

## **A Diagnostic Profile of Art Understandings Based on Verbal Responses to Works of Art**

*Carol Stavropoulos*

This research was undertaken to lessen the gap in art education research on student assessment literature. The assessment instruments that were available to assess written statements about works of art did not account for cognitive conceptions of learning. Furthermore, they did not discriminate between knowledge base and knowledge-seeking strategies, higher-order and lower-order understandings, and misunderstandings.

This paper will summarize the steps taken to develop a diagnostic assessment profile as follows: (a) the theoretical foundation directing the inquiry, (b) methodologies used in development of categories, (c) validity and reliability studies employed to establish instrument quality, and (d) its significance in assessing understandings of art.

The research by Koroscik, Short, Stavropoulos, and Fortin (1992), referred to as the Chagall study, motivated the development of the assessment instrument. Results of the Chagall study indicate student understanding can be facilitated by presenting artworks within comparative contexts. Understanding is further facilitated when verbal cues are provided. Through an analysis of qualitative data, the research team found evidence to support the independent functioning of a student's knowledge base and knowledge-seeking strategies. These findings indicate that a diagnostic assessment of art learning must encompass both facets of learning.

Further analysis of written statements from the Chagall Study revealed a continuum of learning outcomes in the formal, descriptive, interpretive, and historical dimensions of understanding. The written statements were reduced and organized into categories by adopting Guba's (1978) naturalistic inquiry methodology. Through a series of steps, learning outcomes were converted into diagnostic categories as follows: (a) to represent the students application of the knowledge base and knowledge-seeking strategies through the process of transfer, (b) to encompass terminology reflective of the field of art education, and (c) to make discriminations between lower-order understandings, higher-order understandings, and misunderstandings.

### **Summary of Validity Studies**

Validity of the diagnostic profile was estimated by two content validity studies, two criterion-related studies, and a construct validity study. Results of each validity study are summarized below.

## **Content Validity**

Experts were recruited to serve as independent judges for the two content validity studies. Dr. Andra Johnson, an evaluation specialist in art education and Dr. Jacqueline Chanda, an art historian in art education assessed the content of the diagnostic profile. A summary of the results of their analysis follows.

**Importance and appropriateness of content.** Judgments were rendered on the importance and appropriateness of categories in assessing art understandings. With a few minor exceptions, the diagnostic categories were considered important and appropriate.

**Practicality.** The complexity of the analysis performed by the diagnostic profile justified its length, according to Dr. Johnson. In addition, Dr. Johnson responded positively to the procedural and administrative aspects of the diagnostic profile. She found the training procedures to be reasonable, and the methods of coding and reporting data were judged to be clear and concise.

**Applicability to artforms.** Dr. Chanda was asked to rate the diagnostic categories in terms of their applicability to a variety of artforms. Since the categories were originally derived from Western art, the results provided by Dr. Chanda were encouraging. She verified 98% of the categories could be applied to Western art and fine art and 90% could be used with Non-Western art. In addition, the diagnostic profile was highly rated for its applicability to popular arts, folk art, crafts, antiques and heirlooms, and cultural artifacts.

**Applicability to data.** Both judges indicated that a diagnostic assessment could be applied to a variety of data. These might include written statements about a single work of art or multiple works of art, written statements that have been verbally cued, and oral statements that have been transcribed.

**Diagnostic capabilities.** The independent judges agreed the diagnostic profile was capable of determining formal, descriptive, interpretive, and historical understandings of art. The experts indicated the diagnostic profile could assess understandings of students in kindergarten through grade 12. Furthermore, the diagnostic profile can assess the understandings of undergraduate and graduate students.

**Future Implications.** Dr. Johnson found the diagnostic profile to be clear "conceptually and in its application." She found both the instrument and the theoretical model to be "well-conceived." There was also support of using the diagnostic profile in art education research, program assessment, teacher assessment, student assessment, student self-assessment, and curriculum development. In conclusion, Dr. Johnson stated, "I find the implications for

teacher training and on-going self evaluation (by the teacher) to be extremely intriguing."

### **Criterion-Related Validity**

Two criterion-related studies were conducted with a 3rd/4th-grade art class and an 8th-grade art class. The focus of these studies was on the relationship between an external criterion and the diagnostic profile findings. The principal investigator observed 3rd/4th-grade and 8th-grade students and the instruction they received over a period of several weeks. A summary of these observations served as an external criterion.

Data in the form of written statements were gathered from students at the end of the observational periods. The written statements were then scored with the diagnostic profile. The degree of criterion-related evidence of validity was reflected in how well the external criterion predicted student performance as assessed with the diagnostic profile.

**3rd/4th-grade class.** Student writing samples were expected to be heavily influenced by the art teachers' instruction. During the nine-week observation/instructional period, students participated in the following activities:

**lecture format** - students viewed slides in a dark room as teacher read a script

**questioning** - students were asked to recite historical facts about an artist and her/his artwork work

**reinforced vocabulary terms** - students repeatedly identified various art terms

Results of this study demonstrated an extremely strong relationship between the external criterion and the assessment of the 3rd/4th-grade students' written statements with the diagnostic profile. This criterion-related evidence of validity offers verification regarding the diagnostic profile's ability to discriminate lower-order understandings.

**8th-grade class.** Observations of the 8th grade class revealed that students were receiving instruction that encouraged higher-order thinking skills such as:

**research** - students searched through classroom resources to learn about works of art

**"put it in writing"** - students found the words to express their ideas in a cohesive written format

**constructing arguments** - students were required to defend their stance with reasons

**criticism** - describing, analyzing, interpreting, and judging works of art

aesthetic inquiry - considering social and cultural issues in selection of an artworks for a museum

The teacher's implementation of these instructional strategies encouraged students to become active participants in their own learning. As predicted, higher-order thinking skills were exhibited by the vast majority of students in the 8th-grade class. Results of this study demonstrated an extremely strong relationship between the external criterion and the assessment of the 8th-grade student's written statements with the diagnostic profile. This criterion-related evidence of validity provided convincing evidence of the diagnostic profile's effectiveness in discriminating higher-order understandings.

#### **Construct Validity Study**

The objective of the construct validity study was to determine whether constructs in the diagnostic profile could be supported with a comprehensive data base. Results indicated that all diagnostic categories within the diagnostic profile could be supported by responses from the data base. A strong relationship was found between the data and the constructs in the diagnostic profile. Furthermore, the diagnostic profile discriminated lower-order understandings from higher-order understandings, and identified misunderstandings. Construct validity of the diagnostic profile was further amplified by the favorable results of the content validity and the criterion-related evidence validity studies.

#### **Summary of Reliability Studies**

Reliability of the diagnostic profile was estimated by an inter-rater reliability study and an intra-rater reliability study. Results of the reliability studies are summarized in the following sections.

##### **Rater Reliability**

Reliability tests of the diagnostic profile focused on inter-rater consistency and intra-rater consistency.

**Participants.** The principal investigator, an expert at using the diagnostic profile, served as the criterion rater. Since written statements about works of art are generated in art classrooms, it made sense for art teachers to serve as raters.

**Data base.** The diagnostic profile can be used to assess written or transcribed verbal statements about works of art. Since people respond in many different ways to works of art, a comprehensive data base of written and transcribed verbal statements was gathered for the studies. The data base reflected variance in several areas: size of class, age, sex, ethnicity, grade level, economic standard, cultural background, and academic abilities.

### **Inter-rater Consistency**

The focus of the inter-rater reliability study was on how consistently three different raters assigned scores to written statements using the diagnostic profile.

The teachers learned how to use the diagnostic profile during several practice sessions. Calculations of average reliability between pairs of raters were based on 109 checks. In the three comparisons, average reliabilities verified a "very strong association" in assignment of scores with the diagnostic profile as follows:

- .97 - teacher A and teacher B
- .96 - teacher A and the criterion rater
- .98 - teacher B and the criterion rater

The study presented convincing evidence that with a reasonable amount of training art teachers can be taught to reliably score written responses with the diagnostic profile.

### **Intra-rater Consistency**

The intra-rater consistency study involved the comparison of scores assigned to identical written statements at different points in time. Four weeks elapsed between the two scoring periods, and no practice sessions were held for the intra-rater study. The average reliability between time one and time two (i.e., four weeks later) was based on 42 checks. A "very strong association" was indicated between the scores as follows:

- .86 - Teacher A
- .88 - Teacher B
- .98 - Criterion Rater (e.g., principal investigator)

Although there was a noted decline in the abilities of the teachers to apply the diagnostic profile, the retention rate held strongly even after a four week interval. The data suggest that intervening time (four weeks), does not adversely affect an art teacher's ability to reliably score written responses to art using the diagnostic profile.

### **Significance of the Diagnostic Profile**

Grounded in current conceptions of learning, the diagnostic profile extends our ability to gauge student understandings of artworks. Categories within the diagnostic profile represent the students' knowledge in terms of knowledge base and knowledge-seeking strategies facets of learning. The instrument is also diagnostic in its ability to discriminate among lower-order understandings, higher-order understandings, and misunderstandings. The diagnostic profile encompasses terminology reflective of the field of art

education, and it is responsive to educators who have backgrounds in the arts. Studies indicated that the diagnostic profile was consistent and effective in assessing formal, descriptive, interpretive, and historical understandings of artworks. Based on traditional test and measurement techniques, the diagnostic profile is considered a reliable and valid assessment tool. The diagnostic profile has the potential to assess understandings of students in kindergarten through grade 12. Moreover, the diagnostic profile has been judged to be useful in assessing the understandings of college undergraduate and graduate students. The diagnostic profile assesses student understandings from written and/or transcribed verbal statements about works of art. Written statements can be stimulated by art reproductions, slides of artworks, actual works of art in museums or galleries, and the students' own artwork.

### References

- Koroscik, J. S. (in press). A comparative framework for designing visual art curricula. *Design for Arts in Education*.
- Koroscik, J. S., Short, G., Stavropoulos, C., and Fortin, S. (1992). Frameworks for understanding art: The function of comparative art contexts and verbal cues. *Studies in Art Education*, 33 (3), p. 154-164.