Analogue Reasoning Under Different Methods of Test Administration

Ronna F. Dillon
Southern Illinois University

One hundred eighty-five college undergraduates were given the Advanced Progressive Matrices under one of five conditions of testing: standard, simple feedback, examinee verbalization during problem solution, elaborated feedback, and full elaboration. The Group Embedded Figures Test, Paragraph Completion Test, and Zelniker and Jeffrey's revision of the Matching Familiar Figures Test were also administered. The study was designed (1) to investigate the differential effects of method of test administration on performance for college students and (2) to examine the relationship of individual differences dimensions and varying conditions of testing. Analysis of variance coupled with orthogonal comparisons revealed higher levels of performance under the more elaborative testing conditions. The cognitive style variables were differentially related to performance in the different testing conditions. The processing dimensions were related to performance to a higher degree under partially elaborative conditions than under either nonelaborative procedures or full elaboration. Results are discussed in terms of an activation model.

A number of investigations over the past several years (e.g., Carlson & Dillon, 1978, 1979; Dillon, 1980; Dillon & Carlson, 1978) have been guided by the desire to "test for competence," i.e., to cut through the performance variables and the nontarget demands that might be responsible for a discrepancy between an individ-

ual's actual cognitive competence and the behavioral manifestation of the capabilities in question. As Smedslund (1969) has noted, the impeding factors fall into two broad categories:

1. Factors related to the testing situation, e.g., the use of tasks with excessive nontarget demands, such as those imposed by language or memory, or the use of tasks having demands which elicit inappropriate strategies from the examinee; and
2. Factors related to the individual being tested, e.g., motivational, personality, or socio-cultural factors.

The goal of the testing for competence work is to render currently available testing instruments more sensitive indicators of the examinee's possession of cognitive abilities. Although other approaches have been reported, the assessment process in these investigations has either involved elaborate familiarization training (a requirement rendering the procedures burdensome for routine assessment; Budoff & Corman, 1975) or procedures that have changed task variables such that meaningful comparisons with the target task were difficult (Miller, 1976). The present paradigm involves the use of a series of elaborative testing conditions designed to activate underlying competence, i.e., to facilitate the elaboration of existing cognitive organization with no pretraining. Previous research has indicated that elaborative procedures result in en-

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 5, No. 3, Summer 1981, pp. 341-347
© Copyright 1981 Applied Psychological Measurement Inc.
0146-6216/81/030341-07\$1.35

hanced performance and criterion-related validity for normal hearing children (Carlson & Dillon, 1979; Carlson & Wiedl, 1976; Dillon & Carlson, 1978), educable mentally retarded persons (Carlson & Wiedl, 1978), and hearing-impaired children (Carlson & Dillon, 1979; Dillon, 1979) on Piagetian and Raven matrices tasks.

Accurate assessment of intellectual capability requires attention not only to the conditions of test administration but also to the relationship of these conditions to dimensions of individual differences. Carlson and Wiedl (1976) found that the detrimental effects that dimensions such as cognitive impulsivity and external locus of control have on the performance of normal hearing children in Grades 2 and 4 were reduced or eliminated under certain elaborative testing conditions but not under others. Dillon (1980) reported a similar finding for hearing-impaired children 8 to 12 years old. For example, results of both investigations indicated that the negative effects of cognitive impulsivity on performance were reduced under testing procedures involving examinee verbalization or elaborated feedback. The present investigation was designed to examine the extent to which this result could also be observed for college students.

No reported attempts have been made to consider the differential effects of method of test administration for individuals differing along the dimension of field dependence-independence (Kagan, Rosman, Day, Albert, & Phillips, 1964). Such a consideration is clearly warranted in the present context, as one salient respect in which the testing conditions differ concerns whether the structure is imposed from outside the examinee, in the form of examiner feedback, or from within the individual, in the form of examinee verbalization.

Consideration of the relationship of individual differences in integrative complexity to performance under different methods of test administration stems from evidence in the integrative complexity literature (Schroder, Driver, & Streufert, 1967) that individuals manifesting a relatively high degree of integrative complexity

respond favorably to learning environments that are complex and in which the possibility for integrating feedback from a number of sources (e.g., examiner- and examinee-generated structure) exists. The fully elaborative testing condition employed in the present investigation should constitute an optimal testing environment for such individuals.

The figural analogies of the Advanced Progressive Matrices (Raven, 1962) require logical problem-solving not necessary for solution of items on the processing style tasks. The latter items tap the efficient visualization of information: information that need not be conceptually transformed in any way. Thus, the *manner* in which visual information is processed is tapped by the style indices, as opposed to general problem-solving ability that is assessed through performance on the Matrices test.

The present investigation constitutes an extension of previous work. No reported attempts have been made to establish the differential effects of method of test administration, as well as the interaction between selected individual differences dimensions and varying conditions of testing on performance for college students.

Method

Sample

The sample consisted of a total of 185 upper-division undergraduate college students. Males and females were approximately equally represented.

Tests

The Advanced Progressive Matrices (APM; Raven, 1962) was administered to all examinees. The APM consisted of 36 items. Each item included a 3×3 stimulus array where horizontal and vertical dimensions varied with respect to specific attributes. The examinee was instructed to find from a set of eight distractors the piece that correctly completed the

stimulus array. All items were scored 1 if correct, 0 if incorrect.

The Group Embedded Figures Test (GEFT; Oltman, Raskin, & Witkin, 1971) was used to assess individual differences in field dependence-independence. The timed test contained a practice set and 30 dichotomously scored items. Test items were in paper-and-pencil format and required the examinee to demonstrate the ability to disembed a figure from an embedding context by tracing the outline of the simple figure. Following the 2-minute practice set, examinees were allotted 5 minutes for each of two sets of items.

The Matching Familiar Figures Test (MFFT; Kagan, Rosman, Day, Albert, & Phillips, 1964), as revised by Zelniker and Jeffrey (1976), was used to assess individual differences in reflection-impulsivity. For each item, the examinee was shown a picture of an object and a set of six similar variants, only one of which was identical to the standard. The task was to select the identical picture. For 14 of the 28 items, variants differed from the standard in some aspect of detail. The remaining items' variants differed from their respective standard with respect to contour or outside border. Latency to the first response and number of errors overall, taken separately, were the indices of interest. Due to methodological problems with the double median-split procedure (Ault, Mitchell, & Hartmann, 1975), the two measures were considered as independent continuous variables. Global and detail items were analyzed separately.

The Paragraph Completion Test (PCT; Schroder, Driver, & Streufert, 1967) was given to examinees to assess individual differences in integrative complexity. The examinee was given a series of terms and asked to generate a three-sentence paragraph for each term. Individual paragraphs were scored with respect to level of conceptual structure, i.e., the degree of differentiation and number of degrees of freedom in the rules of integration in un-

derlying mediational processes. Responses were scored along a 7-point scale.

Procedures

The GEFT and PCT were group administered, while the MFFT and APM were individually administered. Order of administration of the different tests was varied randomly. The measures of cognitive style were administered under standard procedures only, whereas the APM was administered in five conditions. Examinees were randomly assigned to one of the following conditions: (1) standard, (2) simple feedback as to correctness or incorrectness of response, (3) verbalization of solution strategies during solution of each item, (4) elaborated feedback by the examiner as to reasons for the correctness or incorrectness of each choice, and (5) examinee verbalization coupled with elaborated feedback. In all conditions, the examinee's response to each item was scored prior to elaboration of any sort. Testing conditions were equated with respect to time by adding post-response latency (deemed appropriate during pilot testing) to items presented under Conditions 1 and 2. No minimal time limits were required. Inter-rater reliability for the PCT was .90.

The logical operations being tapped were extremely complex. Unless the structures are present in competence, elaborated feedback is not likely to provide appropriate cues as to necessary solution processes. During such feedback, the examinee was told why a particular response alternative is incorrect, although the correct choice was not indicated. Further, the specific task demands (i.e., the rule or rules that must be generated for specific item solution) were sufficiently different so that no direct application from one item to the next was possible. Under the verbalization condition, examinees were directed to "think aloud." No externally generated mediation was possible. Budoff and Corman (1975) concurred in their discussion of the non-reactivity of such procedures.

Results

Table 1 displays means and standard deviations of APM scores for examinees assigned to the five testing conditions. The data were subjected to a one-way analysis of variance with five levels of the treatment variable. Analysis indicated that the groups' means differed, $F(4,190) = 25.05, p < .001$.

Orthogonal comparisons indicated that the groups' means did not differ for Condition 1 (standard), as compared to Condition 2 (simple feedback). Performance under examinee verbalization during item solution (Condition 3) exceeded performance under simple feedback, $F(1,180) = 10.33, p < .05$, while performance under elaborated feedback (Condition 4) was higher than performance under Condition 3 (examinee verbalization), $F(1,180) = 4.25, p < .05$. Performance under Condition 5, which combined examinee verbalization with elaborated feedback, exceeded performance under elaborated feedback alone (Condition 4), $F(1,180) = 4.16, p < .05$.

Table 2 displays zero order correlation coefficients between the style predictors and APM performance under the different testing conditions, and corresponding tests for the significance of each correlation. Under Condition 1, the only processing style that was significantly related to APM performance was measured by GEFT score, $r = .43, p < .05$. For Condition 2,

significant relationships were identified for GEFT, $r = .56, p < .001$; for analytic MFFT error, $r = -.35, p < .05$, and for global MFFT error, $r = -.37, p < .05$. Under Condition 3, significant relationships were again identified for analytic MFFT error, $r = -.42, p < .05$; and for global MFFT error, $r = -.56, p < .05$. The relationship of GEFT and APM approached significance, $r = .28, p < .10$. For elaborated feedback (Condition 4) significant relationships were identified for GEFT, $r = .51$; analytic MFFT error, $r = -.38$; and global MFFT error, $r = -.44, p < .05$ in all cases. The relationship of integrative complexity to APM approached significance, $r = .35, p = .05$. Under Condition 5, the fully elaborative mode, a significant relationship was identified between APM and integrative complexity, $r = .35, p < .05$, with no other variables bearing a relationship to APM score.

Discussion

Performance on the Advanced Progressive Matrices varied as a function of method of test administration. As predicted, higher performance on the APM resulted under the more elaborative testing conditions. Thus, the cognitive competence possessed by college students may not be veridically assessed under standard testing procedures, as was found for normal children (Carlson & Dillon, 1979; Dillon & Carl-

Table 1
Means and Standard Deviations for APM Performance
Under Five Testing Conditions

Condition	n	Mean	Standard Deviation
1	38	17.05	5.66
2	37	18.35	5.35
3	37	21.97	4.17
4	37	24.30	4.63
5	36	26.61	4.19

Table 2
Correlation Coefficient for Style Variables and APM Performance
Under Five Testing Conditions

Variable	Testing Condition				
	1	2	3	4	5
Analytic Match-to-Sample Latency	.10	.12	-.12	.04	-.07
Analytic Match-to-Sample Error	-.11	-.35*	-.42*	-.37*	-.06
Global Match-to-Sample Latency	.10	.07	.01	.12	.00
Global Match-to-Sample Error	-.17	-.37*	-.56**	-.44*	.00
Locus of Control	-.12	.14	.15	-.06	.08
GEFT	.44*	.56**	.28	.51*	.05
Integrative Complexity	-.02	.09	.27	.35*	.35*

* $p < .05$

** $p < .01$

son, 1978), for educable mentally retarded persons (Carlson & Wiedl, 1978), and for hearing-impaired children (Dillon, 1979).

The data supported the second hypothesis, that the cognitive style variables were differentially related to performance under the five testing conditions. Field dependence-independence, as measured by the Group Embedded Figures Test score, was related to performance under standard procedures, simple feedback, and elaborated feedback. Thus, relatively field dependent individuals may be penalized by treatment effects under these conditions. Cognitive impulsivity, as measured by the modified Matching Familiar Figures Test, was related to performance under simple feedback, examinee verbalization, and elaborated feedback. Individuals manifesting a high degree of cognitive impulsivity may be most appropriately tested under full elaboration, as in Condition 5. Integrative complexity, assessed by performance on the Paragraph Completion Test, was related to Condition 5 performance. Individuals on the low end

of this processing continuum may be optimally assessed under Conditions 3 or 4.

For cognitive assessments to reveal cognitive competence as closely as possible, a differential approach to assessment is needed. The data support an activation model, in which treatment effects vary with respect to the extent to which a given treatment serves to orient the examinee toward constituent processes. Under the non-elaborative procedures of Condition 1, performance was low, with the impact of individual difference dimensions being relatively small and insufficiently salient to militate against the effects of nontarget demands. Under partially elaborative conditions, the effects of individual differences in cognitive style were maximized. The match of individual difference characteristics to testing procedure is important because only partial activation is accomplished, i.e., activation is more easily facilitated when treatment components or task demands are consonant with the examinee's preferred processing style. Thus, the distinction between match-to-sample

items requiring global processes and items of the same type but having different task demands (i.e., requiring analysis of detail) is of greatest importance under partially elaborative conditions.

Under the fully elaborative Condition 5, performance was maximized. The deleterious effects of nontarget demands were most likely to be overcome. As under standard procedures, the importance of individual differences along indices of cognitive style again diminished. Treatment effects in Condition 5 involved both examinee-generated and examiner-generated elaboration, such that the examinee gained exposure to reasoning via less preferred strategies (i.e., strategies present in the examinee's repertoire but less likely to be activated under nonelaborative conditions) when task demands necessitated the use of such strategies. The regulatory function of language is likely to be optimal.

In the only other reported attempt to examine the relationship of indices of cognitive style to method of test administration (Dillon, 1980), findings consistent with the present investigation were noted. Cognitive style variables failed to predict the performance of hearing-impaired children, 8 to 12 years old, under both nonelaborative and fully elaborative conditions. Performance under three partially elaborative conditions was predicted by cognitive style on the Raven CPM and a Piagetian battery of concrete operations tasks.

Thus, under full elaboration, treatment effects overshadowed the influence of individual differences in information-processing styles by fully activating underlying competence. Researchers need to continually consider the importance of (1) harnessing individual differences in determining the appropriateness of treatments for particular individuals and of (2) the possibility of developing psychodiagnostic treatments of sufficient potency as to lessen the negative effects of particular processing characteristics for individuals on the less preferred ends of the style continua. Potent treatments of this sort would not only result in optimal levels of performance but

would also be of considerable psychodiagnostic utility. Veridical assessment of cognitive abilities is necessary if we are to gain information relevant to theories of underlying process, to study fruitfully the course of cognitive growth, or to determine the efficacy of instructional interventions or the developmental appropriateness of treatments and curricula.

References

- Ault, R., Mitchell, C., & Hartmann, D. P. *Some methodological problems in reflection-impulsivity research*. Paper presented at the meeting of the Society for Research in Child Development, Denver, April 1975.
- Budoff, M., & Corman, L. A manual for the Raven learning potential test. *Studies in Learning Potential*, 1975, 59 (Monograph).
- Carlson, J. S., & Dillon, R. The assessment of intellectual capabilities in hearing impaired children: Effects of testing-the-limits procedures. *Volta Review*, 1978, 80, 216-224.
- Carlson, J. S., & Dillon, R. The effects of testing conditions on Piaget Matrices and order of appearance problems: A study of competence vs. performance. *Journal of Educational Measurement*, 1979, 16, 19-26.
- Carlson, J. S., & Wiedl, K. H. Modes of presentation of the Raven Coloured Progressive Matrices Test: Towards a differential testing approach. *Trierer Psychologische Berichte*, 1976, 3 (Monograph).
- Carlson, J. S., & Wiedl, K. H. Use of testing-the-limits procedures in the assessment of intellectual capabilities in children with learning difficulties. *American Journal of Mental Deficiency*, 1978, 82, 559-564.
- Dillon, R. Improving validity by testing for competence: Refinement of a paradigm and its application to the hearing impaired. *Educational and Psychological Measurement*, 1979, 39, 363-371.
- Dillon, R. Cognitive style and elaboration of logical abilities in hearing-impaired children. *Journal of Experimental Child Psychology*, 1980, 30, 389-400.
- Dillon, R., & Carlson, J. Testing for competence in three ethnic groups. *Educational and Psychological Measurement*, 1978, 38, 437-443.
- Kagan, J., Rosman, B. L., Day, D., Albert, J., & Phillips, W. Information processing in the child: Significance of analytic and reflective attitudes. *Psychological Monographs*, 1964, 78(1, Whole No. 578).

- Miller, S. A. Nonverbal assessment of Piagetian concepts. *Psychological Bulletin*, 1976, 83, 405-430.
- Oltman, P. K., Raskin, E., & Witkin, H. A. *Group Embedded Figures Test*. Palo Alto: Consulting Psychologists Press, 1971.
- Raven, J. C. *Advanced Progressive Matrices*, Set II. London: H. K. Lewis, 1962.
- Schroder, H. M., Driver, M. J., & Streufert, S. *Human information processing*. New York: Holt, Rinehart, & Winston, 1967.
- Smedslund, J. Psychological diagnostics. *Psychological Bulletin*, 1969, 71, 237-248.
- Zelniker, T., & Jeffrey, W. E. Reflective and impulsive children: Strategies of information processing underlying differences in problem solving. *Monographs of the Society for Research in Child Development*, 1976, 41(5, Serial No. 168).

Acknowledgments

This research was supported by Research Grant 2-10839 from the Office of Research Development and Administration, Southern Illinois University at Carbondale. The author wishes to thank Patricia Hartman for her assistance in data collection and analysis.

Author's Address

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