The Predictive Validity of Eye Movement Indices for Technical School Qualifying Test Performance

Ronna F. Dillon
Southern Illinois University at Carbondale

Robert A. Wisher
Navy Personnel Research and Development Center

The study was designed to determine the predictive validity of a series of eye scan indices, collected during analogical reasoning, with respect to technical school qualifying test performance. The scan patterns of 11 male Navy recruits were recorded during solution of 12 items taken from the 1962 Advanced Progressive Matrices. Significant positive relationships with performance on the Armed Forces Qualifying Test were identified for (1) the percentage of the total number of fixations occurring within the main stimulus array prior to the first excursion to the distractor field; (2) the proportion of total row scans to binary row scans; and (3) the proportion of total column scans to binary column scans. Significant negative relationships with qualifying test performance were identified for (1) the number of excursions from a point in the stimulus array to a point in the distractor field and (2) the total number of fixations. Results are discussed within the framework of employing indices of cognitive processes to develop selection and training procedures from an individual differences perspective.

A need has recently been identified for analysis of individual differences in cognitive abilities in terms of underlying information-processing capabilities (Pellegrino & Glaser, 1979). Federico and Landis (1979) noted that such indices would be particularly illuminating if information-processing capabilities were examined with respect to their relative contributions to explained variance in different indices of professional and technical school performance as well as later "on-the-job" success.

Although simple analyses of outcomes (i.e., test scores) provide information regarding the products of performance, such analyses do not elucidate intra-item solution processes. Sternberg (1977, 1979) has pointed out the need for such information if understanding of individual differences in cognitive abilities—abilities which are believed to underlie intelligent behavior—is to be enhanced.

The investigation reported herein was undertaken as an initial step toward understanding how intellectual performance can be improved and how prediction of such performance may also be enhanced through analysis of the sequential distribution of eye fixations executed during predictor task performance. A necessary step toward this goal comes from an interpretation of individual differences in terms of the processes that enhance or retard performance. A series of scanning indices were identified during solution of analogical reasoning items (Raven, 1962), each of which reflected one or more underlying psychological mechanisms. Figural analogies were chosen because of the important role in intellectual behavior ascribed to analogical reasoning (Miller, 1960; Polya, 1957; Reit-

men, 1965; Sternberg, 1977). The data were analyzed with respect to the magnitude of the relationships of the scanning indices collected during solution of the matrices to relevant aptitude. With respect to performance on the aptitude measure of interest, the Armed Forces Qualifying Test (AFQT; Swanson, 1978), the scanning indices collected during solution of figural analogies were hypothesized to possess significant predictive validity. Indices bearing predictive power would then be used for later data collection to form models of the scan predictors for use in multiple regression analyses with a greater number of subjects.

Method

Sample

The sample consisted of 11 male Navy recruits, 18 to 21 years old. All subjects had normal vision without corrective lenses.

Stimuli

Twelve items were taken from the Advanced Progressive Matrices (Raven, 1962). The items were 3×3 figural analogies with 8 distractors, one of which correctly completed the stimulus array. Each item was reproduced on a 35mm slide.

Instrument

All subjects were given the AFQT, a subset of tasks taken from the Armed Services Vocational Aptitude Battery (ASVAB; Swanson, 1978). The AFQT is used in selection and initial assignment of recruits. The three ASVAB tests composing the AFQT are Word Knowledge, a 30-item vocabulary test; Arithmetic Reasoning, a 20-item arithmetic test requiring the examinee to generate a principle for solution of each problem; and Space Perception, a 20-item pictorial test in which examinees must select a pattern to form a three-dimensional stimulus figure. The raw scores on each subtask are converted to a score

with a mean of 50 and a standard deviation of 10. Subjects are instructed to work through the tasks at a rate that they find comfortable. Each task typically takes approximately 40 to 50 minutes to complete.

Procedure

Each slide was presented to the subject on a rear-screen projector. A Whittaker oculometer system, Model 1992, was used to track eye fixations. The subject positioned his head against a chin and forehead rest and viewed an eye-level screen, 57cm away. The system computes the subject's point of regard from two reference points by monitoring corneal and pupil reflection 60 times per second, at 60Hz. A horizontal and vertical crosshair were superimposed on the video output of each Raven item. The intersection of the lines defined the fixation point. This information was videotaped during problem solution and was input to a PDP-11/70 for data analysis.

Subjects were tested individually, following a brief calibration procedure. The order of presentation of the 12 Raven items was varied randomly across subjects. Prior to the onset of each trial, subjects were instructed to fixate a point at the center of the screen. Upon a signal from the experimenter, the subject was instructed to solve the problem. A trial was terminated when the subject reported the number corresponding to the chosen response alternative.

Results

A pilot session was conducted with a separate group of subjects. Examination of the video-taped eye movement records and thinking-aloud protocols recorded during solution of the Raven items led to delineation of five scanning indices. The indices were

1. Percentage of the total number of fixations occurring within the main stimulus array prior to the first excursion to the distractor field;

2. Number of excursions from a point in the stimulus array to a point in the distractor field;
3. Total number of fixations;
4. Ratio of total row scans (i.e., $\leqslant 3$ successive fixations between 3 consecutive points in a row) to binary row scans (i.e., $\geqslant 3$ successive fixations between 2 points in a row); and
5. Ratio of total column scans to binary column scans.

Total row or column scans were not included in the count of binary row or column scans. These scanning indices then were used to assess the scan patterns of the experimental subjects.

Table 1 displays the means and standard deviations for each of the scanning indices measured during solution of the Raven items, and score on the Raven and AFQT items. Each index is related to efficient (i.e., nonredundant) scanning and reflects one or more underlying psychological mechanisms.

Correlational analyses were used to ascertain the relationship of the scanning indices to AFQT performance. The results are presented in Table 2. The data indicate the presence of a relationship of the percentage of fixations in the main stimulus array occurring prior to the first distractor excursions during Raven solution and AFQT performance, $r = -.70$, $p <.05$. A negative relationship was found between number of distractor excursions during Raven solution and AFQT performance, $r = -.73$, $p <.05$. A negative relationship was also found between total number of fixations during Raven solution and AFQT score, $r = -.59$, $p < .05$. The proportion of total row scans to binary row scans was related to success on the AFQT, $r = .66$, $p < .05$; a similar tendency was identified for the proportion of total column scans to binary column scans, $r = .53$, $p < .10$.

Discussion

The five scanning indices constitute powerful predictors of performance on the AFQT. The percentage of fixations occurring prior to the

Table 1

Means and Standard Deviations for Scanning
Indices, APM Score and AFQT Performance
for 12 Figural Analogies

Variable	N	Mean	Standard Deviation
% Fixations prior to Distractor Excursion	11	21.25	12.00
Total Fixations	11	720.55	308.18
Number of Distractor Excursions	11	51.18	22.74
Total Row/Binary Row	11	0.74	0.33
Total Column/ Binary Column	11	0.46	0.35
Raven Progressive Matrices Items	11	9.55	2.11
Armed Forces Qualifying Test	11	177.00	20.92

initial distractor excursion, and the proportion of total row scans and total column scans to binary row scans and binary column scans, were positively correlated to success on the AFQT. Significant negative relationships were identified between the number of distractor excursions and AFQT score, and between the total number of fixations during solution of the Raven items and AFQT score.

The percentage of fixations occurring prior to the first excursion to the distractor field reflects the extent to which examinees form complete hypotheses regarding the combined rules governing the simultaneous transformations between elements in the main stimulus array. Individuals scoring high on the AFQT were more likely to form a template of rules in the first and second rows of the stimulus array and to apply this template to the remaining portions of the array—to solve the item and then search the distractor field for the correct choice. Average ability subjects engaged in a greater percentage of early excursions (and a greater number of excursions, as noted below). Average ability subjects were more inclined to test hypotheses about single attributes or single transformations between stimulus elements (i.e., to exhibit attribute testing as opposed to hypothesis testing).

The number of distractor excursions reflects the frequency with which the subject seeks information from the distractor set. Although information relevant to encoding of attributes, rule formation, and rule application may be derived by processing portions of the stimulus array, the distractor set provides only confirmatory information. The relationship of AFQT performance to the point in intra-item solution at which the initial distractor excursion occurs (i.e., the previous index) indicates either the point at which confirmatory information is first sought (by attribute or hypothesis testing) or the examinee's attempt to gain rule-relevant information, rather than strictly confirmatory information, from the distractor set. Similarly, the number of distractor excursions reflects the frequency with which the examinee abandons on-

Table 2
Correlation Coefficients for Scanning Indices, APM Score, and AFQT Score.
Probability $|R|$ Under H_0: Rho=0

	Fixations prior to Distractor Excursion	Total Fixations	N Excursions	Total Row/ Binary Row	Total Column/ Binary Column	AFQT
Advanced Progressive Matrices	0.81**	-0.83**	-0.91**	0.61*	0.42	0.92**
Fixations prior to Distractor Excursion		-0.68*	-0.88**	0.48	0.34	0.71*
Total Fixations	-0.68*		0.84**	-0.54	-0.09	-0.59
N Excursions	-0.88**	0.84**		-0.57*	-0.21	-0.73*
Total Row/ Binary Row	0.48	-0.54	-0.57*		-0.04	0.66*
Total Column/ Binary Column	0.34	-0.09	-0.21	-0.04		0.53
Armed Forces Qualifying Test	0.71*	-0.59	-0.73*	0.66*	0.53	

* $p < .05$
** $p < .001$

going processing either to conduct a test or in an attempt to derive rule-relevant information outside the main stimulus problem. Again, protocol examination indicates that frequent distractor excursion reflects the testing of single attributes or the trial-and-error search for something that "looks right."

All items used embodied two transformations between elements in a row or column. A given number of fixations are necessary to derive the requisite information for item solution. Fixations in excess of this number provide no new information and thus are redundant. The finding that high AFQT ability was negatively related to number of fixations on the figural analogies reflects the tendency on the part of lower ability subjects to refixate a portion of the stimulus array, either because of inefficient processing (i.e., the inability to extract sufficient information from a single scan of a particular portion of the item) or because of the examinee's lack of confidence in his/her ability to process information efficiently.

The ratios of total row and total column scans to binary row and binary column scans were both significantly related to AFQT ability. The large versus small scan indices reflect "chunk" size and corroborate existing evidence (see Norman, 1976) that efficient processing of information is marked by an ability to deal with stimulus elements, not as isolated bits of information, rather as meaningfully interrelated units. Individuals executing a relatively greater ratio of large scans on the figural analogy items may experience success on the AFQT because of an ability to integrate material into larger functional units.

Eye scan data should prove useful in interpreting the theoretical meaning of Raven scores. To the extent that Raven items tap similar logical operations to other analogical reasoning items and to items on other inductive reasoning tasks, interpretable consistencies in eye scans might be expected to emerge. In this way, eye scan data may offer a useful means of construct validation.

Corroborative evidence in support of factor analytic data may be gathered.

Further research across different subject groups also may be designed to permit systematic examination of individual differences in scanning patterns. For example, different individuals may be identified who exhibit different scan patterns or constellations of scanning indices. Individual differences of this sort may then be related to reference ability measures, indicating the possibility of identifying, for example, "row scanners" versus "column scanners."

Scanning indices have considerable promise with respect to their predictive validity vis-à-vis simple outcome information resulting from a single test score. This initial investigation was designed to isolate and to test the predictive power of a series of scanning indices. A product measure, in the form of a test score, cannot reveal intra-item solution processes. The eye movement analysis reported herein permits examination of the processes underlying efficient analogical reasoning. The sequential distribution of eye fixations define the efficiency indices isolated. Efficient problem solving is a result of several processes; the scanning indices reflect these processes. As a single product measure would not capture the dynamics of problem solving, a single scanning index does not provide a complete picture of information processing.

With the pilot nature of this study in mind, further research is underway to establish the robustness of the technique. The extent to which the scanning indices provide enhanced predictive information when scanning is examined during solution of a range of reasoning tasks is under investigation. The extent to which resulting process information is useful in predicting performance on a variety of criterion measures is being investigated across a range of subject groups.

Multiple criterion measures are being selected for further research efforts to provide evidence of divergent, as well as convergent, processes. In the former regard, efficient scanning of a parti-

cular type should be associated with successful performance on one criterion measure but not on a second criterion believed to require disparate operations.

Further research is aimed at determining the relative contributions to explained variance on the AFQT and other aptitude measures made by each of the scanning indices isolated in this investigation. Further research is also designed to determine the relative proportions of variance in technical school and college performance, assessed by school grades and teacher assessments of mastery, accounted for by the scanning indices presently identified. Eye scan data will be recorded during solution of competency-related tasks. As the scanning indices reflect basic cognitive processes, their predictive power should outstrip simple outcome analyses, irrespective of the task during which they are recorded.

The implications of these findings for selection and training are clear and promising. Constellations of eye scan components provide information regarding the manner in which individuals solve complex cognitive tasks. Particular modes of processing may be likely to result in optimal performance for tasks having specific demands. If so, individuals exhibiting appropriate processes may be most likely to experience job success, although an analysis of test score may not predict success. Under conditions where training is warranted, eye movement data provide information regarding specific processing strengths and deficiencies. Thus, process relevant training prescriptions can be developed from an individual differences perspective.

For diagnostic situations in which eye scan recording is not possible, the sensitivity of selection procedures as well as the understanding of basic psychological processes may be positively affected by making use of other process data, as for example, through protocol generation and analysis, and by examining subtask data, as in the work of Sternberg (1977).

References

Federico, P. K., & Landis, D. B. Discriminating between failures and graduates in a computer-managed course using measures of cognitive styles, abilities, and aptitudes. (NPRDC Technical Report 79-21). San Diego, CA: Navy Personnel Research and Development Center, June 1979.

Miller, W. S. *Manual of the Miller Analogies Test.* New York: Psychological Corporation, 1960.

Norman, D. A. *Memory and attention* (Vol. 2). New York: Wiley & Sons, 1976.

Pellegrino, J. W., & Glaser, R. Cognitive correlates and components in the analysis of individual differences. In R. J. Sternberg & D. K. Detterman (Eds.), *Human intelligence: Perspectives on its theory and measurement.* Norwood, NJ: Ablex, 1979.

Polya, G. *How to solve it.* Garden City, NY: Doubleday-Archer, 1957.

Raven, J. C. *Advanced Progressive Matrices, Set II.* London: H. K. Lewis, 1962.

Reitmen, W. *Cognition and thought.* New York: Wiley, 1965.

Sternberg, R. J. *Intelligence, information processing, and analogical reasoning: A componential analysis of human abilities.* Hillsdale, NJ: LEA, 1977.

Sternberg, R. J. Intelligence research at the interface between differential and cognitive psychology: Prospects and proposals. In R. J. Sternberg & D. K. Detterman (Eds.), *Human intelligence: Perspectives on its theory and measurement.* Norwood, NJ: Ablex, 1979.

Swanson, L. Armed Services Vocational Aptitude Battery Forms 6 and 7: Validation against school performance—Interim report. (NPRDC Technical Report 78-24). San Diego, CA: Navy Personnel Research and Development Center, June 1978.

Author's Address

Send requests for reprints or further information to Ronna F. Dillon, Department of Guidance and Educational Psychology, Southern Illinois University, Carbondale, IL 62901.