The importance of dynamic testing procedures for understanding reasoning and its underlying cognitive processes as well as for increasing validity is investigated. An experiment tested the premise that African American and White college students demonstrate equivalent inductive reasoning ability when tested under induced mediation procedures, where they are encouraged to structure their own thought processes. Subject verbalization procedures foster induced mediation, while experimenter verbalization foster imposed mediation. Participants (77 White and 45 African American undergraduates) performed better under the two dynamic conditions than under standard procedures. In addition, an interaction was found between ethnic group membership and testing condition. White participants performed equally well under either dynamic condition, while African American participants performed better under subject-directed, induced mediation than under experimenter-directed imposed mediation. As hypothesized, no differences are found between African American and White college students under induced mediation procedures. Also, validity data are consistent with earlier findings that point to the importance of considering the role of learner attributes in testing for elucidating intelligence or ability. Previous reports of White college undergraduates outperforming their African American peers on tests of general intellectual ability are due to a greater incompatibility between demands of the testing situation and preferred processing tactics for African American participants than for White participants. (Contains 1 table and 38 references.) (Author/SLD)
Dynamic Assessment of Inductive Reasoning: A Test of Induced Versus Imposed Mediation

Ronna F. Dillon
Southern Illinois University

Correspondence concerning this article should be addressed to Ronna F. Dillon, Department of Educational Psychology and Special Education, Southern Illinois University, Carbondale, Illinois 62901.
Dynamic Assessment of Inductive Reasoning

Abstract

The importance of dynamic testing procedures for understanding reasoning and its underlying cognitive processes as well as for increasing validity is investigated. An experiment was conducted to test the premise that African American and white college students demonstrate equivalent inductive reasoning ability when tested under induced mediation procedures, where they are encouraged to structure their own thought processes. Subject verbalization procedures foster induced mediation, while experimenter verbalization foster imposed mediation. Participants performed better under the two dynamic conditions than under standard procedures. In addition, an interaction was found between ethnic group membership and testing condition. White participants performed equally well under either dynamic condition, while African American participants performed better under subject-directed, induced, mediation than under experimenter-directed, imposed mediation. As hypothesized, no differences are found between African American and white college students under induced mediation procedures. Also, validity data are consistent with earlier findings that point to the importance of considering the role of learner attributes in testing for elucidating intelligence or ability. Previous reports of white college undergraduates outperforming their African American peers on tests of general intellectual ability are due to a greater incompatibility between demands of the testing situation and preferred processing tactics for African American participants than for white participants.
Psychologists have recognized the difficulty of gaining valid assessments of cognitive abilities across racial or ethnic groups for at least two decades (Dillon, 1997). Research findings consistently have indicated that incompatibilities between some aspect of the learners' cultural context and demands of the learning situation contribute to ethnic group differences in conceptual learning (e.g., Fisher, Liu, Velozo, & Pan, 1992; Vernon, 1975). In addressing the problem of gaining valid assessments of individuals' current abilities, researchers have claimed that attention should be paid to the relevance of the situation in which the abilities and processes in question are tested and used (e.g., Dillon, 1997).

The early work of Cole, Gay, Glick, and Sharp (1971) exemplifies research findings that points to the sources of ethnic group differences in reasoning. Data from this work indicates that ethnic group differences in complex reasoning and Piagetian task performance result from cultural/experiential factors. More recent studies of abilities across ethnic groups provide consistent results (Bond, 1990; Bresia & Fortune, 1989; Cole & Moss, 1989; Lonner, 1990; Schmitt, 1988).

A related set of investigations provided data regarding the relationship between ethnicity and clinical assessment (Campbell, & Williams, 1989; Johnson & Brems, 1990; Lambert, 1981; Neighbors, Jackson, 1989; Reid, 1995; Sonuga-Barke, Minocha, Taylor, & Sandberg, 1995; Vasquez, 1988; Walters, 1986; Westermeyer, 1987). Moreover,
several experiments focused on assessment in cross-cultural counseling (del Carmen, 1990; Lonner & Ibrahim, 1989; Prediger, 1993).

Specific to the present experiment, Butler-Omololu, Doster, & Lahey (1984) found that African American, male high school students scored higher on identification tasks than did white males, although the pattern of results was not found on other IQ test-type tasks. While interesting, the Butler-Omololu et al. (1984) work does not provide a direct test of the effects of different testing methods for assessing ethnic group differences on the same test. Without such a test, it is difficult to make meaningful statements about reasoning performance across ethnic groups. In synthesizing this literature, it appears that important sources of learner differences across ethnic groups lie in differences in the extent of compatibility between task demands and learner attributes. The experiment reported in this paper provides a direct test of this proposal.

Testing and Intelligence Theory

Under the definition of intellectual abilities that guides this work, intelligence is viewed as elaborated thinking. The interaction between an individual's capabilities and environmental influences is stressed (Clinkenbeard, 1997; Dillon, 1997; Snow, 1982; Sternberg, 1997; Sternberg, Wagner, Williams, & Horvath, 1995; Williams, Blythe, White, Sternberg, & Gardner, 1996). At the level of information-processing abilities, intelligence is characterized as the execution of a small set of elementary information processes.

Individual differences in intelligence can be seen at five levels. The stage level concerns the efficient and effective execution of distinct information-processing components, or stages; i.e., encoding stimulus attributes, forming rules regarding the inferences governing the item, applying previously inferred rules, and selecting the correct response. The
sequence level pertains to the sequential distribution of processing stages. The strategy level taps the execution of optional strategies or strategy components such as response elimination or double-checking. The learning level concerns gross adaptations over time. Finally, the flexibility level taps the examinee’s ability to maintain information-processing efficiency as the demands of a task change (Dillon, 1981, 1986, 1997). The environment is viewed to play an important role in facilitating, or hindering, the manifestation of intellectual competence in performance.

Dynamic Assessment and Intelligence

Each dynamic assessment approach derives from a particular model or view of intelligence. Two main classes of dynamic assessment models and derived testing approaches have received the greatest amount of attention (see Dillon, 1997 for a review of dynamic assessment research). Test-train-test approaches derive from a view that intelligence is the ability to profit from experience. Consequently, these dynamic assessment paradigms involve some type of training prior to the assessment of aptitude or intelligence. Testing-the-limits approaches derive instead from the view that intelligence is the genotype-environment interaction. The testing-the-limits form of dynamic assessment, used in this experiment, derives from the information-processing model of intelligence. In the model, intelligence is equated with the individual’s current level of functioning. Therefore, the goal of dynamic assessment is the activation of current underlying information-processing competence, not training or altering cognitive abilities (Jitendra & Kameenui, 1993).

The effects of verbalization during item solution on components of intelligence are well documented. As far back as Luria (1966, 1973) and Vygotsky (1978, 1986),
researchers identified the role of speech processes in the development of thought and problem solving. Verbalization is known to direct the participant’s attention selectively to the properties of information (i.e., encoding) and to help participants recognize (i.e., rule inference) and identify (i.e., rule application) the existence of different facets of a problem. These processes collectively have been referred to as mediation (see Dillon, 1997).

Recording eye movement patterns during solution of inductive reasoning problems that were administered under examiner and examinee verbalization conditions, Dillon (1985) also found that these testing-the-limits procedures affect the sequential distribution of information processes. For example, under dynamic as opposed to standard procedures, examinees execute a greater percentage of their processes in the main stimulus array prior to their first attempt at response selection. Such processing is related to performance effectiveness. With respect to the strategy level of individual differences, Dillon also reported that participants tested under dynamic procedures executed more intraitem double-checking and less interitem double-checking, another index of processing effectiveness, than under standard procedures. Dynamic procedures involved either induced mediation, wherein the examinee provides the verbal mediation, or imposed mediation, wherein the examiner provides the verbal mediation. Testing-the-limits procedures are believed to yield test performance that is closer to true abilities than occurs under standard procedures. The testing-the-limits procedures work by facilitating use of multiple problem representations, fostering dual-coding of stimulus materials, and serving an organizational function for thinking (Dillon, 1989, 1992).

The proposal being made here is that mediation during dynamic assessment enhances performance only to the level of current capacity. Thus, to the extent that
examinees spontaneously employ the cognitive mediation embodied in the dynamic procedures, intellectual capacity is reached during standard procedures, and dynamic testing does not yield further improvement in performance. However, when examinees do not employ cognitive procedures such as dual-coding of stimulus materials and use of multiple mental representations during thinking, dynamic procedures can foster these mental activities. Moreover, individuals may experience greater benefit under either induced or imposed mediation procedures.

As a case in point, Dillon (1980) tested college students under both verbalization and elaborated feedback conditions. Eighty-five percent of the participants demonstrated a clear performance differential under either the imposed mediation condition or the induced mediation condition. When subjects were either matched or mismatched to their preferred condition for additional testing, matched subjects outperformed their unmatched counterparts.

Two findings motivate the work reported in this paper. First, ethnic group differences in complex reasoning -- which have been reported under standard testing procedures -- are not found for children when the same tasks are administered under examinee verbalization or examinee-verbalization conditions (Dillon & Carlson, 1978). Second, reasoning performance of college students is greater when examinees are matched to their preferred dynamic conditions (Dillon, 1980).

The logical next step in this research is to test directly several premises. First, fostering mediation among college students during inductive reasoning is expected to enhance inductive reasoning performance (see Dillon, 1997). Second, differences in reasoning among African American and white college students that have been reported
under traditional psychometric methods are not expected to be found under subject-directed, i.e., induced, mediation wherein subjects can structure their own thought processes. An interaction between ethnic group and testing condition is hypothesized, wherein African American examinees are expected to perform better under induced mediation than under imposed mediation. Performance of white examinees is expected to be the same under either dynamic condition. Finally, subject verbalization is expected to provide the most valid measurements for African American college-age adults, while no differences are expected in measures of ability for white examinees.

The experiment reported in this paper is designed to test the premise that African American and white college students demonstrate equivalent reasoning ability when they are encouraged to structure their own thinking, through induced mediation. This work is motivated by the assumption that differences in complex reasoning performance between African American and white college students are due in significant measure to greater incompatibility between the type of mediation imposed by test techniques, examples, and procedures for African American students than for their white peers.

These dynamic testing conditions are believed to provide examinees with a testing environment that is more similar to the manner in which thinking and learning typically occur. Therefore, predictors derived in this manner should be more valid indices of important criterion measures.
Method

Participants

The sample was comprised of 122 upper-division college undergraduates, 77 white participants and 45 African American participants, from 19-32 years old. All participants were volunteers who earned regular admission, under the same admission requirements. Participants were enrolled in an upper-division educational psychology course and were in normal academic standing. Males and females were approximately equally represented.

Instrument

The Advanced Progressive Matrices (APM; Raven, 1962) was given to all participants. For solution of each APM item, participants were instructed to select the response option that correctly completed the test item. Each item is a 3 x 3 figural analogy with a missing portion in the lower right corner of the configuration. An 8-item response set is located beneath the main stimulus array. Three criterion measures were selected for examination: (1) High school rank; (2) cumulative grade-point average (GPA); and (3) ACT Composite score.

Procedure

All testing was done on an individual basis. Examinees were randomly assigned to one of the following testing conditions. Examinees were given the same test items under all three testing conditions. Correct answers were not disclosed under any testing condition.

Standard condition. Under standard procedures, the examinee was given a test item, asked to study the item for 2 minutes, and asked to provide the experimenter with the letter corresponding to the chosen response option. After response selection, the
examinee was instructed to review the item and the response option that had been selected for 30 seconds. No interaction with the subject was elicited. Total time per item was 2 minutes, 30 seconds.

**Examinee verbalization.** The examinee was instructed to “think aloud” during item solution, for 2 minutes per item, describing solution tactics. Examinees described the relevant task features or dimensions (i.e., encoding), they discussed how they formed the rules that governed the test item (i.e., inference), they explained how these rules were applied to other parts of the item (i.e., rule application), and they described their methods of response selection (i.e., confirmation). The thinking-aloud period lasted 2 minutes per item, and each participant then reported his or her response option and described the reasons for selecting the particular option for 30 seconds. Total time per item was 2 minutes, 30 seconds.

**Experimenter verbalization.** Examinees studied each test item and selected a response option. The study and response selection activities lasted 2 minutes per item. Following the subject’s reported solution to each item, the experimenter then directed questions to the subject concerning salient item features. Experimenter-directed questions concerned encoding, rule inference, rule application, and confirmation. The verbalizations were given verbatim, in their entirety, to all subjects in this condition. Experimenter verbalization lasted 30 seconds per item. Total time per item was 2 minutes, 30 seconds.
Dynamic Assessment of Inductive Reasoning

Results

Table 1 contains means and standard deviations for APM score for participants assigned to one of the three testing conditions. An alpha level of .05 was used for all statistical tests.

Insert Table 1 about here

Inductive Reasoning Under Dynamic Assessment

Analysis indicates that the groups means differed, $F(2, 121) = 23.35$, $p < .001$, with performance under the dynamic testing conditions being greater than performance under the standard condition, $F(1, 121) = 45.95$, $p < .001$. Also as hypothesized, an interaction is found between ethnic group membership and optimal testing condition, $F(2, 121) = 3.67$, $p < .05$. No difference is found for white participants under induced and imposed mediation conditions $F(1, 49) = .06$, $p > .05$), while black participants’ performance is greater under the induced condition than under the imposed procedure $F(1, 27) = 2.45$, $p < .05$. Finally, as hypothesized, no difference in performance is found between the two ethnic groups under the induced mediation condition $F(1, 35) = 1.48$, $p > .05$.

Validity of Dynamic Testing Session

Concerning the validity of this dynamic assessment of inductive reasoning, the criterion-related validity of the two elaborative conditions outstripped the standard condition for both ethnic groups, for all three criterion measures. For white participants using ACT$^1$ as the criterion, performance under induced mediation, $F(1, 10) = 9.34$, $p <$
Dynamic Assessment of Inductive Reasoning

.05 (R² = .51), and performance under imposed mediation, F(1, 11) = 20.04, p < .05 (R² = .67), are excellent predictors of achievement, while as hypothesized, performance under the standard condition appears to be a less powerful predictor of academic achievement, F(1, 14) = 5.44, p < .05 (R² = .30). For African American participants, performance under the standard condition does not predict ACT composite, as expected, F(1, 8) = .23, p > .05, while performance under induced mediation is an excellent predictor of ACT performance, F(1, 11) = 8.46, p < .05, accounting for 46% of the variance in ACT Composite. Consistent with the hypothesized pattern of relationships, performance under imposed mediation does not account for significant variation in ACT performance for African American participants F(1, 9) = .34, p > .05.

With high school rank as the criterion, reasoning performance for white subjects under the standard condition does not account for significant variation in criterion task performance F (1, 10) = 2.24, p > .05, while, again consistent with the hypothesized pattern of relationships, performance under the two dynamic conditions is a good predictor of high school achievement; under verbalization, i.e., induced mediation, F(1, 11) = 5.99, p < .05 (R² = .37), and under elaborated feedback; i.e., imposed mediation, F(1, 11) = 9.42, p < .05 (R² = .49). For African American participants on this same criterion, reasoning performance under the standard condition again, as hypothesized, does not account for significant variation in high school performance F(1, 6) = .45, p > .05, while reasoning performance under the induced mediation condition is a powerful predictor of criterion task performance, F(1, 10) = 11.21, p < .01 (R² = .55). Again as hypothesized, performance under the imposed mediation condition does not account for significant variation in high school rank, F(1, 13) = 3.64, p > .05.
Reasoning performance under the standard condition does not account for significant variation in GPA for white subjects, as hypothesized, \( F(1, 26) = .29, p > .05 \). Also as expected, performance under induced mediation, \( F(1, 23) = 40.08, p < .001, R^2 = .65 \), and under imposed mediation, \( F(1, 24) = 48.33, p < .001, R^2 = .68 \), are powerful predictors in both instances. For African American participants, again as hypothesized, reasoning performance under induced mediation is a powerful predictor of GPA, \( F(1, 11) = 13.61, p < .01, R^2 = .60 \).

Discussion

Data from the experiment reported in this paper indicate that college students’ inductive reasoning performance is greater under dynamic testing conditions than under standard procedures. Moreover, while white college students perform equally well under either induced or imposed mediation, African American college students perform better under subject-directed, induced, mediation. The latter finding is consistent with the literature on assessment across ethnic groups and supports the proposal being made here that a greater incompatibility exists between content and/or processes verbalized by the examiner under imposed mediation for African American college students than for white students. Moreover, the two groups perform equally well the subject-directed; i.e., induced mediation procedures employed in the subject verbalization condition. Thus, African American college students perform as well as white students on complex reasoning items when they are encouraged to structure their own thought processes, as opposed to having a structure imposed by the experimenter.

The importance of dynamic testing procedures is underscored also by noting that the assessments performed using these procedures provide more valid measures than
standard administrations of the same test. These more valid administrations of the test result from use of a testing approach that is more similar to typical thinking and learning environments than is the case under standard testing procedures, that activates underlying intellectual capacity, and that is compatible with differences in examinee attributes. Differences identified in previous reports between groups of subjects may be due to a greater mismatch between preferred processing tactics and situational demands for one group of subjects than for another group of individuals.

Clearly, equivalent inductive reasoning ability is identified across ethnic groups under subject-directed procedures. In short, the data reported in this experiment support the proposal that conventional testing, at least at times, fails to reveal the true abilities of African American examinees. Previous reports of ethnic group differences might be considered in light of this finding, whether standard procedures or other testing conditions are used. Further research is underway to extend this work to other intellectual tasks.
Dynamic Assessment of Inductive Reasoning

References


Schmitt, A. P. (1988). Language and cultural characteristics that explain
differential item functioning for Hispanic examinees on the Scholastic Aptitude Test.

Montague (Eds.), Aptitude, learning, and instruction: Cognitive process analysis of

bias in teachers' ratings of childhood hyperactivity. British Journal of Developmental
Psychology, 11, 187-200.


children: A construct validation of the triarchic model. Gifted Child Quarterly, 40(4),
220-228.


Methodological issues, challenges, and suggestions. In K. S. Pope (Chair). Doing valid
and useful assessments: New demands on practitioners. Symposium conducted at the
96th Annual Convention of the American Psychological Association. Atlanta, GA.

(Eds.), Race and education across cultures. London: Heinemann.


Footnote

'American College Admissions Test (ACT), and high school rank (HSR) data were used as available.
### Table 1

**Means and Standard Deviations for Inductive Reasoning Under Three Testing Conditions**

<table>
<thead>
<tr>
<th>Testing Condition</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n = 27 (whites)</td>
<td>18.15</td>
<td>5.26</td>
</tr>
<tr>
<td>n = 16 (blacks)</td>
<td>16.19</td>
<td>5.74</td>
</tr>
<tr>
<td><strong>Subject Verbalization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n = 24 (whites)</td>
<td>24.21</td>
<td>4.37</td>
</tr>
<tr>
<td>n = 12 (blacks)</td>
<td>26.58</td>
<td>6.60</td>
</tr>
<tr>
<td><strong>Experimenter Verbalization</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n = 26 (whites)</td>
<td>25.19</td>
<td>4.32</td>
</tr>
<tr>
<td>n = 17 (blacks)</td>
<td>21.00</td>
<td>5.65</td>
</tr>
</tbody>
</table>
Title: Dynamic Assessment of Inductive Reasoning: A Test of Induced Versus Imposed Mediation

Author(s): Dillon, Ronna F.

Corporate Source: Southern Illinois University

Publication Date: October, 1998

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Ronna Dillon


Southern Illinois University

Date: 10/16/98