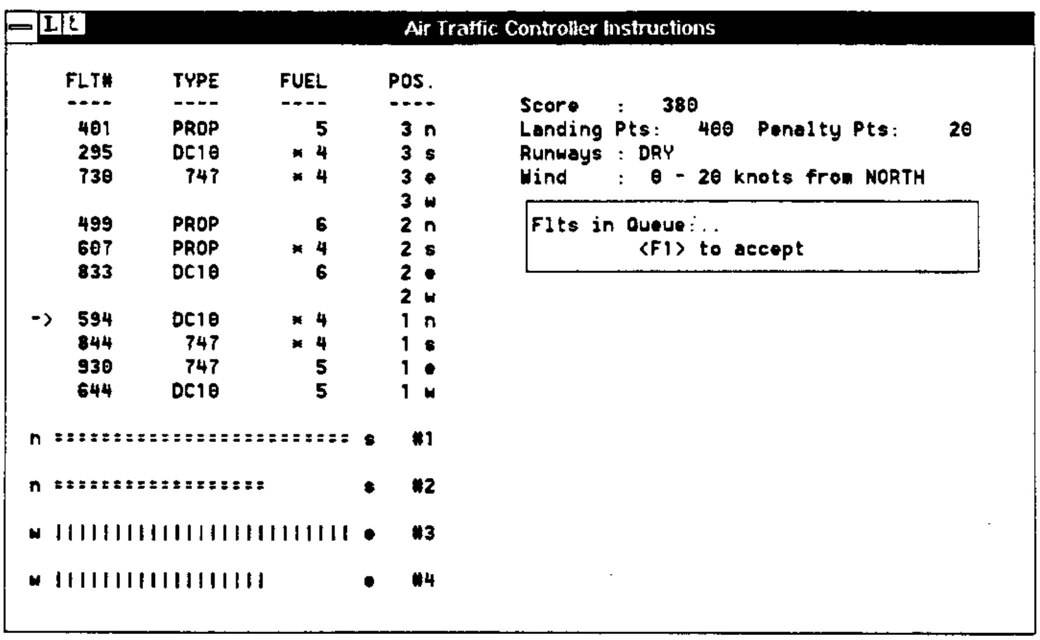


Figure 11.1 Main screen from our modified version of the Kanfer-Ackerman Air Traffic Controller Task (© Kanfer and Ackerman 1989).

There are six basic rules governing this task:

1. Planes must land into the wind (e.g., use a north-south runway rather than an east-west runway if the wind is coming from the north or south);

2. Planes can only land from lowest hold level (hold 1);

3. Planes can only move down one hold level at a time, and only into an unoccupied position in that level;

4. Ground conditions and wind speed determine the minimum runway length required by different plane types (747s always require long runways; 727s can use short runways either when the runways are dry or when wind speed is 0–20 knots; DC10s can use short runways only when the runways are dry or wet and the wind speed is less than 40 knots; and PROPs can always use short runways);

5. Planes with less than 3 minutes of fuel remaining must be landed immediately; and

6. Only one plane can occupy a runway at a time.

Our critical manipulation was to vary the proportion of 747s (i.e., planes that always required the long runway for landing). In our view, it would be adaptive to treat the long runway as a scarce resource and use it *only* for planes that require a long runway when these are many, but not to hold it in reserve when planes requiring it are few because the long runway is inherently easier to use. First, fewer keystrokes are required to land a plane on a long runway because the long runway is above the short runway. Second, using the short runway requires knowing and accessing the rules for when the short runway is legal, and requires checking the current wind and weather conditions, whereas one can always land a plane on the long runway independent of wind, weather, or plane type.

Thus the primary measure of interest was whether participants modified their tendency to treat the long runway as a scarce resource. We analyzed the data to see how often participants made sure to use the short runway when either was available and could be used, as a function of the proportion of planes that where 747s.

11.6 METHOD

Participants

There were 68 participants recruited from a temporary-employment agency near Brooks Air Force Base and paid for their participation. There were 57 participants in condition A, and 11 participants in condition B. The majority of the participants were assigned to condition A because we were primarily interested in predicting individual differences, and condition B was a simple control condition. The gender distribution in the sample was 79% male, 21%

female. All but 4% of participants had a high school diploma, and 22% had at least some college experience.

Design and Procedure

The task consisted of a series of 10-minute trials. A trial began with the computer screen similar to the one in figure 11.1—that is, with planes already in various hold positions and other planes in the queue. The number of minutes of fuel left for each plane was clearly marked at all times and decreased in real time. At the end of each trial, the participant was allowed a short break (time effectively came to a halt) and the screen display was reconfigured anew. The cursor pointed to the top position at the start of the trial.

The set of possible keystrokes on the computer keyboard were: the UP arrow, the DOWN arrow, F1, and ENTER. The UP arrow and DOWN arrow keys moved the cursor up and down (respectively) between the different hold positions and runways. The *F1* key accepted the planes from the queue into a holding pattern, and the *enter* key selected the plane in the hold corresponding to the placement of the cursor, or placed a selected plane (either from the queue or from another hold position) into an empty hold position, or landed a plane on the runway.

Participants were given nine 10-minute trials. These nine trials were divided into three blocks of three trials. There were two between-subject conditions that manipulated the proportion of 747s (and propplanes) across blocks in different orders. Two orders were used to ensure that the results were not peculiar to one particular order, nor simply due to changes which would have occurred naturally as a function of practice with the task (i.e., independent of our manipulation). In condition A, the proportions of 747s over the three blocks were 25%, 5%, 50%. In condition B, the order was 25%, 50%, 5%. The frequency of propplanes was set to be 55% minus the frequency of 747s (i.e., 30%, 50%, 5% in condition A, and 30%, 5%, 50% in condition B). Because the propplanes were the only planes that could always land on the short runway, varying the ratio of 747s to propplanes was the most direct way of manipulating the scarcity of the long runway resource. The proportion 727s and DC10s was held constant at 5% and 40%, respectively.

Because we were interested in predicting individual differences and sensitivity to one particular manipulation, the remaining structure of the task was held as constant as possible across blocks and participants while still maintaining the overall dynamic structure of the task.

Measures of interest The study used a variety of dependent measures. First, we collected keystroke information. These included keystrokes for the UP and DOWN arrow, ENTER, and F1 keys described above and also for

Client: MIT-U	J-8878 Gopher	Neuro 45	Palten_M	(.01) .29 .02 .05
(IDP) TPS 23/7	PMU: WSL 30/7	pp. 315-342	Ch11_P	(p. 322)

the six rule keys, which allowed a participant to review any of the six rules of the task. In addition, we analyzed performance on the basis of score (points accumulated) and the number of planes crashed.

The most critical dependent measure was one we defined for the purpose of our experiment, *op-short*, which referred to the occasions participants opted to use the short runway when both runways were open and could be used, based on current wind and weather conditions at landing time. Because the DC10 is the only plane type that occurred with a constant, high frequency in all blocks, we only calculated the proportion of op-shorts for CD10s.

Individual difference measures To establish that individual differences in adaptivity were not just due to random variation, and to discover more about the precise nature of these differences, we examined whether the differences in adaptivity could be predicted by performance on an assessment battery of various cognitive skills. We used the "Cognitive Abilities Measurement" (CAM; Kyllonen 1993, 1994, 1995) battery because it provided a broad range of tests plausibly related to adaptivity in strategy use. Moreover, the CAM battery has been used to predict learning and performance in a large number of training environments (see Shebilske, Goettl, and Regian, chap. 14, this volume; Shute 1993).

The full CAM is quite large, and requires several days to complete. We focused on a subset of the CAM that seemed most relevant for our purposes. We used eleven tests in all, for working memory, fact (or declarative memory) learning, skill (or procedural) learning, processing speed, and inductive reasoning. For all but inductive reasoning, we included one test in the verbal domain and another in the spatial domain. The spatial and verbal variants were very similar to one another in structure. For the inductive reasoning, only spatial reasoning tests were available at that time; we included the three that were available.

We used two criteria for selection of CAM tests. First, we wanted to include measures we thought were plausibly related to adaptivity in strategy selection. For example, it seemed plausible that working memory and inductive reasoning might be associated with greater sensitivity to shifts in base rate of 747s which might in turn affect adaptivity. Second, we felt it was important to include factors that have historically been used to predict performance in psychometric studies so that we could compare our results with others in the individual difference literature and determine whether the same factors or different factors predict adaptivity. Table 11.1 presents the tests we used; the chapter appendix provides more detailed descriptions of the tasks.

For each of the tests, we used only one global score per participant, overall percentage correct on that measure, except on processing speed, where we used median reaction time in addition to overall percentage correct. This produced thirteen predicting variables per participant.

| Client: MIT-U | J-8878 Gopher | Neuro 45 | Palten_M | (.01) .29 .02 .05 |

Table 11.1 The Eleven CAM Tests Used in the Study, Classified by Skill Type and Content Domain

Skill	Verbal	Spatial
Working memory	4-term ordering: Furniture/Animals	4-term ordering: Blocks
Processing speed	2-term ordering: Furniture/Animals	2-term ordering: Blocks
Fact learning	Word recognition	Figure recognition
Skill learning	Reduction: Future-past-present	Reduction: Circles
Inductive reasoning		Figure sets
		Figure series
		Figure matrices

Another obvious source of potential individual differences in task performance is the degree to which participants pay attention to the instructions. To assess and control for such differences, we gathered data on how long the participants spent reading each of the 61 instruction pages. These timing data were compressed into four variables: the time spent reading each of the three rules pages (which displayed rules 1–3, 4–5, and 6 respectively), and the median time spent reading the remaining pages.

11.7 RESULTS

What is the Adaptive Response?

In section 11.6, we argued that the adaptive response to our manipulation of the ratio of plane types was to change the op-short (the proportion of DCIDs landed on the short runway when both the long and short runways were available and legal). In particular, we argued that the participants should have a low op-short when there are few 747s, and a high op-short when there are many 747s. This particular response pattern will form the basis of dividing the participants into adaptive and nonadaptive groups. However, as a manipulation check, we wanted to first test our assumption that differing levels of op-short across blocks, corresponding to the differing plane ratios, actually produced better scores than not doing so, that is, were we correct that such shifts were adaptive?

To answer this question, we regressed each participant's score for each 10-minute trial against the proportion of op-short for that trial, and this was done separately for blocks 2 and 3, within each condition. Trials in which the participant had fewer than three opportunities to select among two available runways were excluded. Within condition A, block 2 (5% 747s), op-short was negatively correlated with score ($r = -.32$), $\beta = -754.0$, $t(122) = -3.7$, $p < .001$). Within block 3 of condition A (50% 747s), op-short was positively correlated with score ($r = .35$, $\beta = 843.1$, $t(135) = 4.4$, $p < .0001$). A similar pattern emerged when we regressed op-short against number of plane crashes. The regression analyses for condition B provided a similar

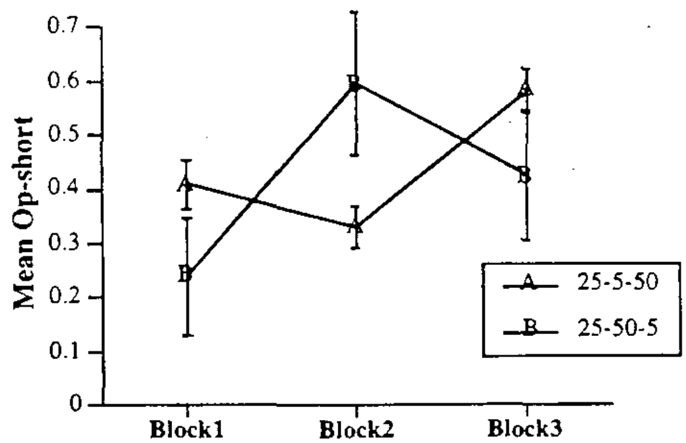


Figure 11.2 Mean proportion (and S.E.) of op-short with each condition within each block of three trials.

story, although not as strongly because there were far fewer data points. Within block 3 of condition B (5% 747s), op-short was uncorrelated with score ($r = .04$, $\beta = 48.0$, $t(22) = .2$). By contrast, within block 2 of condition B (50% 747s), op-short was positively correlated with score ($r = .46$, $\beta = 355.9$, $t(18) = 2.2$, $p < .05$).

Thus the regression analyses indicate that it was generally better for participants to use the short runway whenever there were many 747s. Conversely, it was more adaptive not to hold the long runway in reserve when there were few 747s. Furthermore, as we will report later, adaptivity (defined as shifting the tendency to use op-short in response to changing plane ratios) is associated with much higher scores.

Did Participants Generally Adapt in the Expected Direction?

To assess the impact of the manipulation, we conducted an ANOVA on the op-short measure with condition and block as factors. Overall, the main effect of condition was not significant ($F(1, 48) < 1$), whereas the main effect of block, and the interaction with condition were significant ($F(2, 96) = 11.25$, $p < .0001$, and $F(2, 96) = 7.41$, $p < .001$, respectively). For condition A, the participants had the predicted pattern of medium-low-high (see figure 11.2). By contrast, the participants in condition B had the predicted pattern for only blocks 2 and 3 (high, then low)—their block 1 was somewhat lower than expected, perhaps reflecting a lack of knowledge of the rules at the beginning of experiment. In general, however, participants adapted in response to the manipulation, and they did so in the expected directions.

Because the remaining analyses focus on individual differences in adaptivity, they were conducted only on the data from condition A ($N = 57$), the main condition for which we have a sufficient N to study individual differences. As we said earlier, condition B ($N = 11$) was included in the experiment just

Client: MIT-U	J-8878 Gopher	Neuro 45	Palten_M	(.01) .29 .02 .05

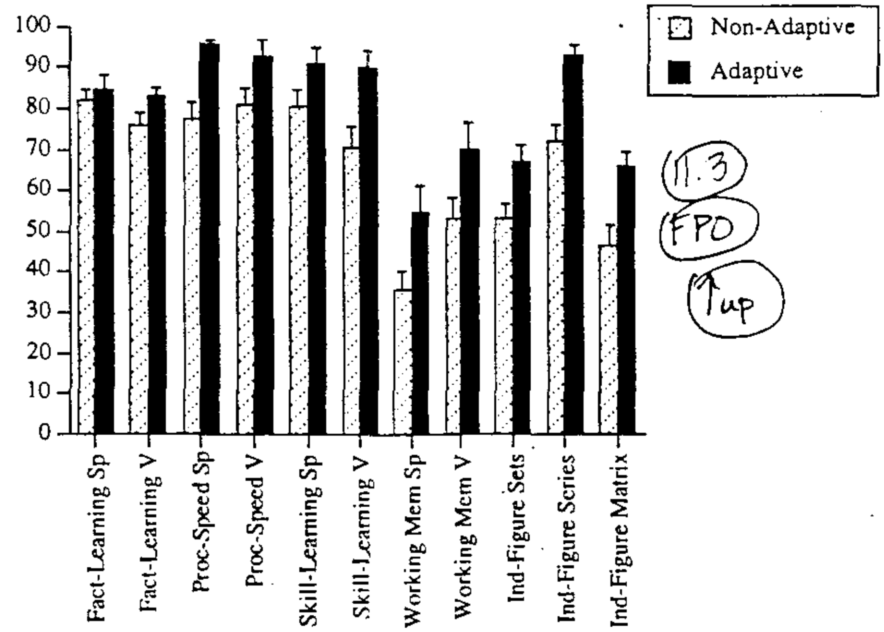
to ensure that the effect of the manipulation across blocks was not due to an artifact of the order of blocks.

This overall pattern of performance presented in figure 11.2 was used to define what was considered adaptive in this task. In particular, we divided participants into adaptive and nonadaptive groups according to the following criterion: participants were classified as "adaptive" if their op-short for the second block was lower than in the first block and their op-short was higher in the third block than in the first block (i.e., a medium, low, high pattern).[2] In fact, those participants who were classified as "adaptive" had much higher mean block scores than those classified as "nonadaptive" (4,104 versus 2,544), supporting our use of the label "adaptive."

Fourteen of the participants could not be classified because they had too few opportunities (less than three) to exhibit a preference in one or more of the blocks. To understand how this might occur, we return to the definition of op-short: the proportion of DC10s landed on a short runway when both runways were open and currently legal. Note that two conditions must be met for a behavior to be included in this proportion: (1) landing a DC10; and (2) both runways must be open and legal (based on the current wind and weather conditions). While 90% of DC10s landed met the second condition, for participants who landed few or no planes, there were too few opportunities to evaluate their op-short use. These participants performed very poorly (mean total score per block of −2,302 versus 3,052 for the other participants) and crashed many planes (mean number of crashes per block of 16.5 versus 1.8 for the other participants). Although one could simply assume these participants were not adaptive, we did not want to include them in our analyses because (1) we were not sure they understood the task (none of these participants had a positive mean block score) and (2) we wanted to separate, at least partially overall performance from adaptivity. These participants were therefore excluded from all the subsequent analyses (and from the score means presented in the preceding paragraph).

One alternative interpretation to why certain participants were nonadaptive is that some other difference made it less useful for them to change op-short levels in response to our manipulations. For example, perhaps these participants were less familiar with the rules, or perhaps they had lower working memory capacities, such that they may have made more errors if they had tried to change op-short levels in response to our manipulations. In other words, it may have been optimal for these particular participants not to change their strategy use. To examine this possibility, we recomputed the correlation between op-short use and score separately for the nonadaptive participants within blocks 2 and 3. We found that the same patterns held for nonadaptive participants as for all participants combined: there was a negative relationship between op-short use and score in the second block, and there was a positive relationship between op-short use and score in the third block. Thus it appears that the participants classified as "nonadaptive" truly were performing suboptimally.

Client: MIT-U	J-8878 Gopher	Neuro 45	Palten_M	(.01) .29 .02 .05



11.3
FPO
up

Figure 11.3 Mean percentage correct (and S.E.) on each CAM test for adaptive and nonadaptive participants.

What Predicts Adaptivity?

Of the remaining 43 participants, 14 (33%) were coded as "adaptive," and the remaining 29 (67%) were coded as "nonadaptive." Of primary interest was whether we could predict which participants were adaptive and which were nonadaptive. If one compares profiles of the two groups of participants on the CAM battery, the two groups differ on numerous dimensions (see figure 11.3); because many CAM scores are positively correlated with each other, however, it is unclear from these profiles which of these measures are actual independent predictors of adaptivity.

To assess which measures were independent predictors of adaptivity, we conducted a stepwise regression. All thirteen CAM scores and the four measures of instruction reading time were used as possible predictors, and the dependent variable was the binary outcome: adaptive or nonadaptive. One individual difference measure entered into predicting adaptivity: the inductive reasoning test, figure series ($r = .46$, $\beta = 0.010$, $t(39) = 3.2$, $p < .003$).

Was this relationship between inductive reasoning and adaptivity derivative of some other relationship between inductive reasoning and overall op-short use? For example, perhaps only those with high levels of op-short overall or high levels in the first block could show the adaptive pattern. To test this possibility, stepwise regressions were performed with absolute op-short use in each block as the dependent measure. A different pattern was found within each block.

For the first block, one factor entered into predicting proportion op-short: time spent reading the rule for when each plane type could land on the short runway ($r^2 = .21$, $\beta = 0.006$). Participants who spent more time reading the rules used the short runway more often. For the second block, four factors entered, in order of predictiveness (multiple $r^2 = .49$): skill learning spatial ($\beta = -0.009$), fact learning verbal ($\beta = 0.11$), and processing speed spatial ($\beta = -0.004$), and time spent reading rules 4 and 5 ($\beta = -.0046$).[3] For the third block, no factors were significant predictors. These regression analyses suggest that inductive reasoning is not related to overall op-short use. Instead, inductive reasoning is specifically predictive of the differential use of the short runway in response to the manipulation—that is, adaptive strategy use.

What Predicts Extent of Adaptivity?

The binary classification of participants into "adaptive" and "nonadaptive" does not distinguish between participants who adapted only slightly and those who shifted their strategy use a great deal in response to the manipulations. Furthermore, there were two components to the definition of adaptivity: a drop in op-short use from the first to second block, and a rise in op-short use from the second to third block. It is possible that separate factors predict these two transitions. We therefore constructed two new measures of adaptivity: the difference in mean op-short use between the first and second blocks, and the difference in mean op-short use between the second and third blocks.

Did the factors entering into a stepwise regression differ for the two transitions? For the first transition, one factor (time spent reading the rule for when people could land planes) only weakly predicted the difference in op-short use between the first and second blocks ($r^2 = .11$, $\beta = -0.004$). It seems that participants who did not know the rules very well had a stronger incentive to stop conserving the long runway in the second block (when conserving was not very beneficial). By contrast, for the second transition the inductive reasoning, "Figure Series" measure strongly predicted the difference in mean op-short use between the second and third blocks ($r^2 = .33$, $\beta = 0.008$). Thus the same factor that predicted presence of adaptation also predicted extent of adaptation.

What Predicts Speed of Adaptivity?

Another measure of adaptivity is how fast people adapt. We calculated rate as the proportion of an eventual adaptation made immediately following the transition in 747/prop plane ratios (i.e., at trials 4 and 7). This was done as follows. We measured amount of eventual adaptation as the difference between (1) op-short use on the trial immediately prior to the transition (i.e., blocks 3 or 6) and (2) the most extreme op-short value on the three trials following the transition (i.e., the minimum of trials 4, 5, 6 or maximum of trials, 7, 8, 9). This was simply a variant of extent of adaptation discussed in the previous section. The amount of immediate adaptation was calculated as the difference between (1) the op-short use for the last trial prior to a transition (i.e., trials 3 or 6), and (2) the op-short use for the first trial after that transition (i.e., trials 4 or 7). We divided the amount of immediate adaptation by the eventual transition amount, giving a proportion. A score of 1 reflected adaptation completed entirely immediately, zero reflected no immediate adaptation, and values in between reflected intermediate adaptation rates. People who adapted in the wrong direction—either overall, or on the first transition block—or who did not adapt at all were assigned a zero on this measure.

For the first transition—from trial 3 to trial 4—no factor predicted adaptation rate. For the second transition—from trial 6 to trial 7—two factors predicted adaptation rate significantly (multiple $r^2 = .55$): processing speed spatial RT ($\beta = -0.25$) such that spatially fact participants adapted more quickly, and working memory verbal ($\beta = 0.006$), such that high verbal participants adapted faster. Thus separate factors predicted rate of adaptation than predicted amount of adaptation.

Were Adaptive Participants Just More Biased to Use the Short Runway?

It is possible that the nature of participants' adaptivity was just a bias to use the short runway when there was a large demand for the long runway, but that they were actually no better at using the short runway correctly. To use the short runway appropriately required greater cognitive effort because one had to have committed the requisite rules to memory, verified the weather and ground conditions, and verified that DC10 could use the short runway under the current weather conditions. Conceivably, adaptive participants might have risked more errors and tried to use the short runway indiscriminately while nonadaptive participants might not have risked making the errors and thus did not look as adaptive.

To assess this, we defined a hit as landing a DC10 on the short runway when both runways were open and legal for the DC10. A false alarm was defined as trying to land a DC10 on the short runway when both runways were open, but the short runway was not legal for a DC10.

There was a significant effect of block on hit rate ($F(2, 82) = 32.6$, $p < .0001$), showing that overall participants were more likely to use the short runway when it mattered, (i.e., late in the experiment). Adaptive participants had a marginally higher hit rate than nonadaptive participants, ($F(1, 41) = 3.5$, $p < .07$). There was a significant interaction of participant type by block ($F(2, 82) = 7.3$, $p < .002$) showing a higher hit rate for the adaptive participants in the third block (see table 11.2). This effect is to be expected because we defined adaptivity in terms of the shift in use of the

Client: MIT-U	J-8878 Gopher	Neuro 45	Palten_M	(.01) .29 .02 .05

Table 11.2 Mean Hit and False Alarm Rates within Each Block for Adaptive and Nonadaptive Participants

Measure	Block 1	Block 2	Block 3
Hit rate			
Nonadaptive	.22	.24	.38
Adaptive	.30	.20	.64
False alarm rate			
Nonadaptive	.32	.27	.33
Adaptive	.30	.21	.37

short runway, although it is interesting to note that nonadaptive participants showed some rise in op-short from blocks 2 to 3 as well.

Consistent with the view that participants were trying to deal with the greater need to conserve the long runway when there was high demand (block 3, when there were many 747s), the false alarm rate also rose for all participants ($F(2, 82) = 5.5$, $p < .006$). However, there was neither a significant effect of participant type ($F(1, 41) < 1$) nor an interaction of participant type with block ($F(2, 86) = 1.2$, $p > .3$). Despite the suggestion of more false alarms for the adaptive participants in block 3, they were still much more sensitive in terms of d', overall, ($F(1, 41) = 6.6$, $p < .05$) and especially in block 3, where their d' was 1.0 ($\beta = 13.0$) and the nonadaptive participants had a d' of 0.2 ($\beta = 3.7$). In other words, the adaptive participants' higher hit rate did not come at the expense of an equal growth in false alarms; instead, they invested extra cognitive effort to use the short runway appropriately when necessary.

Individual Differences among the Nonadaptive

Thus far, we have examined which factors predict whether we classify a participant as "adaptive" or "nonadaptive" and what variables predict extent of adaptivity. However, it may also be informative to examine individual differences among the large group of nonadaptive participants. In particular, one could ask whether there are qualitative differences between those nonadaptive participants at the extremes of op-short strategy use. At the one end are those that usually opted to use the short runway when it was available (regardless of the proportion of 747s). At the other end are those that rarely opted to use the short runway when both were available (regardless of the proportion of 747s).

Rarely opting short Two nonadaptive participants were essentially at the floor in terms of tending to reserve the long runway for 747s (i.e., they opted for the short runway less than 10% of the time when both runways were available, across all three blocks). Interestingly, these participants performed almost as well as adaptive participants (i.e., much higher than the

Table 11.3 Mean Score, Keypresses, and Crashes per Block, as Well as Mean Time (in Seconds) Spent Reading Rule 4 during Instruction Phase and Mean Percentage Correct on the CAM Tests, by Participant Subgroup

Measure	Low op-short	High op-short	Unadapt. other	Adaptive
N participants	2	5	22	14
Score	3,768	2,243	2,501	4,104
Keypresses	2,563	2,260	2,528	3,276
Crashes	0.67	1.73	2.77	0.52
Study rule 6	35.0	54.7	43.7	51.9
Overall CAM	87.5	63.2	64.5	80.3

other nonadaptive participants). However, these participants were the same as the other nonadaptive participants on their number of keypresses. These participants scored fairly high on all of the CAM tests, and spent the least amount of time of all the participants reading the rule page that explained when one could land planes of different types (see table 11.3). Thus it appears that these were bright participants who did not bother encoding the rule for when to land planes (rule 4).

Frequently opting short Opting to use the short runway for DC10s requires much more cognitive effort than using the long runway in that the weather conditions dictate whether the short runway is allowable for these planes, whereas the long runway is always acceptable regardless of weather conditions. Participants were therefore classified in the "high op-short" category if their op-short use was greater than 50% across all blocks. Five participants were so classified. These participants performed as well as the other nonadaptive participants (see table 11.3) in terms of scoring fairly low on all of the CAM tests and on the KA-ATC. They spent the most amount of time of all participants on the rule page that explained when one could land planes of different types. Thus it appears that these not-so-bright participants learned the rules for when one could land planes but could not discover the strategic aspects of runway allocation.

Unadaptive other The largest subgroup of participants fell into the grab bag category "unadaptive other." One could argue that these were the most interesting of the nonadaptive participants. These participants knew the rules for when to land planes on the short runway and frequently used both strategies of conserving and not conserving the long runway. Yet they were simply unable to adaptively modify this strategy selection process in response to changing features of the environment.

Although the relative success of the rarely opting short subgroup appears problematic for the interpretation that the "optimal" behavior in this task is to adapt strategy use in response to our plane ratio manipulations, it is important to note that for block 2, the adaptive response is to use low levels of op-short. It is only in block 3 that higher levels of op-short become important. Correspondingly, the score differences between the rarely opting short and adaptive groups is smallest in block 2 (5,170 versus 5,437) and largest in block 3 (3,490 versus 4,102). Moreover, the two participants in that subgroup had very high CAM scores (i.e., strategy use and general ability are confounded). When the adaptive participants are matched with these participants on CAM scores, the difference between the groups becomes quite large (mean block scores of 3,768 versus 4,501). Thus adapting strategy use in response to the plane ratio manipulations truly is the optimal response.

11.8 GENERAL DISCUSSION

In this chapter we have attempted to build a bridge between the two traditional approaches to the study of individual differences—parameter differences, on the one hand, and strategy differences, on the other. We have argued that some differences in performance among individuals may be understood as differences in ability to adapt the selection of strategies to optimize performance in the local context. That is, some differences in performance may be a result of differences, not in people's repertoire of procedures (strategies) nor in their ability to execute a particular strategy, but rather in their ability to shift their strategy use in response to changing features of the environment. In other words, we have proposed a new kind of process parameter along which individuals might differ, and this parameter is cast in terms of strategy use. In this way, we have provided a link between the two traditional approaches.

In support of our speculations, we have found a situation in which people do differ in their strategy adaptiveness. In particular, we found that only a minority of our participants were able to adjust their op-short use appropriately in response to task changes that strongly called for shifts in op-short use. Participants differed not only in whether they shifted, but also in how much and how quickly they shifted.

An alternative interpretation to the strategy adaptiveness view is that the adaptive participants had more complex strategies rather than selecting among the same set of strategies more adaptively. For example, the adaptive participants may have had different explicit strategies for the different plane ratios (e.g., "If the ratio of 747s is high, then use the short runway whenever possible"). This alternative interpretation would make the observed differences consistent with the strategy differences view of individual differences. However, it is important to note that the adaptive participants never shifted their op-short levels in a binary fashion (e.g., from always using the short runway to never using the short runway). Instead, the participants merely changed the degree of runway use. This kind of continuous shift is much more consistent with changing ratios of strategy use than shifting from one strategy to another.

LCS

Client: MIT-U	J-8878 Gopher	Neuro 45	Palten_M	(.01) .29 .02 .05

In this study, a considerable portion of the full set of tested participants was excluded from the adaptivity analyses. Several factors motivated this exclusion. First, the excluded participants landed so few planes that it was not possible to reliably estimate their op-short levels and therefore to evaluate their adaptivity. Second, these participants performed abysmally—negative trial scores even after 90 minutes of practice, and showed little evidence of any improvement. It is likely that they were not motivated to learn the task at all, and thus should have been excluded even if they could have been evaluated on adaptivity of op-short use. Although the use of "temp agency" participants is the likely source of having so many poorly motivated participants, if one had to rate these participants on adaptivity, the label "extremely nonadaptive" seems most appropriate. What individual difference factors were associated with this group? To assess this question, we conducted a stepwise regression using the CAM factors as potential predictors. Only one factor entered into predicting which participants were excluded: inductive reasoning, "Figure Matrices" ($r = -.39$, $\beta = -.007$, $t(52) = 3.5$, $p < .005$). Thus inductive reasoning continues to be a good predictor at all levels of adaptivity.

For the remainder of the participants not excluded from the main analyses, we used a strict criterion for classifying participants as "adaptive" versus "nonadaptive," and using this strict criterion, we found that only a third of them were adaptive—a finding that is in sharp contrast to recent claims that people generally perform rationally or optimally (e.g., Anderson 1990; Reder 1987; Siegler and Shipley 1995). However, we do not wish to imply that our nonadaptive participants did not adapt at all. Adaptiveness is likely to be a continuum, and individuals below an arbitrarily chosen cutoff point are likely to also show some adaptivity. Indeed table 11.2 revealed that even the nonadaptive participants as a group did show a trend toward the desired strategy adaptiveness. Using a less strict criterion of simply increasing op-short levels from the second to third blocks, we found that fully three-quarters of the participants were adaptive. To ensure that the same individual differences measures were associated with this lower criterion of adaptivity, we recomputed the stepwise regressions. Again, only one factor entered into predicting adaptivity: the inductive reasoning test, "Figure Series" ($r = .49$, $\beta = 0.011$, $t(39) = 3.6$, $p < .002$). Thus adaptivity is a general characteristic, and inductive reasoning is predictive of adaptivity, even at the lower thresholds of adaptivity.

There are many conceptions of adaptivity. In this chapter, we have focused on adaptivity in strategy use, in particular adaptivity to changing base rates of strategy success. It is an open question as to how individual differences in strategy adaptiveness might correlate with other kinds of adaptiveness. For example, it may be correlated with individual differences in the ability to adaptively control attention (Gopher 1982, 1996; Gopher and Kahneman 1973) or in the ability to adapt to instructions (Reder 1987; Shebilske Goettl, and Regian, chap. 14, this volume).

How do our findings compare with other investigations of individual differences in the KA-ATC task? Lee, Anderson, and Matessa (1995) found that differences in overall strategy use accounted for a large proportion of performance differences, in contrast to our focus on differential strategy use. We view their work as complementary to ours in that we both emphasize the importance of strategies in the analysis of performance. However, they did not study the impact of within-subject manipulations, and thus did not have the opportunity to observe adaptivity differences.

Focusing on the relationship between predictors of performance and extent of training within the KA-ATC task, Ackerman (1988, 1989, 1990) found that different factors predicted performance at different phases of training: first reasoning ability, then perceptual speed, and finally reaction time ability. Although we did not have separate tests of perceptual speed and reaction time ability, we also found that reasoning ability and processing speed were important components of performance. One contribution of our study has been to provide a mechanism for how a cognitive skills assessment battery might predict performance: via differences in strategy adaptivity.

How general is the difference in adaptiveness we found in the KA-ATC task? Is adaptiveness a general trait that will hold across other tasks? We chose a dynamic task for our investigation because we thought it especially important to demonstrate adaptiveness in the context of a task that generally requires strategy adaptiveness (i.e., has frequent environment change). Indeed, we found that the same factors that predicted individual differences in adaptiveness also predicted overall performance in the beginning of the task, suggesting a common strategy adaptiveness component.

A remaining open question is whether the differences in strategy adaptiveness we observed are related to an enduring, stable individual difference factor or whether it is an unstable factor that varies with situational factors (e.g., with amount of training or interest in the task). That the differences in strategy adaptiveness correlated strongly with other stable individual differences factors (e.g., inductive reasoning skills) is suggestive of a stable adaptiveness factor, although not conclusive. What we have established is that individuals in a given task situation do vary significantly in their strategy adaptiveness and that strategy adaptiveness differences have large consequences for task performance.

A variation on the issue of stability of individual differences in adaptiveness is the role of levels of training. That is, are differences in strategy adaptiveness equally important at high performance levels as well as during early training phases? We found that our adaptive participants had the highest performance levels even in the first block, suggesting that strategy adaptiveness is important at early training levels. It may be, however, that at very high levels of expertise, differences in strategy adaptiveness become less important. For example, working memory loads of a task typically become lower with expertise, and thus the role that working memory capacity

Ucs

Client: MIT-U	J-8878 Gopher	Neuro 45	Palten_M	(.01) .29 .02 .05

played in rate of strategy adaptivity may be greatly reduced at high levels of training. Yet because it seems less clear that the role played by the inductive reasoning component of adaptivity differences would decrease with increasing expertise, we expect individual differences in strategy adaptivity to be present at all levels of expertise.

Another interesting issue in strategy adaptiveness is the role of awareness. Many researchers have assumed that metacognitive control comes with metacognitive awareness (Davidson, Deuser, and Sternberg 1994; Metcalfe 1994; Nelson and Narens 1994; although see Reder and Schunn 1996 for an alternative view). Perhaps the nonadaptive participants simply did not notice that the plane ratios had changed. Certainly, the participants were not explicitly told what the plane ratios were nor that the plane ratios would change. However, several pieces of information argue against this interpretation. First, pilot participants who were asked about the plane ratios were all aware of the changes. Second, our manipulation was far from subtle—there were large changes in plane ratios, and the difficulty of the task differed dramatically with the different plane ratios. Third, the participants had to encode the plane types in order to land the planes on the short runway, and all the participants did make some use of the short runway. Instead, it is likely that the ability to make use of base rate information *while* working on the task is the basis of adaptivity. Yet, with more subtle base rate manipulations, it is possible that awareness might play a more important role.

The practical importance of these findings seems clear: because individuals differ in how well they adapt their strategy use to changes in the task environment, one would want to select adaptive individuals for jobs in which the task environment frequently and rapidly changes (e.g., air traffic controllers). Performance in this modified version of the KA-ATC task was strongly related to adaptiveness—adaptive participants had much higher scores. This study also suggested which factors are likely to be good predictors of adaptiveness: inductive reasoning predicted whether people adapt, and working memory capacity and processing speed predicted how quickly people adapt. Of course, further research is required to establish the generality of these particular predictors. All of our predictors were positively correlated with one another, and our sample size was not very large. Thus, although inductive reasoning was consistently the best predictor of strategy adaptiveness, the predictors we found may not generalize as the consistently best predictors of adaptivity.

In particular, the stepwise regression analyses presented in this chapter have relied on good discriminability among the CAM measures. It is possible that a general *g* factor is the primary determinant of the adaptivity, and that the inductive tests merely load heavily on *g*. This possibility is especially likely given how well the adaptive participants performed overall. In further support of this hypothesis, a factor analysis on the CAM measures revealed a strong general factor accounting for 46% of the total variance and upon which all measures loaded positively. However, when the stepwise regressions were recalculated, it was still inductive spatial, and not the general g factor, that predicted adaptiveness.

A related issue is whether adaptiveness and performance are synonymous. In other words, have we predicted adaptiveness specifically, or simply overall performance, which happens to be correlated with adaptiveness? To test this possibility, we performed stepwise regressions predicting score within each of the blocks. Within the first block, two of the inductive spatial tests predicted score ($r^2 = .56$), namely, the inductive reasoning figure series ($\beta = 39.5$) and figure matrices ($\beta = 34.1$). In the second block, five factors predicted score, in order of predictiveness (multiple $r^2 = .79$): processing speed spatial ($\beta = 54.8$), skill learning spatial ($\beta = 45.2$), time spent reading the rule about when planes could land ($\beta = 25.6$), fact learning verbal ($\beta = -46.5$), and the "Figure Sets" inductive spatial test ($\beta = 25.2$). In the third block, two factors predicted score (multiple $r^2 = .54$): skill learning verbal ($\beta = 34.6$) and working memory spatial ($\beta = 22.4$). Thus, while inductive measures also predict score as well as adaptiveness, their predictiveness does not extended to the same region for which these measures predict adaptiveness. That is, inductive measures predicted adaptive behavior across the course of the study, with changing task characteristics. On the other hand, inductive measures predicted score only at the beginning of the study. It seems likely that these relationships between inductive reasoning measures and score in the earlier phases of the task—similar to Ackerman's findings (1988, 1990)—reflect initial adaptiveness to other aspects of the task, which then remain constant across the task.

Another possibility is that the adaptive subjects were simply more motivated than the nonadaptive subjects. In support of this hypothesis, adaptive subjects did have the highest keypress rates, as was shown in table 11.3, and this group difference was true of all three blocks. However, we found that even when partialing out the number of keypresses, the adaptive participants still obtained higher scores than nonadaptive participants. It appears that in a complex, dynamic task such as the KA-ATC task, the keystroke rate is more of a reflection of the ability to figure out quickly what to do next rather than a simple psychomotor factor. Providing further support for this interpretation, we found that one of the inductive spatial tasks, figure series, was the best predictor of total number of keypresses. These findings should be compared with Ackerman's finding that (1988, 1990) toward the end of the KA-ATC task, psychomotor speed becomes the best predictor of total keypresses and score. Perhaps our disparate findings can be explained by the fact that Ackerman held the structure the task constant across trials (i.e., did not manipulate plane ratio). Not only were our participants required to continually adapt, but they had less practice with any scenario and much less practice overall. Presumably, if we had held the task environment constant after block 3 and the participants had continued to practice, psychomotor speed would ultimately have become the best predictor.

One potentially interesting interpretation of the general positive correlation between strategy adaptiveness and each of the CAM measures is that the causal arrow goes from adaptiveness of CAM performance. That is, if there is such a thing as a general trait dimension of strategy adaptiveness, then one would expect individuals high on strategy adaptiveness to be able to quickly select good strategies for performing on cognitive ability tests. Taking this argument further, it maybe that inductive reasoning per se has nothing to do with predicting or enabling strategy adaptiveness. Instead, it may be that there are certain very useful strategies for solving the induction test problems that adaptive people are quick to find. This is an interesting possibility, which requires further exploration.

Assume for the moment, however, that inductive reasoning skill is the best predictor of adaptiveness, and processing speed and working memory capacity are the best predictors of rate of adaptation. What are the theoretical consequences of such relationships for our understanding of strategy selection mechanisms? Inductive reasoning could play a role in adaptiveness in at least two ways. First, it could be that inductive reasoning skill is related to being able to notice shifting patterns in the environment. In our case, this possibility seems unlikely because our manipulation was so heavy-handed that all the participants that we debriefed in a pilot study were explicitly aware of the manipulation. Second, inductive reasoning might be related to being able to quickly understand the relationship between a strategy and its effect, or diagnosing when a strategy is no longer appropriate. Further research is required to test this possibility, perhaps manipulating the transparency of the relationship between strategy and outcome.

The role of working memory capacity in rate of adaptation also has several possible causal chains. Presumably, it involves an increased ability to keep some information in mind while simultaneously performing the task. The theoretical framework for strategy selection that we presented in the introduction suggests that it is information about base rates (i.e., extrinsic information) that would be key to strategy adaptiveness. In a related fashion, processing speed could be related to adaptation rate by allowing dual tasking: keeping track of outcomes while working at the basic task. Ability to retain the recent set of outcomes would predict how quickly the pattern of change can be detected, and hence how fast one could adapt. Reasoning ability, in contrast, would predict whether, given a pattern, one understood what strategy to adopt for best performance with the new pattern.

In conclusion, we have tried to provide a bridge between the psychometric and strategy approaches to individual differences. This new approach has some of the advantages of each of the others. Like the strategy approach, it can be used to provide a detailed account of performance on any particular task. Like the psychometric approach, it should be able to account for correlated performance differences across many tasks and domains. In addition, we hope that this new approach will bring further insights to our understanding of strategy selection, psychometric testing, and individual differences.

APPENDIX

The verbal and spatial test of each factor were quite similar to one another in structure, with differences only in the items presented on the screen (i.e., words versus stick figures). For lack of space, we will only describe either the verbal or the spatial test.

Working Memory Capacity Verbal

Four-term ordering requires participants to relate what is described in three statements to the order of four key words later presented. Two of the sentences describe the order in which the four key words are to appear. The third sentence describes the sequence of the other two sentences by using category names. For example:

The ANIMALS come after the FURNITURE.

The cow does not come before the bird.

The chair does not come after the lamp.

After the three statements are presented, eight numbered alternatives appear on the screen, and participants are to select the number corresponding to the correct order. The correct response for this example would be: "Chair, lamp, bird, cow." Participants are allowed 15 seconds to make a response. With incorrect responses, participants are given feedback and are allowed to review the three sentences and alternative answers. This test contains twenty-four items.

Processing Speed Verbal

Two-term ordering requires participants to decide as quickly as possible whether two presented words conform to the sequence specified in a sentence at the top of the screen. Each sentence explains that world 1 will be either before or after word 2 (e.g., "The bird comes before the cow"). The test words are then presented in the middle of the screen. Participants are to respond as quickly as possible by typing "L" if the word order matches the sentence, and "D" if it does not. This test contains twelve items.

Fact Learning Spatial

Figures recognition requires participants to study and memorize 12 geometrical figures in a 3 × 4 matrix, and then determine whether individually presented figures were in that matrix. Participants are given 60 seconds to study the figures. During practice, participants are given the hint to try to make associations with what the figures may resemble (e.g., a letter, a flag). At test (beginning immediately after study), participants are shown individ-

Ա Ϛ

ual figures and asked whether it was one studied. This test contains two sets of figure matrices, asking twenty-six questions for each set.

Skill Learning Spatial

Reduction of circles presents participants with two circles that must be combined to form one circle, according to the following rules:

Rule 1: If both circles are solid (i.e., each is either solid white or solid black), then the combined circle will keep the black parts of both circles. For example, if one circle is solid white and the other is solid black, then the result is a solid black circle.

Rule 2: If either circle is a mix, then the combined circle will keep the white parts of both circles. For example, if one circle is solid black and the other is black on the left half and white on the right half, the result is a circle with the right half white.

Participants choose the answer from four numbered alternatives pictured at the bottom of the screen. This test consists of four sets of problems, each containing twenty-four items.

Inductive Reasoning Spatial

Figure sets presents participants with three sets of figures. Two of the sets will be related according to various themes. Participants must determine which set is the odd set. Various patterns include figures formed with straight lines as opposed to curved lines, internal shading versus no shading, and so on. In all, there are ten items for participants to solve within a 5-minute period.

Figure series presents participants with a series of shapes at the top of the screen and must choose the next shape occurring in the series from three numbered alternatives. For example, if the series was "/, *, //, ____," the answer would be "**". Participants must solve all ten problems within a 5-minute period.

Figure matrices presents participants with a 3×3 matrix in which a figure is contained in all but one of the cells. Participants must look at the figures and apply horizontal and/or vertical rules to determine what figure belongs in the empty cell. The matrix is shown on the screen at the same time the eight alternative responses are shown. Themes used include gradual shading of figures, successive additions or deletions to figures, rotation of figures, and so on. Themes are used in combination to form both horizontal and vertical rules for participants to induce. Participants have 10 minutes to solve all nine problems.

Ley

NOTES

The work reported here was supported by N00014-95-1-0223 from the Office of Naval Research. Data collection was made possible with the generous assistance of the (TRAIN) laboratory of AFOSR, and in particular Barry Goettl and J. Wesley Regian, Jr., of Armstrong Laboratory. We would like to thank Frank Lee for making available the code for his variant of the (KA-ATC) task, and John R. Anderson, Patrick Lemaire, John Tangney, Daniel Gopher, and two anonymous reviewers for comments on the manuscript. Authors are listed in alphabetic order and made equivalent contributions.

1. A participant's cumulative score is calculated as follows: 50 points for landing a plane; minus 100 points for crashing a plane; minus 10 points for violating one of the six rules that govern the task.

2. Only the direction of the difference (i.e., greater or less than zero) was used to differentiate the adaptive and nonadaptive groups; the magnitudes of the differences were not used.

3. Note, this last predictor is most likely spurious because it is the reading time for rules irrelevant to this aspect of the task and entered last into the stepwise regression.

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. (1989). Individual differences and skill acquisition. In P. L. Ackerman, R. J. Sternberg, and R. Glaser (Eds.), *Learning and individual differences*, pp. 165–217. New York: Freeman.

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., and Kanfer, R. (1994). University of Minnesota. *CD-ROM database, data collection program, and playback program.*

Anderson, J. R. (1982). Acquisition of a cognitive skill. *Psychological Review, 89*, 369–406.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Case, R. (1985). *Intellectual development: A systematic reinterpretation*. New York: Academic Press.

Cooper, L. A., and Regan, D. (1982). Attention, perception, and intelligence. In R. Sternberg (Ed.), *Handbook of human intelligence*, pp. 123–162. New York: Cambridge University Press.

Davidson, J. E., Deuser, R., and Sternberg, R. J. (1994). The role of metacognition in problem solving. In J. Metcalfe and A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing*, pp. 207–226. Cambridge, MA: MIT Press.

Farah, M. J., and Kosslyn, S. M. (1982). Concept development. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior*, vol. 16, pp. 151–186. New York: Academic Press.

Fitts, P. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning*, pp. 244–288. New York: Academic Press.

Gardner, H. (1983). *The mind's new science: A history of the cognitive revolution*. New York: Basic Books.

Gopher, D. (1982). A selective attention test as a predictor of success in flight training. *Human Factors, 24*, 173–183.

Gopher, D. (1996). Attention control: Explorations of the work of an executive controller. *Cognitive Brain Research, 5*, 23–38.

Gopher, D., and Kahneman, D. (1973). Individual differences in attention and the prediction of flight criteria. *Perceptual and Motor Skills, 33*, 1335–1342.

Hunt, E., Joslyn, S., and Sanquist, T. (1996). *Evaluating individual differences in response to emergency situations*. ONR Technical Report. Washington, DC: Office of Naval Research.

Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122–149.

Kail, R. (1988). Development functions for speeds of cognitive processes. *Journal of Experimental Child Psychology, 45*, 339–364.

Kanfer, R., and Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology—Monograph, 74*, 657–690.

Kyllonen, P. C. (1993). Aptitude testing inspired by information processing: A test of the four-sources model. *Journal of General Psychology, 120*, 375–405.

Kyllonen, P. C. (1994). Cognitive abilities testing: An agenda for the 1990s. In M. G. Rumsey, C. B. Walker, and J. H. Harris (Eds.), *Personnel Selection and Classification*, pp. Hillsdale, NJ: Erlbaum.

Kyllonen, P. C. (1995). CAM: A theoretical framework for cognitive abilities measurement. In D. Detterman (Ed.), *Current topics in human intelligence*: Vol. 4, *Theories of intelligence*, pp. 307–359. Norwood, NJ: Ablex.

Lee, F. J., Anderson, J. R., and Matessa, M. P. (1995). Components of dynamic skill acquisition. In the *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*, pp. 506–511. Hillsdale, NJ: Erlbaum.

Lemaire, P., and Reder, L. M. (~~forthcoming~~ in press). How do subjects control their strategy selection in arithmetic? The example of parity effects on product verification. *Memory & Cognition*

Lovett, M. C., and Anderson, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology*.

Lovett, M. C., Reder, L. M., and Lebiere, C. (1996) Modeling individual differences in a digit working memory task. Paper presented at the *International Conference on Memory*. Padua.

MacLeod, C. M., Hunt, E. B., and Matthews, N. N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior, 17*, 493–507.

Metcalfe, J. (1994). A computational modeling approach to novelty monitoring, metacognition, and frontal lobe dysfunction. In J. Metcalfe and A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing*, pp. 137–156. Cambridge, MA: MIT Press.

Nelson, T. O., and Narens, L. (1994). Why investigate metacognition? In J. Metcalfe and A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing*, pp. 1–25. Cambridge, MA: MIT Press.

Reder, L. M. (1979). The role of elaborations in memory for prose. *Cognitive Psychology, 11*, 221–234.

Reder, L. M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review, 89*, 250–280.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19*, 90–137.

u >

Reder, L. M. (1988). Strategic control of retrieval strategies. In G. H. Bower (Ed.), *The psychology of learning and motivation*, vol. 22, pp. 227–259. San Diego: Academic Press.

Reder, L. M., and Schunn, C. D. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In L. M. Reder (Ed.), *Implicit memory and metacognition*, pp. 45–78. Mahwah, NJ: Erlbaum.

Salthouse, T. A. (1994). Aging associations: Influence of speed on adult age differences in associative learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1486–1503.

Schneider, W., and Shiffrin, R. M. (1977). Controlled and automatic human information processing: 1. Detection, search, and attention. *Psychological Review, 84*, 1–66.

Schunn, C. D., Reder, L. M., Nhoyvanisvong, A., Richards, D. R., and Stroffolino, P. J. (1997). To calculate or not calculate: A source activation confusion (SAC) model of problem-familiarity's role in strategy selection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1–27.

Shute, V. J. (1993). A macroadaptive approach to tutoring. *Journal of Artificial Intelligence in Education, 4, 61–93.*

Siegler, R. S. (1988). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General, 117*, 258–275.

Siegler, R. S., and Shipley, C. (1995). Variation, selection, and cognitive change. In G. Halford and T. Simon (Eds.), *Developing cognitive competence: New approaches to process modeling*, pp. 31–76. New York: Academic Press.

Snow, R. E., Kyllonen, P. C., and Marshalek, B. (1984). The topography of ability and learning correlations. In R. Sternberg (Ed.), *Handbook of human intelligence*. New York: Cambridge University Press.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15*, 72–101.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning*. Hillsdale, NJ: Erlbaum.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Client: MIT-U	J-8878 Gopher	Neuro 45	Palten_M	(.01) .29 .02 .05
(IDP) TPS 23/7	PMU: WSL 30/7	pp. 315-342	Ch11_P	(p. 342) (V5)