(2)

**Individual Differences in Adaptive Processing in Complex Learning and Cognitive Performance**

Richard E. Snow,
Robert L. Chastain, and
Douglas N. Jackson III

School of Education
Stanford University





**Final Report**
**March 17, 1992**

To
U.S. Air Force Office of Scientific Research
Department of the Air Force
Bolling Air Force Base
Washington, D.C. 20332-6448
Grant No. AFOSR-89-0517

Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under grant or cooperative agreement number, AFOSR-89-0517. 

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No 0704-0188*



**1. AGENCY USE ONLY** *(Leave blank)*

**2. REPORT DATE**
17 March 1992

**3. REPORT TYPE AND DATES COVERED**
Final Report 1 Sep 89 To 30 Apr 91

**4. TITLE AND SUBTITLE**
Individual Differences in Adaptive Processing in Complex Learning and Cognitive Performance

**5. FUNDING NUMBERS**
Grant
AFOSR 89-0517
61102F
2313
A7

**6. AUTHOR(S)**
Richard E. Snow
Robert L. Chastain
Douglas Jackson III

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
School of Education
Stanford University
Stanford, CA 94305

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Dr John F. Tangney
U.S. Air Force Office of Scientific Research/NL
Building 410
Bolling AFB, DC 20332

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**
Unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 words)*

Theories of intelligence and learning ability emphasize individual differences in adaptation of information processing during performance on novel, changing tasks, but measures of adaptation have been lacking. This research sought to develop and evaluate such measures in samples of Air Force recruits. Computerized tests of verbal and spatial reasoning and of spatial-perceptual-motor performance were designed to provide scores reflecting raw and residualized differences between performance in homogeneous versus heterogeneous tasks. Reference tests represented conventional ability constructs. A complex learning task based on logic gates in electronics yielded criterion scores. Reliabilities of adaptation indices were moderate but within acceptable ranges. Low but significant correlations were obtained between some adaptation scores and some learning indices. Apparently, adaptation estimates can make a small but unique contribution to learning prediction. The research suggested further steps toward instrument improvement that seem clearly justifiable on both theoretical and practical grounds.

**14. SUBJECT TERMS**
Individual differences, adaptive processing, flexibility, learning, complex cognitive abilities, dynamic spatial abilities

**15. NUMBER OF PAGES**
22

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | (U) |

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

2 0 MAR

## OVERVIEW

This brief final report summarizes a series of studies designed to explore individual differences in complex learning and cognitive performance and, particularly, the role of flexible adaptation to changing task characteristics in such differences. The main aim of this research was to develop a more explicit assessment of the individual differences in adaptive processing involved in flexible performance, and to explore the degree to which such differences might account for correlations among complex ability and learning tasks. Two provisional hypotheses guided this exploratory work: 1) individual differences in flexible adaptation can be measured and analyzed as a common construct underlying several kinds of complex ability tasks, and can be distinguished from other constructs such as component processing skills, speed of processing, attentional resource allocation, or memory capacity; 2) this sort of flexibility remains important in later as well as earlier learning stages whenever learning tasks are inconsistent (in the sense of Ackerman, 1987) from item to item, or trial to trial, i.e. when they involve transfer.

The results reported here are promising but by no means conclusive. They suggest an approach to measurement of adaptive processing that might profitably be further developed and used in research on individual differences in learning. The hypothesis that individual differences in flexible adaptation represent a unique source of variance in learning, distinct from conventional ability differences, remains viable. However, difficult problems of reliability improvement and construct validation remain to be addressed. It may be that adaptive processing is to a significant degree specific to the task at hand and highly probabilistic.

Technical details pertaining to each study conducted in this research are not given in the present report. They are available in two separate reports; see Chastain (1992) and Jackson (1991).

## BACKGROUND

Massive evidence shows that general ability measures correlate strongly with individual differences in complex learning and cognitive performance (Snow, 1982). Recent research has sought to identify the underlying information processing components and strategies that might account for such relations (Snow & Lohman, 1989; Sternberg, 1985). However, most traditional definitions of intelligence also hypothesize that flexible adaptation of cognitive processing, particularly in novel situations, also plays an important role beyond the sum of elementary processing components and capacities (Snow, 1978). Kyllonen, Lohman, & Woltz (1984) sought to explore this possibility by demonstrating that several kinds of within-person strategy shifting exist in complex spatial performance. Other evidence also suggests that complex performance cannot be understood in terms of stable information processing organizations for different tasks (Bethell-Fox, Lohman, & Snow, 1984; Snow & Lohman, 1984; Snow, Kyllonen, & Marshalek, 1984). This hypothesis deserves further research because within-person flexibility needs to be understood and explicitly assessed in ability tests and learning tasks if improvements in theory and practical personnel testing and training are to be realized.

A fairly simple and direct way of considering information processing hypotheses about ability and learning differences is based on a radex model of ability and learning correlations. A two-dimensional radex is usually obtained when correlation matrices including representative cognitive tasks are subjected to nonmetric multidimensional scaling (see Snow, Kyllonen, & Marshalek, 1984). Tests or tasks are arrayed in this space so that proximity reflects strength of intercorrelation; the array resembles a dart board. The most prominent feature of this model of ability organization is the complexity continuum from the center to the periphery; more complex tasks appear toward the center while less complex tasks are distributed around the periphery. It is near the center where tests are highly loaded on G and ability-learning correlations are highest. The complexity continuum can be defined as an ordering of ability tests along a continuum according to their correlation with G (Marshalek, Lohman, & Snow, 1983). The key to a theory of intelligence is, thus, an understanding of what is increasing as one moves from test to test along this complexity continuum.

A number of hypotheses have been advanced to explain the sources of increasing complexity in this continuum: 1) an increasing number of processing components involved in task performance; 2) the increasing involvement of one or more central components, such as inductive reasoning; 3) an accumulation of speed differences in component processing; 4) increasing demand on attentional or memory capacities; and 5) increasing demand on adaptive functions, including executive or metacognitive control of these functions. It is possible that the complexity continuum may reflect combinations of many or all of these hypothesized sources. However, the focus of this research is on adaptive processing because, although it has been hypothesized to be at the heart of general ability differences and ability-learning relations, it has been relatively neglected in previous research. No viable approach to the measurement of such processing seems to have been developed.

## METHOD

Three experimental studies were conducted using United States Air Force recruits tested in the Air Force Human Resources Laboratory at Lackland Air Force Base, San Antonio, Texas. Paper-and-pencil test scores for these recruits were available from the Armed Services Vocational Aptitude Battery (ASVAB), including measures for Arithmetical Reasoning, Mathematical Knowledge, Numerical Operations, Coding Speed, Word Knowledge, Paragraph Comprehension, and General Science, as well as the Armed Forces Qualification Test (AFQT) total score. In addition, a battery of computer-based tests of processing speed, working memory capacity, and spatial ability were administered. Together with the ASVAB scores, this provided 16 reference ability measures. Recruits were also tested for approximately three hours with computerized tasks designed to measure complex learning and cognitive performance.

Computerized Learning Task

Learning was assessed using a logic gate task (Kyllonen & Woltz, 1989) in which subjects learn about three types of electronic input/output gates that operate with a set of logical rules. For this research, logic gate problems were given in four blocks of 72 problems each. The fourth block of logic gate problems introduced negative (reverse) gates where subjects had to reverse the input and/or output values mentally and then apply the appropriate rules. Learning measures were obtained directly from block performance and also as slope (or difference) scores reflecting increases or decreases in performance across the logic gate blocks. Of particular interest was the difference in performance between the third and fourth blocks, since this might reflect transfer adaptation differences most directly.

Computerized Performance Tasks

As noted, three measures of information processing speed were given; each measure represented a different content area (math, spatial, and verbal). Three measures of working memory capacity were given in the same manner, again using math, spatial, and verbal content. Two spatial reference measures were added because Lohman, Pellegrino, Alderton, and Regian (1987) have hypothesized that most complex spatial tasks require much flexible adaptation and therefore may be better characterized as figural reasoning tests with a spatial component rather than as measures of unique spatial processes. This battery of reference measures together with the ASVAB composites represented the four factor model of cognitive performance differences posited by Kyllonen and Christal (1989).

Our experimental measures of adaptive processing derived from three additional computerized performance tasks. One task was created using items from the Wonderlic Personnel Test (WPT, 1942),[1] a 12-minute paper-and-pencil test of 50 items. The WPT was chosen because it samples a broad range of problem types, such as verbal analogies, disarranged sentences, arithmetic computation, and geometric figure analysis, and presents them in intermingled order. We produced computerized test formats that minimized or maximized inter-item variation due to changing problem type and problem difficulty. To minimize inter-item variation, a more homogeneous format was constructed where similar problem types were grouped together in blocks. For example, all verbal analogy items would be administered first, then all arithmetic items, then all geometric figures, etc. To maximize inter-item variation, a more heterogeneous format was constructed in which problem types were alternated or mixed as much as possible so that no two adjacent items represented the same type of problem. It was hypothesized that a homogeneous (blocked) format would require less inter-item adaptation of processing strategies than would a heterogeneous (mixed) format. In the mixed format, subjects would need to adapt more frequently by shifting strategies between items; maximum performance would require flexible and efficient, strategy shifting. In a blocked format, such strategy shifting would be significantly reduced if not minimized.

The second computerized performance task was the Figure Encoding Test, (FET), developed along the lines of the Kyllonen, Lohman, and Woltz (1984) study which used Lohman's (1979) spatial visualization task. Computer-generated stick figures

---

[1]Permission to computerize the WPT was granted by E. F. Wonderlic Personnel Test, Incorporated.

were rated for labelability; some looked like objects that would be easy to label (i.e., a flag, an arrow, a question mark, a chair, etc.), whereas others were difficult to label. These figures were also categorized according to their complexity, based on scoring rules for the number of line segments, number of lines, number of subfigures and a combined symmetry score. The 96 chosen figures were then divided into four quadrants based on the two dimensions of labelability and complexity: easy to label and simple (ES); easy to label and complex (EC); hard to label and simple (HS); hard to label and complex (HC). The task was to remember an initial target figure and then to judge a second comparison figure as the same or different.

The rationale behind FET was to present four conditions that would afford choices among two strategies for remembering the figure, verbal labeling or feature analysis (analyzing the featural complexity for cues). Optimal performance should require shifting between strategies, a form of adaptive processing. The first condition, ES, was designed to be the easiest of the four conditions to accommodate labeling, feature analysis, or strategy shifting. Subjects could afford to choose and shift strategies easily. Then came the EC condition, designed to accommodate labeling but not feature analysis. Subjects who were labelers in ES could continue to use the same strategy in EC and show comparable performance. Subjects who were feature analyzers in ES would find the EC harder and show a decrement in performance. Subjects using both strategies in ES by shifting between the two should show comparable performance in EC. After the EC, HS was designed to afford feature analysis. Labelers should show poor performance in HS relative to ES or EC but feature analyzers should show increased performance relative to EC and comparable performance relative to ES. Shifters should continue to have comparable and high performance across all three conditions. Lastly, the difficult HC condition would not afford labeling, feature analysis, or shifting. In this sense, HC represents a baseline condition from which to compare performance across the other three conditions because subjects should score better across the first three conditions. Two adaptation scores were created: a) a difference score subtracting the HC score from the average of scores in the other three conditions; b) a residual score created by regressing the averaged score onto the HC score. Larger averaged scores across the first three conditions relative to the HC baseline score (i.e., larger positive differences) would indicate more flexible adaptation.

The third computerized test was a dynamic spatial coordination task (DST; Jackson, 1989), chosen for inclusion here because it favored the use of flexible strategies, included an interesting game-like character (explosions and color graphics), and other features that needed pilot investigation as part of another research project concerned with dynamic spatial judgment. The object is to shoot and hit a moving target from a stationary base. The complexity and flexibility requirements of the task were systematically manipulated by varying horizontal speed of the gun, vertical as well as horizontal movement of the target, the number and movement of barriers between the base and the target, and the addition of distracters, e.g., the target may fire at the base. The target can be regarded as a small spaceship that moves along the upper portion of the screen. The subject must shoot at the target by judging time-rate-distance relations while ignoring distractions. The dependent measure was the number of hits in different trials. There were 72 trials, 12 for each of the six problem types. The six problem types were constructed by varying task dimensions to yield versions that minimized or maximized inter-item variation (blocked versus mixed format). The varied dimensions were: speed of target (slow, medium, or fast); height of target (low or high); target vertical movement (none or up and down); barricades (none, 1 stationary, or three moving); distracter firing from target (no or yes); direction of target movement from (right or left). In the more homogeneous (blocked) condition, each trial was exactly the same within any one problem type. In the heterogeneous (mixed) condition, pairs of identical trials from each problem type were randomly selected with the condition that no two pairs of adjacent trials be the same problem type. Order of administration of blocked and mixed formats was randomly assigned before testing. Adaptation scores were derived from the simple difference between formats (mixed minus blocked) and a residual score created by regressing the mixed format score onto the blocked format score.

Subjects

All subjects were United States Air Force recruits in their sixth day of basic training at Lackland Air Force Base, San Antonio, Texas. There were 146 recruits in Study 1 (51% male, 75% single, 82% white, and 100% high school graduates) with mean AFQT percentile score of 62.90 (N=116). There were 103 recruits in Study 2 (66% male, 80% single, 81% white and 102 high school graduates) with average AFQT percentile score of 65.75 (n=85). There were 303 recruits in Study 3 (63% male, 83% single, and 79% white, and 301 high school graduates) with average AFQT percentile score of 65.75 (n=252).

## RESULTS

Reliability

Studies 1 and 2 were conducted primarily to guide instrument development and design administration procedures for the experimental WPT and FET tests. Of primary concern were the distributional properties and reliabilities of the adaptation scores. Although most WPT subtests showed adequate internal consistency reliability (using coefficient alpha), some showed floor or ceiling effects, and substantial inconsistency. Particularly troublesome were the antonyms and synonyms subtests of WPT. Item revisions and substitutions based on Studies 1 and 2 sought to rectify these problems.

Table 1.

Odd-Even Split-Half Reliability Estimates for Blocked, Mixed, Raw Difference (DIFF), and Residual Difference (RESID) Scores for Experimental Tests and Subtests in Three Studies.

| | Test | Score Type | | | |
|---|---|---|---|---|---|
| | | Blocked | Mixed | Diff | Resid |
| Study 1 (N=146) | WPT Antonyms | 35 | 37 | 26 | 35 |
| | WPT Spatial | 84 | 80 | 28 | 41 |
| | WPT Number Series | 73 | 65 | 17 | 31 |
| | WPT Syllogisms | 28 | 57 | 11 | 57 |
| | WPT Arithmetic | 80 | 78 | 60 | 68 |
| | WPT Total | 84 | 88 | 28 | 46 |
| Study 2 (N=103) | WPT Antonyms | 15 | 62 | 46 | 64 |
| | WPT Spatial | 72 | 77 | 29 | 47 |
| | WPT Number Series | 68 | 73 | 37 | 52 |
| | WPT Syllogisms | 52 | 52 | 13 | 37 |
| | WPT Arithmetic | 74 | 70 | 41 | 52 |
| | WPT Total | 86 | 86 | 48 | 57 |
| Study 3 (N=303) | WPT Total | 88 | 88 | 45 | 54 |
| | FET Total | 48 | 86 | 68 | 62 |
| | DST Total | 82 | 77 | 68 | 72 |

Note: Decimals omitted. Coefficients stepped up by Spearman-Brown formula.

These efforts were partially successful, although several additional points of needed improvement were noted for consideration in future work. Some further experimental changes were made in items and procedures for Study 3.

To obtain reliability estimates for the difference and residual scores --- the adaptation indices of primary interest --- an odd-even split-half method was used. Resulting coefficients for all three studies are shown in Table 1. For Studies 1 and 2, the table also includes estimates for subscores in both blocked and mixed conditions Given these reliability considerations, it was decided to focus attention in further work on WPT total scores rather than subscores, and on the residual scores rather than the raw difference score.

Validity

Residual scores for WPT, FET, and DST in Study 3 were not highly correlated. It was hoped that these scores for WPT and FET would intercorrelate sufficiently to indicate a common construct; r was .32 which is significant beyond the .01 level with N=303, and is also equal to the correlation between WPT and FET total scores. Corrected for attenuation in both measures, this r became .55. This relation is quite promising as an initial trial, but hardly sufficient to allow WPT and FET adaptation scores to be used as alternative indicators. The DST residual score was not expected to correlate highly with WPT or FET; r was .12 and .06 respectively (.12 just exceeds $r_{05}$ with N=303). Corrected for attenuation, these coefficients became .19 and .09, respectively.

Tables 2 and 3 provide correlations of WPT, FET, and DST total and residual scores with the reference ability measures and the learning measures, respectively. Also shown are correlations corrected for unreliability in the residual scores only.

The reference ability measures in Table 2 provided patterns of correlations that conformed roughly to expectations. Total and residual scores for WPT should correlate

Table 2.

Correlations of WPT, FET, and DST Total, Residual, and Disattenuated Residual Scores with Reference Ability Measures (N=303).

| | WPT | | | FET | | | DST | | |
|---|---|---|---|---|---|---|---|---|---|
| **Learning Measure** | **Total** | **Resid** | **DResid** | **Total** | **Resid** | **DResid** | **Total** | **Resid** | **DResid** |
| AFQT Score | 61 | 22 | 30 | 21 | 13 | 17 | 22 | 01 | 02 |
| Arithmetic Reasoning | 63 | 32 | 44 | 20 | 13 | 17 | 26 | 03 | 04 |
| Math Knowledge | 53 | 12 | 16 | 20 | 14 | 17 | 19 | 02 | 02 |
| Word Knowledge | 34 | 13 | 18 | 10 | 08 | 10 | 11 | -01 | -01 |
| Para. Comprehension | 17 | 05 | 07 | 08 | 08 | 10 | -05 | -08 | -10 |
| General Science | 33 | 11 | 15 | 10 | 15 | 18 | 17 | 05 | 06 |
| Numerical Operations | 19 | 16 | 22 | 01 | 00 | 00 | 00 | -06 | -07 |
| Coding Speed | 16 | 12 | 17 | 05 | 04 | 05 | 04 | -02 | -03 |
| Proc. Speed Math | -30 | -11 | -16 | -01 | 02 | 03 | -23 | -13 | -15 |
| Proc. Speed Verbal | -25 | -08 | -11 | 00 | -05 | -07 | -10 | -08 | -10 |
| Proc. Speed Spatial | -33 | -08 | -11 | -09 | -04 | -05 | -32 | -14 | -17 |
| Work Memory Math | 59 | 17 | 23 | 16 | 11 | 14 | 27 | 12 | 14 |
| Work Mem Verbal | 18 | 05 | 07 | -04 | -01 | -01 | 07 | 05 | 06 |
| Work Mem Spatial | 56 | 17 | 24 | 24 | 20 | 26 | 33 | 05 | 06 |
| Spatial Transform | 38 | 12 | 16 | 18 | 17 | 21 | 15 | 00 | -01 |
| Spatial Synthesis | 31 | 08 | 11 | 28 | 22 | 28 | 26 | 10 | 12 |

Note: Decimals omitted. $r_{05} = 11$ and $r_{01} = 15$.

Table 3.

Correlations of WPT, FET, and DST Total, Residual, and Disattenuated Residual Scores with Learning Measures (N=303).

| | WPT | | | FET | | | DST | | |
|---|---|---|---|---|---|---|---|---|---|
| **Learning Measure** | **Total** | **Resid** | **DResid** | **Total** | **Resid** | **DResid** | **Total** | **Resid** | **DResid** |
| Logic Gates | | | | | | | | | |
| Block 1 | 51 | 11 | 15 | 32 | 19 | 24 | 32 | 11 | 13 |
| Block 2 | 49 | 12 | 16 | 32 | 21 | 27 | 30 | 11 | 12 |
| Block 3 | 45 | 08 | 11 | 31 | 22 | 29 | 29 | 09 | 10 |
| Block 4 | 54 | 18 | 25 | 37 | 24 | 31 | 24 | 03 | 04 |
| Total | 58 | 17 | 22 | 39 | 28 | 31 | 30 | 08 | 09 |
| Slope 12 | -04 | 00 | 00 | 00 | 04 | 05 | -03 | -02 | -02 |
| Slope 123 | -11 | -08 | -10 | -03 | 05 | 07 | -07 | -04 | -05 |
| Slope 34 | 37 | 20 | 27 | 24 | 13 | 16 | 12 | 00 | 00 |

Note: Decimals omitted. $r_{05}$ = 11 and $r_{01}$ = 15.

more highly with complex ability measures (e.g., arithmetic reasoning or mathematical knowledge) than with measures of simpler abilities (e.g., numerical operations or coding speed). Correlations with processing speed should be significantly negative; the speed measures are expressed as reaction times. It is noteworthy that two of the three working memory measures showed substantial correlation with WPT; it is not clear why the verbal content measure gave much lower correlations. FET and DST yielded patterns similar to WPT, but with lower correlations on average. Both FET and DST showed their highest correlations with spatial reference measures. However, these correlations still leave much specific variance in FET and DST distinct from conventional spatial measures.

Not shown in Table 2 are separate correlations for blocked and mixed score components of the total and residual scores. For both WPT and FET, mixed scores correlated about as high or higher with reference ability measures as did blocked scores. DST showed lower correlations with reference measures than did WPT and FET on average, with blocked scores often yielding higher correlations than mixed scores. Residual scores in Table 2 yielded eleven significant correlations for WPT, eight for FET, and three for DST. These were usually lower in value than the corresponding blocked and mixed score correlations, which in turn were usually slightly lower than the total score correlations.

The learning task gave several interesting results even though correlations with its slope scores may have been limited by ceiling effects; average performance on the four learning trial blocks was 81, 89, 92, and 77 percent correct, respectively. Nonetheless, as shown in Table 3, all three experimental tasks yielded substantial correlations with logic gate block performance. The total score correlations were strongest. The residual score correlations were often statistically significant; those for FET were particularly noteworthy. Correlations with the difference between Blocks 1 and 2 and the slope across Blocks 1, 2, and 3 were mostly near zero. The exception was the relation of WPT to Slope 123. In Table 3, the -.11 for WPT total is statistically significant, and the -.08 and -.10 show that the residual scores deserve some attention. These correlations are negative presumably because of ceiling effects in Blocks 2 and 3. Those subtests high in WPT total and residual did particularly well in Block 1, therefore gained less in Block 2, and again in Block 3; the result would be lower gain scores and slopes for these initially high performers, relative to those lower in WPT total, residual, and Block 1 performance.

Most noteworthy were the relations of WPT and FET with the difference in performance between Blocks 3 and 4 (Slope 34). This learning score was intended to reflect recovery from or adaptation to the shift in learning from positive to negative logic gates. This kind of adaptation process should be central to effective transfer. As shown in Table 3, the correlations of WPT and FET scores with this learning-transfer measure were mostly moderate in size but highly significant. They clearly deserve further study.

Table 4 reports the results of multiple regression analyses using the total scores from the three experimental measures. The first column shows $R^2$ for a reference predictor battery composed of AFQT, processing speed, working memory, and spatial ability measures; this battery can be interpreted as providing the best available prediction from conventional tests. Against this standard, WPT total scores added significantly to prediction whereas WPT residual scores did not. FET scores, on the other hand, offered significant additions to prediction from total scores but also from residual scores, at least on Blocks 3, 4, and Total Learning Score. DST total scores also added in predicting Block 1, 2, and 3 performance. The addition to $R^2$ from .44 to .48 in predicting Total Learning Score by adding WPT or FET, or to .51 by adding all extra predictors, is certainly a substantial improvement in criterion variance accounted for. However, since the residual scores seem to play a minor role in this, it is not at all clear that the adaptation aspect of the experimental measures is the source of this improvement. In any event, the strong showing of WPT and FET total scores in this multiple regression analysis deserves follow-up. Not shown here are further analyses distinguishing mixed and blocked scores in multiple regressions. These analyses suggested that both contribute to prediction. Mixed scores seemed to offer slightly better prediction than did blocked scores, especially for FET.

Table 4.

Percentage of Variance ($R^2$) Accounted for by Reference Battery and Experimental Predictors (N=303).

| | Reference | ADD ONLY | | | | | | ADD |
|---|---|---|---|---|---|---|---|---|
| | Predictor | WPT | | FET | | DST | | ALL |
| **Learning Measure** | **Battery** | **Total** | **Resid** | **Total** | **Resid** | **Total** | **Resid** | |
| Logic Gates | | | | | | | | |
| Block 1 | 31 | 35* | 31 | 33* | 31 | 33* | 31 | 37* |
| Block 2 | 29 | 34* | 29 | 32* | 30 | 31* | 30 | 36* |
| Block 3 | 22 | 27* | 22 | 26* | 24* | 24* | 22 | 30* |
| Block 4 | 39 | 44* | 39 | 44* | 40* | 39 | 39 | 46* |
| Total | 44 | 48* | 44 | 48* | 45* | 45 | 44 | 51* |
| Slope 12 | 03 | 03 | 03 | 03 | 03 | 03 | 03 | 03 |
| Slope 123 | 04 | 04 | 04 | 04 | 04 | 04 | 04 | 04 |
| Slope 34 | 19 | 21* | 20 | 21* | 19 | 19 | 19 | 23* |

* = Significant increase in $R^2$ compared to Reference Predictor Battery at .05 level.

## DISCUSSION

Virtually all cognitive tasks are inconsistent across items or trials in their information processing requirements. Most do not require that one mental program be applied iteratively. Rather, they typically require some degree of adaptation of that program from item to item or trial to trial. Novel tasks may require the assembly and adaptation of entirely new processing programs, and complex, heterogeneous, changing tasks may require this assembly and adaptation process repeatedly. Even tasks that are consistent involve adaptations within or between items or trials as learners acquire the ability to perform them. Thus, adaptation of information processing within or between tasks or parts of tasks is the norm, not the exception. Adaptations will often be qualitative, as in shifts of strategy, rather than quantitative, as in speed changes. Given this, it can be expected to be extremely difficult to obtain isolated measures of adaptation processes and their effects. It is not surprising that this problem remains unsolved, even though most theories of intelligence claim adaptation of cognitive processing as a central aspect of performance.

The present research sought to develop and evaluate such measures by designing contrasts between homogeneous or blocked tasks and heterogeneous or mixed tasks. The approach assumed that although both kinds of tasks would require adaptive processing, the heterogeneous mixed versions would place heavier demands on this kind of processing. Residual scores would capture this difference; these were obtained by partialling blocked performance out of mixed performance statistically. The result was an admittedly gross estimate with modest reliability. Since blocked performance also involves some adaptational functions, the residualizing process overcorrects. Since the result is a difference score, it reflects unreliability in both entering scores.

Given these weaknesses, the procedure nonetheless produced a faint signal; we conclude that it <u>is</u> a signal. Residual scores for two of the three experimental tasks were substantially correlated. Both total and residual scores showed correlation patterns with reference abilities that were reasonable and interpretable; correlations were neither so low nor so high as to cast doubt on the usefulness of the adaptation scoring approach. Correlations with learning measures and multiple regression analyses based on these were encouraging. Significant correlations were obtained between some adaptation scores and some learning indices and adaptation estimates made small but unique contributions to learning prediction. Of particular note was the relation of adaptation estimates to learning differences under conditions where a shift to more difficult, transfer trials had to be negotiated.

If these results are accepted as demonstrating that measures of adaptation are possible and that they may have theoretical and practical empirical value, then the call for further research in this direction is obvious. Several suggestions for that work can be derived here.

1. Improvement in reliability is a necessary first step. Lengthening the tasks, clarifying the instructions, and ensuring motivated performers are obvious steps. The present work was conducted with time constraints, a captive audience, and tasks that were rather unusual. A more comprehensive generalizability study of both test facets and testing conditions might help pinpoint improvements in instrument design.

2. The contrast between homogeneous and heterogeneous tasks could be made more extreme, especially as tasks are lengthened. The WPT should not be composed of different subtests, or perhaps subtests should be lengthened and separated into distinct tasks for blocked conditions. More radical shifts might be included. Similarly, another version of FET might be designed to provide a parallel to contrast WPT by interspersing items from different parts; the present version pursued only one of several hypotheses arising from previous research on strategy shifting.

3. More analytic research on within-item as well as between-item adaptation in these or related tasks might also suggest sharpened contrasts. Eye-movement tracking, for example, has shown adaptive variations dependent on interactions of student ability, item type, and item difficulty. One might be able to devise computer-adaptive tests that iteratively focused in on particular subject's adaptation processes in particular kinds of items.

4. Further research should also bring multiple approaches to adaptation assessment together. One approach is the mixed-blocked residualization method exemplified here. A second is exemplified by Sternberg's (1977; 1985) componential analysis, in which special component skills are estimated separately and a "wastebasket" parameter contains the residual, including variance due to adaptations; it is noteworthy that in several studies this parameter produced higher relations with complex reference tests than did the component parameters. Other approaches might include adaptation indices built up for individual subjects from eye track data, think-aloud protocols, and observational measures of performance. It is likely that there are different kinds and forms of adaptation processes. It is also likely that these processes are highly task or situation dependent, and may be highly probabilistic. Designing and interpreting multitrait-multimethod studies in this domain is hardly a straight-forward exercise.

In conclusion, we urge that further research of the sort outlined be undertaken. The present results are hardly conclusive, but they exhibit trends that are not easily dismissed or attributed to other sources despite the methodological problems of the present approach.

## REFERENCES

Ackerman, P.L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. Psychological Bulletin, 102, 3-27.

Bethell-Fox, C.E., Lohman,D.F., & Snow, R.E. (1984). Adaptive reasoning: Componential and eye movement analysis of geometric analogy performance. Intelligence, 8, 205-238.

Chastain, R. L. (1992). Adaptive processing in complex learning and cognitive performance. Unpublished doctoral dissertation, Stanford University, Stanford, CA.

Jackson III, D. N. (1991). Individual differences in performance on dynamic spatial tasks. Unpublished technical report, Stanford University, Stanford, CA.

Kyllonen, P.C., Lohman, D.F., & Woltz, D.J. (1984). Componential modeling of alternative strategies for performing spatial tasks. Journal of Educational Psychology, 76, 1325-1345.

Kyllonen, P. C., & Woltz, D. J. (1989). Role of cognitive factors in the acquisition of cognitive skill. In R. Kanfer, P. L. Ackerman, and R. Cudeck (Ed.), Abilities motivation & methodology (pp. 239-280). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kyllonen, P. C., & Christal, R. E. (1989). Cognitive modeling of learning abilities: A status report of LAMP. In R. Dillon and J. W. Pellegrino, Testing: Theoretical and applied issues. New York: Freeman.

Lohman, D.F. (1979). Spatial ability: A review and reanalysis of the correlational literature (Tech. Rep. No. 8). Stanford, CA: Stanford University, Aptitude Research Project, School of Education. (NTIS No. AD-A075 972).

Lohman, D.F., Pellegrino, J.W., Alderton, D.L., & Regian, J.W. (1987). Dimensions and components of individual differences in spatial abilities. In S.H. Irvine & S.E. Newstead (Eds.), Intelligence and cognition: Contemporary frames of reference (pp. 253-312). Dordrecht, The Netherlands: Martinus Nijhoff.

Marshalek, B., Lohman, D.F., & Snow, R.E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. Intelligence, 7, 107-128.

Snow, R.E. (1978). Theory and method for research on aptitude processes. Intelligence, 2, 225-278.

Snow, R.E. (1982). Education and intelligence. In R.J. Sternberg (Ed.), Handbook of human intelligence (pp. 493-585). Cambridge: Cambridge University Press.

Snow, R.E., Kyllonen, P.C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R.J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 2, pp. 47-104). Hillsdale, NJ: Erlbaum.

Snow, R.E., & Lohman, D.F. (1984). Toward a theory of cognitive aptitude for learning from instruction. Journal of Educational Psychology, 76, 347-376.

Snow, R.E., & Lohman, D.F. (1989). Implications of congitive psychology for educational measurement. In R.L. Linn (Ed.), Educational measurement (3rd ed.)(pp. 263-331). NY: Macmillan.

Sternberg, R.J. (1977). Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, NJ: Erlbaum.

Sternberg, R. J. (1985). Beyond IQ. New York: Cambridge University Press.

Wonderlic, E. F. (1942). Wonderlic Personnel Test. Northfield, IL: E. F. Wonderlic Personnel Test, Inc.