

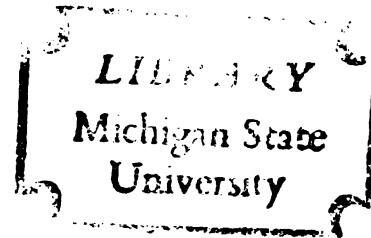
THE RELATIONSHIP OF
AESTHETIC JUDGMENT TO
SURROGATE DECISIONS
ABOUT WORKS OF ART IN
ART-NAIVE AND
ART-SOPHISTICATED
COLLEGE STUDENTS

Dissertation for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY



3 1293 10471 1019



This is to certify that the
thesis entitled
The Relationship of Aesthetic Judgment
To Surrogate Decisions About Works of Art
In Art-Naive and Art-Sophisticated
College Students

presented by
Marilyn Mendelson

has been accepted towards fulfillment
of the requirements for

Ph D degree in Educational Psychology

Major professor

Date June 28, 1974

ABSTRACT

THE RELATIONSHIP OF AESTHETIC JUDGMENT TO SURROGATE DECISIONS ABOUT WORKS OF ART IN ART-NAIVE AND ART-SOPHISTICATED COLLEGE STUDENTS

By

Marilyn Mendelson

The Problem

Standardized tests of aesthetic judgment measure the ability to perceive visual artistic quality through a comparative choice procedure. These tests' operational definition of aesthetic judgment appeared ambiguous in that the suggested decision criteria are artistic quality, pleasure, and satisfaction. The equivalence of aesthetic judgment and personal preference for art appeared to be an underlying assumption of this decision basis. The distinction between judgment and preference posited by aestheticians was seen as a challenge to this assumption and therefore to the construct validity of these tests.

To investigate the congruence between empirically obtained data and the philosophical premise of the unequivalence of judgment and preference, a comparison was made of the correspondence between pairs of decisions selected from the following triad: aesthetic judgment (J); personal preference (P); and ambiguous choice (A), this last representing the type of decision specified in standardized tests of aesthetic judgment. These decisions were taken in the six type-order combinations of AJ, AP, JP, PJ, AA, and JJ.

erend

paire

sente

Prese

tradi

Proce

by th

items

repro

and c

parat

admin

serie

serie

yield

gory

colle

were

order

sophi

backe

each

level

ing w

Additionally, the effects of art background and preference bias on the magnitudes of correspondence between paired decisions were examined. Art background was represented by the levels of art-naivety and art-sophistication. Preference bias was defined by the polarity of modern versus traditional art preference.

Procedures

The Art Decision Relationship Test (ADRT), constructed by the author and used in this study, is composed of 35 items, each a pair of projected 35 millimeter color slide reproductions of paintings in modern (M), traditional (T), and contrast (C) categories. The test task is that of comparative choice. Two randomized series of the same pairs are administered with a type of decision specified for each series. Selections of one of each pair made in the first series are compared with selections made in the second series, yielding a total test correspondence score and painting category subscores.

The ADRT was administered to 177 education and art college students at Michigan State University. Examinees were randomly assigned to the six testing levels of type-order but were admitted to the art-naive (AN) or art-sophisticated (AS) sample only by meeting preestablished art background criteria. In final form, the AN and AS samples each contained 60 Ss evenly distributed in the six type-order levels. AN Ss were characterized by limited or no art training whereas AS Ss were art majors or art minors.

ance (

intera

test o

examin

differ

prefer

Findin

effect

order

sied.

to ide

the no

popula

prefer

prefer

magnit

which o

the sur

A two-way fixed-effects univariate analysis of variance (ANOVA) balanced design was used to test the main and interaction effects of art-level and type-order on total test correspondence scores. Preference bias effects were examined using a one-way ANOVA of a difference score (the difference between the M and T subscores) for each of four preference subgroups.

Findings

F test analysis disclosed no significant interaction effect and no significant art-level main effect. A type-order main effect significant at the .0001 level was identified. The Tukey post hoc procedure at the .05 level was used to identify which type-order comparisons were contributing to the nonnull condition.

Within the qualification imposed by the established populations and measure, the major findings were:

1. Aesthetic judgment is not equivalent to personal preference nor to ambiguous choice.
2. Ambiguous choice is more similar to personal preference than it is to aesthetic judgment.
3. Neither art-level nor preference bias affect the magnitude of correspondence between decisions.

The findings call into question the validity of tests which operationalize aesthetic judgment through the use of the surrogate decision ambiguous choice.

THE RELATIONSHIP OF AESTHETIC JUDGMENT
TO SURROGATE DECISIONS ABOUT WORKS OF ART IN
ART-NAIVE AND ART-SOPHISTICATED COLLEGE STUDENTS

By

Marilyn Mendelson

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Personnel Services
and Educational Psychology

1974

Copyright by
Marilyn Mendelson
1974

The material quoted on pages 9 and 83 was taken from Evaluation of Learning in Art Education by B. G. Wilson and appeared in B. S. Bloom, J. T. Hastings, & G. F. Madaus, Handbook on Formative and Summative Evaluation of Student Learning. Copyright 1971 by McGraw-Hill Book Company. Used with permission of McGraw-Hill Book Company.

the A
Servi
Resea

ment
doctor
Willi
chair

ACKNOWLEDGMENTS

The writer wishes to thank the faculty and staff of the Art Department, the Department of Counseling, Personnel Services and Educational Psychology, and the Office of Research Consultation for their assistance in this study.

The writer is deeply appreciative of the encouragement she received from her apprenticeship research and doctoral committees, and of the guidance provided by Dr. William A. Mehrens in his capacity as doctoral committee chairman.

M. M.

CHAPTER

CHAPTER

CHAPTER I

CHAPTER 2

TABLE OF CONTENTS

CHAPTER I:	THE PROBLEM	1
	Need	1
	Purpose	3
	Hypotheses	5
	Theory	6
	The Nature and Bases of Aesthetic Judgment	6
	The Distinction Between Aesthetic Judgment and Personal Preference for Works of Art	10
	The Bipolar Factor of Modern Versus Traditional Art Preference	13
	Overview	14
CHAPTER II:	REVIEW OF LITERATURE	16
	The Development of Aesthetic Judgment Tests	16
	A General Factor of Aesthetic Judgment	24
	Secondary Factors of Aesthetic Judgment	25
	Summary	26
CHAPTER III:	INSTRUMENTATION	29
	Test Items	29
	Testing Procedure	32
	Test Scoring	35
	Test Reliability	36
	Summary	37
CHAPTER IV:	DESIGN OF THE STUDY	39
	Samples	39
	Design	43
	Testable Hypotheses	47
	Hypotheses I - IV	47
	Hypothesis V	51
	Hypothesis VI	51

CHAPTE

CHAPTER

APPEND

BIBLIOG.

	Analyses	52
	Summary	55
CHAPTER V:	ANALYSES OF RESULTS	57
	Two-Way Analysis of Variance	57
	Analyses of Hypotheses I - IV	60
	Analysis of Hypothesis V	62
	Analysis of Hypothesis VI	63
	Interpretation of Results	65
	Chi-Square Analysis of Questionnaire Responses	68
	Summary	68
CHAPTER VI:	SUMMARY AND CONCLUSIONS	73
	Summary	73
	Conclusions	76
	Discussion	78
	Type-Order	78
	Art-Level	79
	Preference Bias	80
	Implications for Future Research	80
	Implications for Instruction in Art Education	83
APPENDICES		
APPENDIX A:	Test Item Pairs of Paintings .	87
APPENDIX B:	Order of Slide Presentation for Series One and Two	89
APPENDIX C:	Test Packet Top Page: Directions for All Examinees .	90
APPENDIX D:	Level 1 Instructions: Ambiguous Choice-Judgment. . .	91
APPENDIX E:	Level 2 Instructions: Ambiguous Choice-Preference. .	92
APPENDIX F:	Level 3 Instructions: Judgment-Preference.	93
APPENDIX G:	Level 4 Instructions: Preference-Judgment.	94
APPENDIX H:	Level 5 Instructions: Ambiguous Choice-Ambiguous Choice	95
APPENDIX I:	Level 6 Instructions: Judgment-Judgment	96
APPENDIX J:	Questionnaire	97
APPENDIX K:	Statement Read to Examinees Prior to Testing	98
APPENDIX L:	Sequence of Testing	99
BIBLIOGRAPHY		101

TABLE

4

5

6

7

8

9

10

LIST OF TABLES

TABLE 1:	Characteristics of the Art-Naive (AN) and Art-Sophisticated (AS) Samples	42
2:	Major Areas of Study of Art-Naive Subjects	44
3:	Frequency Distribution of Preference Scores for Bipolar Preference Subgroups . . .	48
4:	ANOVA Group Mean Correspondence Scores and Standard Deviations	58
5:	Two-Way ANOVA Tables of Values	59
6:	T-Method Table of Values	61
7:	Mean Scores and F-Ratio Values for Tests of Difference Scores for Bipolar Preference Subgroups	64
8:	Frequencies of Responses to Choice and Instruction Questions of Questionnaire .	69
9:	Chi-Square Tests of Independence Data	70
10:	Summary of Findings and Conclusions	72

FIGURE

LIST OF FIGURES

FIGURE 1:	Two-Way Fixed-Effects Analysis of	
	Variance Design	45
2:	Pilot Study Interaction of Art-Level	
	and Order of Judgment and Preference	66

a long
experi
vigore
renewe
and th
tional

Need

aesthe
period
these
school
as pre
abilit
an inn
tent a
develo
achiev

judgme
task i
work o
examin
throug

CHAPTER I: THE PROBLEM

The development of measures of aesthetic judgment has a long but sporadic history dating back to the inception of experimental psychology. While development has not been vigorous, impetus for progress has come recently through the renewed interest in specifying the objectives of education and through the emphasis on evaluating the effects of educational programs.

Need

The first phase of test development in the area of aesthetic judgment in the United States occurred in the period of the 1920s through the 1940s. The emphasis of these early tests was on measuring artistic aptitude in high school students and young adults. The tests were designed as predictive instruments to aid vocational selection. The ability to make valid judgments about art was felt to be both an innate capability and a prerequisite condition for competent artistic production. In contrast, current test developers have turned from measuring aptitude to measuring achievement gained through instruction (Wilson, 1971).

A usual procedure of published tests of aesthetic judgment is to present sets of visuals to examinees whose task it is to select one of each set as the better or best work of art. The degree of correspondence between an examinee's responses and art expert judgment established through consensus is examined. High agreement is generally

equa

in

acco

ambi

manu

perc

stan

make

basi

plea

is t

of a

theo

ment

is l

been

the

aest

Give

deci

exam

exis

shou

disc

need

equated with aesthetic sensitivity, good taste, or competency in making aesthetic judgments. Examinees are often ranked according to their level of aesthetic judgment.

One of the inadequacies of these tests is their ambiguity in operationally defining aesthetic judgment. Test manuals generally define aesthetic judgment as the ability to perceive artistic quality. Operationally, however, several standardized tests of aesthetic judgment ask examinees to make decisions that are defined as aesthetic judgments on the basis of a work of art's artistic quality or its ability to please or satisfy. The underlying assumption of this request is that aesthetic judgment and personal preference for works of art are synonymous types of decisions. Although art theorists have argued that a distinction exists between judgment and preference, empirical support for this distinction is lacking. The manner in which existing instruments have been constructed and used has rendered them insensitive to the detection of such a difference.

In their instructions to examinees, leading tests of aesthetic judgment actually encourage preference choices. Given that judgment and preference are discrepant kinds of decisions, these tests may not accurately reflect the examinee's level of aesthetic judgment. If a discrepancy exists, the validity of future tests of aesthetic judgment should be increased by explicitly calling for judgments while discouraging preference choices.

Thomas (1965, pp. 171-176) called attention to the need to clarify what it is that tests of aesthetic judgment

meas-

and a

grasp

prese

valid

nonsu

perso

perso

dicti

is al

educa

the i

about

betwe

prefe

needs

Purpo

that

makin

exami

what

ment

and a

of de

measure. He suggested that preference (likes and dislikes) and analysis (understanding, ability to analyze structure and grasp cultural significance) be measured separately. This present study was designed to address this issue of construct validity and to provide empirical grounds for the support or nonsupport of the equivalence of aesthetic judgment and personal preference.

The possible distinction between aesthetic judgment and personal preference has implications that go beyond the predictive utility of aesthetic judgment tests. The distinction is also relevant to the development and evaluation of art education programs. A goal of such programs is to increase the individual's ability to make reasoned critical judgments about artistic quality. Investigations of the differences between and similarities of aesthetic judgment and personal preference could help direct art curriculums by identifying needs and by pointing to promising instructional practices.

Purpose

This study was designed to examine the relationships that exist among types of decisions used as the basis for making selections of works of art. The types of decisions examined were aesthetic judgment, personal preference, and what was called ambiguous choice. At issue was whether judgment and the two surrogate types of decisions -- preference and ambiguous choice -- are equivalent or dissimilar kinds of decisions.

made

quali

basis

uous

afore

left

sever

relat

if di

desig

purpo

equiv

either

degree

prefe

made?

consis

ground

releva

influe

betwee

Judgment was defined as a critical evaluation of art made on the basis of general aesthetic value or artistic quality. Preference was defined as a decision made on the basis of personal appeal, pleasure, or satisfaction. Ambiguous choice was defined as a decision made on either of the aforementioned bases. Ambiguous choice was intentionally left open-ended in keeping with the questioned approach of several tests of aesthetic judgment.

The main purpose of this study was to compare the relationships of these decisions when taken in pairs to see if differences in relationships exist. This study was designed to answer several questions specific to the main purpose. These questions were:

1. Are judgment, preference, and ambiguous choice equivalent kinds of decisions, or is judgment different from either of the other types?
2. Is ambiguous choice like judgment to the same degree that it is like preference?
3. Is the degree of agreement between judgment and preference affected by the order in which these decisions are made?
4. Are judgment and ambiguous choice held with equal consistency?

Two additional purposes guided this study. Art background and preference bias were identified as two variables relevant to decision-making about works of art. The possible influences of these variables on the degree of agreement between pairs of decisions were examined. The purpose of

this

effe

of a

bias

fact

degre

Hypo

as t

(Hyp

amon

amon

to a

pain

of a

arb.

of a

the

Cha

this study pertaining to art background was to examine the effects of art-naivety and art-sophistication on the degree of agreement between decisions. Pertaining to preference bias, the purpose was to examine the effect of the bipolar factor of modern versus traditional art preference on the degree of agreement between decisions.

Hypotheses

The hypotheses investigated in this study¹ dealt with:

1. The relationships among types of decisions used as the basis for making selections of works of art (Hypotheses I - IV to follow).

2. The effect of art background on the relationships among types of decisions (Hypothesis V to follow).

3. The effect of preference bias on the relationships among types of decisions (Hypothesis VI to follow).

Although the hypotheses do not so state, each pertains to a college student population and works of art refer to paintings.

Hypothesis I: Aesthetic judgment decisions about works of art are not equivalent to personal preference nor to ambiguous choice decisions about works of art.

Hypothesis II: Ambiguous choice decisions about works of art are more similar to personal preference decisions than they are to aesthetic judgment decisions.

¹Testable forms of the hypotheses are stated in Chapter IV.

jud

the

with

degr

degr

Theo

purpo

topic

perso

prefe

The Na

knowle

philos

clarif

art (B

and Sm

critici

Hypothesis III: The degree of agreement between judgment and preference is affected by the order in which these decisions are made.

Hypothesis IV: Aesthetic judgment decisions are held with greater consistency than are ambiguous choice decisions.

Hypothesis V: Art background has an effect on the degree of agreement between decisions about works of art.

Hypothesis VI: Preference bias has an effect on the degree of agreement between decisions about works of art.

Theory

The theoretical concerns that are relevant to the purposes and hypotheses of this study fall into the following topic categories:

1. The nature and bases of aesthetic judgment.
2. The distinction between aesthetic judgment and personal preference for works of art.
3. The bipolar factor of modern versus traditional art preference.

These concerns are discussed in turn in this section.

The Nature and Bases of Aesthetic Judgment The field of knowledge known as aesthetics can be thought of as the philosophy of criticism. The concern of aesthetics is to clarify and confirm critical statements made about works of art (Beardsley, 1966, p. 308). Feldman (1967, pp. 470-486) and Smith (1971, pp. 473-480) distinguished four phases of criticism:

Gots

tand

appl

the

that

pers

to d

phas

aest

as b

Thes

1. Description: listing the features presented in a work and analyzing the methods employed.

2. Formal analysis: describing the qualities of parts and their relationships.

3. Interpretation: explaining meaning. Generally, an hypothesis is formulated which attempts to explain the work and its effect on the viewer. Criticism may stop at this point or proceed to the fourth and most difficult phase.

4. Judgment: evaluating to decide aesthetic merit.

Somewhat different phases of criticism were posited by Gotshalk (1966, pp. 343-344), but he too recognized the importance of the judgmental phase where relevant criteria are applied to assess value. Gotshalk (1966, p. 353) stated that the function of criticism is to formulate value statements that, even if not universally true, "are translatable into the personal value language and understanding of others."

Feldman (1967, pp. 489-495) suggested that judging art to decide the question of worth is guided by:

1. Relating the work of art to the widest possible range of similar works.

2. Discerning the purpose of the work.

3. Discerning how the work departs from historical precedents.

4. Discerning the relationship of the work to its own time period.

5. Identifying the artistic problem and the degree of artistic originality.

6. Identifying the level of craftsmanship and technique.

An additional and essential concern of the judgmental phase of criticism is to assess the work of art against aesthetic principles or standards. Laws of composition, such as balance and movement, are examples of such standards. These standards embody empirically acquired information,

facili

used

pp. 2

serve

(1934

as di

judgm

the b

autho

can v

(1966

sary

the a

crite

very

attes

affec

be ac

are a

judgm

posse

of wo

indiv

as:

intui

facilitate the artist's work, and contribute to the criteria used for determining a work's excellence (Bullough, 1957, pp. 24-25).

Whether aesthetic standards are absolute precepts or serve only as guiding principles is controversial. Burt (1934, p. 294), for one, supported their universal validity as did the developers of standardized tests of aesthetic judgment. Dewey (1934, p. 307), on the other hand, rejected the blind acceptance of aesthetic standards, arguing that the authority of rules is to be questioned and that works of art can violate standards and still achieve excellence. Weitz (1966, p. 85) stated that art never has had a "set of necessary and sufficient properties": major art theories² through the ages have revised requisites and have stressed different criteria that direct the experiencing of works of art. The very concept of art is an open one subject to revision and attests to the absence of dogma.

However traditional objective criteria have been affected by prevailing theories and given that they need not be accepted as universal truths, these criteria have survived, are available, and may be applied when formulating aesthetic judgments. While the art-knowledgeable individual may possess much relevant information on which to base decisions of worth, it should not be assumed that the art-naïve individual lacks any basis for making such decisions. On the

²Weitz (1966, p. 84) identified the influential theories as: emotionalism, formalism, organicism, intellectualism, intuitionism, and voluntarism.

contra

are c

expres

abili

than

1938;

Proth

aesth

catic

a fur

intui

criti

the p

in a

seems

indi

poss

ledge

meth

contrary, belief in the premise that art-naive individuals are capable of evaluating the worth of works of art was expressed by Wilson (1971, p. 517):

Reasoned critical judgments of works of art usually are grounded in the standards which follow from the major theories of art, even though students who make aesthetic judgments might be unaware of the theory and, perhaps at least in part, unaware of the standards which they are applying.

Meier (1942, p. 16) defined aesthetic judgment as the ability to discern "universal" qualities intuitively rather than logically. Research data (Burt, 1934, p. 292; Dewey, 1938; Eysenck, 1940, 1957; Child & Iwao, 1966; Child, Ford & Prothro, 1966) tend to support a general, innate factor of aesthetic judgment that is independent of art training, education, intelligence, or culture.

Accepting the assumptions that aesthetic perception is a fundamental innate ability (Meier, 1942, p. 16) and that intuitively known aesthetic standards may be applied without critical awareness (Wilson, 1971, p. 520), one must concede the possibility that even individuals with no formal training in art are capable of making reliable aesthetic judgments. It seems reasonable to conclude that, on the average, art-naive individuals in comparison to art-knowledgeable individuals possess to a lesser degree the intuitive and acquired knowledge necessary for determining value according to critical methods.

The

Pref

and

to v

ence

reac

crit

tinc

refe

subj

The

perce

reti

crat.

indi

poss

the

reco

by a

quite

Burt

that

The Distinction Between Aesthetic Judgment and Personal Preference for Works of Art

Granting that both art-naive and art-sophisticated individuals use aesthetic standards but to varying degrees when evaluating art, the possible influence of personal preference, or taste, on one's total reaction to a work of art needs to be recognized. Taste and critical judgment may be discrete phenomena. A major distinction between judgment and preference is that the former refers to general artistic merit while the latter refers to subjective appeal or the degree and direction of affect felt. The hypothesized criteria that underly judgment are cognitive, perceptual, and critical ones. In contrast, preference theoretically derives from affective bases.

Decisions about works of art in the form of idiosyncratic likes and dislikes are not restricted to art-naive individuals. As Ushenko (1953, p. 204) put it, "it is possible to identify a masterpiece without liking it." Even the critic is not immune to biases and predilections: he may recognize the aesthetic value of a work of art but may be led by a philosophical premise or a personal bias to prefer art quite different in style (Feldman, 1967, pp. 450-451).

Other than biases of this sort, criteria identified by Burt (1934, pp. 280-286) may reflect the kinds of influences that act on personal preference. These criteria are:

1. Subjective: emotional and psychological implications of a work of art.
2. Associative: a work's ability to evoke ideas or recall of experience.
3. Character type: ascribing human qualities to a work's elements; empathizing with portrayed objects.

expe

kind

judg

mean

sati

then

Beca

crit

ambi

judg

ambi

tic

woul

dec

gre

ind.

ind.

acq.

valu

mig.

pre.

that

ence

This

Jud.

It is obvious that appeal to a prior emotional experience is a poor determinant of aesthetic merit. This kind of consideration should have no effect on aesthetic judgment if it is indeed established solely through critical means and innately known standards. Its influence on the satisfaction and pleasure one may derive from art, and therefore its influence on preference, may be profound. Because tests of aesthetic judgment direct examinees to the criteria of pleasure and satisfaction, it was felt that ambiguous choice would reflect preference more than it would judgment. Further, because the directions to examinees are ambiguous in that they also include the criterion of aesthetic quality, it was felt that ambiguous choice decisions would fluctuate more than would specifically defined decisions; i.e., that judgment decisions would be held with greater consistency than would ambiguous choice decisions.

Earlier, the conclusion was reached that art-naive individuals are not as well qualified as art-knowledgeable individuals to determine value according to innately known or acquired knowledge. If objective criteria for determining value are not readily available to art-naive individuals it might be that subjective criteria, of the kinds operating in preference decisions, are substituted in their stead; i.e., that the criteria for judgment and the criteria for preference are not well differentiated in an art-naive population. This suggested that for an art-naive population aesthetic judgment parallels personal preference.

expect

learn

indep

restr

capab

a mul

dual

these

posed

to as

same

on a

expect

respo

knowl

of a

ship

betwe

sophi

when

diffe

prefe

for t

secu

art t

diffe

In contrast, art-sophisticated individuals were expected to possess the critical means, whether innate or learned, that allow them to form judgments of worth quite independent of personal biases. Since art even when restricted to painting is composed of numerous aspects, each capable of diverse treatment, the field's products represent a multitude of styles. While the art-sophisticated individual can appreciate the value and contribution of each of these styles, on a subjective level he is apt to be predisposed to prefer some over others. It was therefore logical to assume that judgment and preference are not one and the same for the art-sophisticated population, that each rests on a different set of criteria. This reasoning led to the expectation that judgment and preference more closely correspond in an art-naive population than they do in an art-knowledgeable population.

This expectation, however, was tempered by the finding of a pilot study (Mendelson, 1973) investigating the relationship between judgment and preference, in which agreement between judgment and preference choices was higher for art-sophisticated individuals than for art-naive individuals when preference decisions preceded judgment decisions. No difference between groups was found when judgment preceded preference decisions. One possible explanation accounting for this occurrence is that the art-sophisticated are more secure in and accepting of their subjective reactions towards art than are the art-naive, and this tends to bias judgments differentially.

The

Pre

jud

to a

been

vers

simpl

(Chi

or c

in f

of j

tion

more

This

ence

trad

pola

judg

that

sent

the

for a

judg

repre

The r

tive

ence

The Bipolar Factor of Modern Versus Traditional Art Preference

In addition to the general factor of aesthetic judgment that has found support, secondary factors that help to account for the variability of aesthetic judgment have been identified. These bipolar factors are: (1) extra-version-introversion (Eysenck, 1941); (2) complexity-simplicity (Child & Iwao, 1968); and (3) openness-rigidity (Child, 1962; Pyron, 1966). Whether several factors exist, or only one, is difficult to determine. These variables may in fact refer to the same dimension of personality. A dualism of judgment does seem to exist that separates complex, emotionally expressive art, on the one hand, from more simple, more representational, and less-colorful art on the other. This dichotomy, in general, is characteristic of the differences that distinguish modern and contemporary art from traditional art.

For those individuals at either extreme of this bipolar preference continuum, the degree of agreement between judgment and preference might be different for works of art that represent the preference bias than for works unrepresentative of that bias. The bipolar factor hypothesis of the pilot study cited earlier (Mendelson, 1973) held that, for art-naive individuals, the degree of agreement between judgment and preference would be higher for those visuals representing the polar preference than for opposing visuals. The rationale underlying this view relates to the comparative lack of differential criteria for judgment and preference hypothesized above for an art-naive population. It was

felt

attr

vidi

ing

art-

The a

was s

polar

prefe

rathe

effec

for m

eithe

Overv

inves

lined

revie

aesth

facto

determ

decis

judgme

presen

in Cha

felt that those works of art providing pleasure would be attributed value consistently, while those works not providing pleasure would be more difficult to "judge," resulting in less consistency between judgment and preference.

A finding in reverse of that predicted occurred for art-naive individuals with a preference for traditional art. The agreement between judgment and preference for this group was significantly lower for those visuals representing the polar preference than for those visuals opposed to the preference. It appeared as if preference served to confound rather than stabilize decisions for this group. No bipolar effect was found for art-naive individuals with a preference for modern art, nor for art-sophisticated individuals at either preference extreme.

Overview

The body of this thesis is a report of an empirical investigation of the hypotheses derived from the theory outlined above. In Chapter II, the pertinent literature is reviewed. The review traces the development of tests of aesthetic judgment and deals with the general and secondary factors of aesthetic judgment.

In Chapter III the instrument and procedure used to determine the relationships that exist between pairs of decisions selected from the triad of decisions -- aesthetic judgment, personal preference, and ambiguous choice -- are presented. A discussion of the design of the study is given in Chapter IV. The manner of selection and the

chara

sampl

The p

eses

analy

and i

eses v

proce

from

prefe

test

prese

characteristics of the art-naive and art-sophisticated samples of college students used in this study are discussed. The plan for data collection is given along with the hypotheses in testable forms and the statistical tests used to analyze the data.

In Chapter V, the findings of the study are analyzed and interpreted for each hypothesis of interest. The hypotheses were generated in part from an examination of the procedures and premises of tests of aesthetic judgment and from research into the personality correlates of judgment and preference for art. A review of the literature concerning test development and the factors that relate to judgment is presented next.

aesth

chapt

and s

along

The

theti

G. T.

and a

pref

aesth

p. 3

1865

rect

the

Bori

tici

Fech

the

simu

1934

CHAPTER II: REVIEW OF LITERATURE

The development, premises, and operations of tests of aesthetic judgment are examined in the first section of this chapter. In remaining sections, a discussion of the general and secondary factors of aesthetic judgment is presented along with a summary of literature.

The Development of Aesthetic Judgment Tests

In modern times, the quantificative analysis of aesthetic judgment began about a century ago with the work of G. T. Fechner (Pickford, 1955, p. 914), a German philosopher and a founder of experimental psychology. Fechner studied preference for proportions and its relationship to the known aesthetic doctrine of the golden section (Eysenck, 1957, p. 326). In Fechner's first reported aesthetic experiment in 1865, examinees were asked to state their preference for rectangles of varying dimensions. The proportion³ known as the golden section was invariably pleasing (Introduction by Boring, 1966, in Fechner, 1860, p. xvi).

One of Fechner's concerns was determining the authenticity of two Madonna paintings, both attributed to Holbein. Fechner conducted a public opinion poll to decide which of the two was more beautiful. The paintings were exhibited simultaneously and viewers were asked to state their

³The equation for this proportion is given in Burt, 1934, p. 298.

preferences. Only 113 out of 11,000 viewers cast opinions. Many of them were submitted by art experts, were felt to be biased, and were therefore rejected (Introduction by Boring, 1966, in Fechner, 1860, p. xvi).

Fechner's experimental methodology influenced Birkhoff (1933, pp. 33-44, 205), an American mathematician who developed a formula for determining the aesthetic worth of simple art forms, such as polygons. The formula is $M = O/C$, where M = aesthetic measure; O = order, harmony, and symmetry; and C = complexity. The formula relates to the traditional aesthetic demand for unity coupled with variety. Even when applied to sufficiently restricted forms, the formula rests on an involved value system and its reliability and validity are questionable.

The approach adopted by British psychologists in the 1930s departed from the use of mathematical models. Burt (1934, pp. 279-289) was interested in the feelings of the spectator, the study of which he called the psychology of artistic appreciation. Burt hoped to find support for the existence of a general factor of aesthetic judgment that overrode irrelevant associations, as well as to identify bipolar factors for specific types of artistic appreciation (Pickford, 1955, pp. 916-917).

In the early 1930s Burt and Pelling devised the Picture Postcard Test, in which examinees ranked fifty picture postcards of works of art in order of preference (Pickford, 1955, p. 917). The art varied in degrees of excellence from master paintings to crude greeting card

illustrations. Agreement of ratings among artists and art critics was high. Burt (1934, pp. 289-290) felt that a fundamental factor guided the ordering and he attributed the minor divergence to points of view.

Bulley (1934, p. 162) and Burt collaborated on a broadcasted test of art judgment for the B.B.C., the test material and questionnaire of which were published in The Listener in January, 1933. The test contained nine pairs of contrasted objects of household use. Each pair had a better and a worse example as judged by a small group of art critics. The purpose of the test, to which 10,000 people responded, was to gain information about public taste regarding the decorative arts. A unique feature of the test was that respondees were asked to state the motives for their choices.

The prototype of standardized aesthetic judgment tests developed in the United States is the Meier-Seashore Art Judgment Test (Meier & Seashore, 1928) which resulted from research conducted at the State University of Iowa. The test, administerable as low as the 7th grade, was designed as a predictive measure for vocational selection. The test was based on the rationale that aesthetic judgment is the most important factor in determining success as a creative artist.

Aesthetic judgment was defined as the capacity for perceiving quality in aesthetic phenomena through sensitivity to universally held aesthetic standards. It was hypothesized that this sensitivity was primarily perceptual and independent of art information, training, maturity,

inte

did

pair

are

some

is o

alth

on t

of e

of t

a re

exam

juni

citie

samp

7 - 1

aesth

ceptu

intel

not c

were

prese

dark

intelligence, or general ability; i.e., that the capacity did not depend on "learned qualities."

The test is in book format and is composed of 125 pairs of black and white pictorials. The two of each pair are a reproduced work of art and an altered version in which some design principle has been violated. The unlike portion⁴ is obvious on the plates and is also briefly described, although the principle involved is not named, for each pair on the answer sheet. Examinees are asked to compare the two of each pair, note the unlike portion, and select the better of the two (the more pleasing, artistic, or satisfying).

The test is untimed. The answer sheet functions as a record of responses as well as a short questionnaire of the examinee's art background. The test was normed on 1850 junior and senior high school students in four midwestern cities. Test scores were also obtained for several adult samples. Prognosis of success in art for students in grades 7 - 12 was based on the percentile rank achieved.

The data provided some support for the hypothesis that aesthetic judgment is primarily an innate, unlearned, perceptual capacity. Low correlations with measures of general intelligence were found. Training and art information were not crucial to performance, although those individuals who were trained in art tended to perform better than those who

⁴In actuality, small unintended differences are present in many pairs. They tend to be differences in light-dark value and in the amount of visible detail.

were not. This does not imply a causal connection between training and test performance, however.

The test achieved satisfactory reliability coefficients ranging from .71 to .85. No data were presented in support of the test's validity as a predictive measure. That the test ever gained wide use is doubtful. In terms of representing the field of art, the pictorials were inadequate in that few schools of art were represented, three dimensional art was almost excluded, and the plates were small, poorly reproduced, and lacked color.

The Art Judgment Test, Unit I of the Meier Art Tests (Meier, 1942), is a 1940 revision of the Meier-Seashore Art Judgment Test and contains the earlier test's best 100 items as determined by biserial correlation. In scoring, the 25 of these 100 items having the greatest diagnostic validity are double-weighted. The black and white pictorials appear again in book format except that they are somewhat larger and improved in reproductive quality. The basic assumptions of the Meier-Seashore test -- the importance of aesthetic judgment for artistic competence and the universality of art principles -- are maintained.

Percentile norms were given for junior high school, senior high school, and college-adult samples. All were based on large sample testing. To increase the test's predictive validity, the norm groups were drawn from art student populations rather than from general student populations.

book

.64

abil

was

art

This

of s

tion.

some

meri

of m

struc

and

domin

tion,

shape

appea

black

the p

sever

prefe

appea

corre

to di

In a review for the Fourth Mental Measurements Year-book Schutz (1953) commended the test's reliability (.70 to .84 using the Spearman Brown correction for split-half reliability) and ease of administration and scoring. The test was criticized for neglecting the comprehensive nature of art in attending to only the compositional aspect of art. This restriction, Schutz felt, made the test a good measure of sensitivity to design but not necessarily to art. Additional criticisms were that: the design changes effected in some visuals were insignificant, making them poor items; the merit of some of the works was questionable; and the field of modern art was ignored.

The Design Judgment Test (Graves, 1948) was constructed to measure the degree to which individuals perceive and respond to the basic aesthetic principles of unity, dominance, variety, balance, continuity, symmetry, proportion, and rhythm. The test items are abstract designs of shapes and lines rather than works of art. The designs appear in sets, each a pair or triad of designs in greys, black, and white. In each set one design is concordant with the principles above, while the other(s) violates one or several of the principles.

Examinees are instructed to select the design they prefer in each set even though the designs may appear equally appealing. The test score is the number of items marked correctly according to expert opinion. The test was found to differentiate between art student and non-art student

grou

meas

rega

and

acte

test

asym

desi

sever

each

test

chara

natur

Welsh

ident

with

artis

of th

measu

ment

were

simpli

these

struct

groups. As with the Meier test, the Design Judgment Test measures sensitivity to compositional organization and disregards other salient aspects of art, such as content, color, and style. Symmetrical balance is such a pervasive characteristic of the correct designs that it is conceivable the test actually measures preference for symmetrical versus asymmetrical organization.

The Welsh Figure Preference Test (Barron, 1953) was designed to aid psychiatric diagnosis. Test items were several hundred black line drawings on 3 x 5 white cards for each of which examinees indicated a liking or disliking. The test was not successful as a diagnostic tool and various characteristics were factor analyzed to investigate the nature of the preferences expressed. Quite accidentally Welsh isolated the bipolar factor of simplicity-complexity identified at an earlier time by Eysenck. Simplicity coupled with bilateral symmetry and complexity with asymmetry. The artists in the control group clustered at the complexity pole of the factor.

Barron came across Welsh's findings in searching for a measure of artistic discrimination to include in an assessment battery. Additional samples of artists and non-artists were administered the Welsh Figure Preference Test and the simplicity-complexity factor maintained consistency with these new groups (Barron & Welsh, 1952, p. 203).

Using the results of item analysis Barron (1953) constructed a revised scale, the Barron-Welsh Art Scale, as a

meas
achi
vali
disc
crim
was
art

ment
Art T
high
mat o
each
compo
to po
ity m
balanc
used i
test m
world

groups
studen
percen
college
supplen
earlier
ability

measure of simplicity-complexity preference. The scale achieved high degrees of reliability and validity when cross-validated. Rosen (1955) reported that the test did not discriminate for levels of art training but that it did discriminate significantly between artists and non-artists. It was also valid for predicting originality and ability among art students.

The most recently published test of aesthetic judgment is the Aesthetic Perception Test, Unit II of the Meier Art Tests (Meier, 1967). The test is suitable for those of high school age and older. Unlike the pair comparison format of the preceding Meier test, Unit II contains 50 items, each of which has four versions with slightly different compositions. The task is to rank order the four from best to poorest. Directions to examinees indicate that superiority may be determined by considering proportion, unity, balance, or satisfaction derived. Unit II is meant to be used in conjunction with the 1940 Unit I test although either test may be used alone. Unit II more broadly represents world art than does the earlier Meier test.

Significant mean score differences were found for groups of artists, college art students, and high school students taking art. The preliminary test manual provides percentile norms for a high school art student sample and a college-adult sample. The 1967 manual is intended as a supplement to the 1942 manual and refers test users to the earlier manual for interpretation of the norm table. Reliability and validity data are not provided and in general

the
set
vic

A G

fac
as v
cal
vis
cult

exis
of a
cult
reach
sive
aesth
deter

suppo
from
accor
were r
agreem
people
about
esthet.

the manual information does not meet minimal standards as set by the American Psychological Association (see French & Michael, 1966).

A General Factor of Aesthetic Judgment

Hypotheses that claim the existence of a general factor of aesthetic judgment have stimulated research studies as well as the development of standardized tests. Basically, these studies ask if there is an objective factor of visual aesthetic sensitivity that is independent of training, culture, and other secondary concerns.

Eysenck (1940, 1957) believed that he established the existence of this general factor and that it was independent of age, sex, intelligence, artistic sophistication, race, or culture. In earlier studies Burt (1934) and Dewar (1938) reached similar conclusions. Although not claiming conclusive proof, Eysenck maintained that the general factor of aesthetic judgment is innate, biologically derived, and determined by a property of the central nervous system.

Child and Iwao's study (1966) of Japanese potters supported the universality of aesthetic values. Individuals from a different culture who engaged in art were more in accord with American expert judgment than were Americans who were not involved in art. Child hypothesized that this agreement "lies partly in the independent discovery, by people in differing cultural traditions, of similar facts about the adequacy of particular works for satisfying esthetic interests." In another study that examined cultural

influences on aesthetic judgment, Child, Ford and Prothro (1966) found additional evidence of transcultural agreement of aesthetic evaluations among people actively interested in art.

The theory of a universal aesthetic factor is not unanimously held. Eysenck's work has been criticized (Pickford, 1955, p. 921) for not including subjects from extremely divergent racial or cultural backgrounds. In a study of aborigines and whites in Australia, McElroy (1952) found no support for the existence of a general factor "based on inherited predispositions." McElroy argued that cultural perceptual conditioning is extremely influential in assigning values to art and that this cultural influence prevents universality.

One can accept McElroy's conclusion and still argue that a general factor exists within the confines of the established cultural tradition. The evidence does seem to support a general factor that overrides cultural influence. This need not be seen as incompatible with what is known about the historical transformations that art has undergone.

Secondary Factors of Aesthetic Judgment

In addition to a general factor, secondary factors that account for the variability of aesthetic judgments have been identified. These bipolar factors are: extraversion-introversion, complexity-simplicity, and openness-rigidity. Whether several factors exist, or only one, is difficult to determine. The different variables may well be reflecting

the same dimension of personality. A dualism of judgment does seem to separate complex, colorful, emotionally expressive art, on the one hand, from more simple, less-colorful, more representational art on the other.

Eysenck (1941) found positive correlations between preference for modern art and extraversion, and between preference for traditional art and introversion. Child (1962) found a high correlation between art preference and tolerance for ambiguity and ambivalence, whereas preference and aesthetic value correlated only slightly.

Studies by Child (1965) and Child and Iwao (1968) found significant positive correlations between aesthetic judgment and independence of cognitive judgment. Subjects high in aesthetic judgment tended to be inquiring, alert to the potentialities of experience, desirous of deep experiences, and receptive to complexity and novelty.

Pyron (1966) compared simplicity-complexity of social orderings and attitudinal rigidity in relation to preference for popular, classical, and avant-garde art forms in literature, painting, and music. He found a significant positive relationship between preference for avant-garde art with complexity and with four aspects of rigidity: low on dogmatism and rejection of people, high on acceptance of change and self-reliance.

Summary

Standardized tests of aesthetic judgment are predicated on the research supported assumption of a general

perC

sona

pref

vers

rigi

facto

valid

durin

slowl

The S

pp. x

tests

new a

of th

may r

rary s

involv

tific

transm

and co

of art

needed

aesthe

structe

or rela

perceptual factor of aesthetic sensitivity. Several personality bipolar variables that correlate with judgment and preference have also been identified. They are: extraversion-introversion, complexity-simplicity, and openness-rigidity. Unfortunately, research studies investigating the factors of aesthetic judgment have not resulted in more validly constructed standardized tests.

The measurement of aesthetic judgment which started during the infancy of experimental psychology has progressed slowly in comparison to the measurement of other abilities. The Seventh Mental Measurements Yearbook (Buros, 1972, pp. xxxi, 521-522) reported 14 new or revised fine arts tests but they represented only 1.2% of the total number of new and revised tests. Only four of the 14 were in the area of the visual arts. Of these four, only one was a new test.

This past lack of vigorous psychometric development may related to the fringe position of aesthetics in contemporary society, the difficulty of measuring the constructs involved, or the degree of content familiarity that scientific inquiry requires. Considering art's role as a transmitter of culture and as a medium of human expression, and considering too the long supported view of the importance of art education for all students, renewed activity is both needed and justifiable.

Indications of an upsurge of interest in measuring aesthetic capabilities exist. Several instruments constructed in the late 1960s (see Wilson, 1971) that measure or relate to aesthetic judgment are not yet available in

published form. This recent activity is seen as reflective of the current emphasis on evaluating the effects of educational programs.

The relatively few extant standardized tests of aesthetic judgment have shortcomings. We cannot say precisely what it is they measure (Thomas, 1965, p. 167). None is an exemplar of what investigators would deem sufficient to the task of measuring aesthetic judgment. The inadequacies of existing tests pertinent to this study may be summarized as follows:

1. The ambiguity of their operational definition for aesthetic judgment. Whether these tests measure judgment of aesthetic value or personal preference, or both indiscriminately, requires clarification.

2. The lack of provision for securing data concerning the underlying dimensions that influence decisions. Related to this problem are the tests' emphasis of composition and their neglect of many aspects that contribute in important ways to art.

3. The insignificantly sized, black and white format of test items that makes them poor substitutes for the works of art they are intended to represent.

CHAPTER III: INSTRUMENTATION

A test designed to measure the degree of relationship among paired decisions about works of art, constructed for an earlier pilot study (Mendelson, 1973), was modified for use in this investigation. The test, entitled the Art Decision Relationship Test (ADRT), and the conditions of its administration are described in this chapter.

Test Items

While the visual arts encompass painting, sculpture, graphics, and other forms, the ADRT visuals were restricted to American and European paintings. Painting has been a decidedly important form of the arts and the one upon which aesthetic criticism has focused. It is also an art form frequently encountered in everyday situations. The restriction of visuals to paintings has characterized tests of aesthetic judgment and is supported by several administrative limitations such as the problem of adequately representing three dimensional art.

The test items consist of 35 pairs of visuals, plus a practice pair, in the form of projected 35 millimeter color slides of paintings varying in subject matter and style. All represented painters were of established merit and only paintings felt to be valid works were chosen.

Such visuals, or tokens, function as reminders of actual art objects. Examinees are in effect being asked to look at the token as if it were the referent object. This

regue

what

where

less

consi

the r

and w

and G

which

in the

and mo

into t

on a c

p. 327

point

the pe

tradit

in Appe

traditi

request seems more reasonable when the token approximates what it is intended to represent than in other instances where correspondence between token and referent object is less (Sloan, 1970, p. 192). This and group administration considerations led to the choice of projected slides and the rejection of the common practice of using small black and white test booklets.

An art object is a composite of many aspects. Norris and Goodwin (1971, pp. 9-10) listed the aspects of art, which sum to expressive content:

1. Sensory qualities, as color, line and shape.
2. Compositional elements, as structure, movement, and balance.
3. Expressive quality, as mood, feeling, and emotion.
4. Subject matter, as objects, themes, events, ideas, and symbols.

Diverse treatments of the aspects of art are evident in the 36 test pairs because they include both traditional and modern-contemporary art. The division of the paintings into traditional and modern-contemporary categories was based on a commonly used classification system (as in Canaday, 1959, p. 327) that established post-impressionism as the starting point of modern art. Thus, modern-contemporary art occupies the period from the late 19th century to the present, and traditional art the longer period preceding.

A listing of the 36 test pairs of paintings is given in Appendix A. The pairs were chosen according to contrast, traditional, and modern classes:

1. Contrast (C) pairs. Sixteen pairs of paintings contrasted a modern or contemporary painting with a traditional painting. As much as possible, subject matter was held constant within pairs. The variable art aspects were sensory qualities, compositional elements, and expressive quality. The contrast was primarily one of style and manner of expression.

2. Traditional (T) pairs. Ten pairs of paintings contrasted two traditional paintings. Within each pair, paintings were those of two contemporaries of roughly the same stature working at the same place and point in time. The time range was the 15th through the 19th centuries. As much as possible, sensory and expressive qualities were held constant within pairs. The variable aspects were compositional elements and subject matter.

3. Modern (M) pairs. Ten pairs contrasted two modern or two contemporary paintings produced during the late 19th through the 20th centuries. Unlike the T pairs, geographic location of the painter pairs was not necessarily the same. The constant and variable art aspects were as in the T pairs except for the four contemporary non-objective pairs (test items M4, M5, M7, and M8; see Appendix A). Because of the difficulty of defining subject matter for non-objective art, the two of each of these pairs were more equivalent than other C, T, and M pairs. The difference between the two of each non-objective pair was primarily compositional.

The classification system just described not only facilitated examining the research questions but also

safeguarded against unduly limiting the period, style, and contrast bases of the visuals. The pairing of only similar visuals or only diverse visuals would have curtailed the basis of stimuli presentation and unnecessarily restricted the generalizability of findings.

Testing Procedure

The ADRT contained 71 items, each a pair of paintings presented through the method of pair-comparison, and required under 40 minutes to administer. Thirty-five items comprised the nucleus of the test. An additional practice item was shown at the start of testing to familiarize Ss with the test task. Each item was shown for 20 seconds, the two of each pair being projected proximately onto a large wall screen.

The test was divided into two series. The first series contained the practice item and the 35 nuclear items. These 35 were randomly ordered disregarding the C, T, and M classes. A random order was established for the left-right positions of the two visuals of each pair. The second series contained the same 35 nuclear items in a different randomized order, with left-right positions again randomized. The order of slide presentation for the two series is given in Appendix B.

For each series, Ss were instructed to choose one of each pair according to a decision-basis and to record their choices on an answer sheet. Three types of decisions -- judgment (J), preference (P), and ambiguous choice (A) -- in

six different pair combinations of type-order were specified. Ss were randomly assigned to the following six levels of type-order, each of which indicated the pair of decisions and the order in which they were made: AJ, AP, JP, PJ, AA, and JJ.⁵

A packet of written instructions was prepared for each level of type-order specifying and explaining the decision-basis on which choices were to be made for each series. When asked to make decisions of aesthetic judgment, Ss were instructed to determine which of each pair was the better work of art in terms of its general aesthetic value. Emphasis was placed on eliciting value judgments of artistic merit. When asked to make decisions of personal preference, Ss were instructed to determine which of each pair was more personally pleasing or satisfying. Emphasis was placed on eliciting choices based on personal appeal and pleasure. For ambiguous choice decisions, the suggested criteria were those of both judgment and preference: artistic quality, pleasure, and satisfaction. So as not to bias decisions, the word "comparative" was used instead of "ambiguous" in packet instructions.

Each packet was stapled and contained the following pages in the order given: a common top page of general directions; two half sheets of instructions specific to the

⁵JA, PA, and PP were not used as levels of type-order; the information they would have provided was not necessary for testing the study's hypotheses.

assigned level of type-order (one for each of the two series); and a folded questionnaire. A standard machine scoring answer sheet was inserted into the packet between Series 1 and Series 2 instruction sheets. The top sheet of general directions is given in Appendix C. The instructions specific to levels of type-order are given in Appendices D through I. It should be noted that the instructions for the two series appear on the same page as given in the Appendices whereas they appeared on separate pages in the test packets. The questionnaire is given in Appendix J.

The packets were arranged in an alternating sequence according to levels of type-order. They were distributed in that order to examinees at the start of testing to enable the random assignment of ss to levels of type-order. Precautions were taken not to discuss the decision instructions and to insure that ss saw the second series instructions only after the first series had been administered.

Prior to testing, an introductory statement that briefly described the study was read to examinees. This statement appears in Appendix K. A detailed report of the sequence of testing is given in Appendix L. Basically, examinees were given time to read instructions and the instructions were reiterated aloud by the experimenter prior to the administration of each slide series. Questionnaires were responded to at the conclusion of testing.

The testing procedure was nearly equivalent for all groups tested. The test was conducted in a room designed for viewing projected visuals. Lighting was controlled for

clarity and brightness of visuals while still enabling Ss to see and mark their answer sheets. Difficulties were encountered in only one testing session where the group was large and a few examinees could not initially see their answer sheets. The condition was corrected by adjusting seating but the standardized procedure was disrupted for the first minutes of testing.

Test Scoring

The testing procedure yielded two choices per item per subject. Correspondence between decisions was quantified by determining the number of items in which the two choices matched. The following data were obtained:

1. A score of correspondence between the two specified decisions. This score reflected the total number of test items in which any one visual of each pair was chosen for both series. The possible range of correspondence scores was 0 - 35. Group mean scores of correspondence were used in analyzing Hypotheses I - V.

2. A preference score based on the C pairs for Ss in the type-order levels of AP, JP, and PJ.⁶ The possible range of preference scores was 0 - 15. High scores indicated a preference for modern art; low scores a preference for traditional art.

⁶Preference scores were not obtainable for the other three type-order levels of AJ, AA, and JJ since preference was not a decision used in these levels.

3. A modern subscore of correspondence based on the M pairs, the possible range being 0 - 10.

4. A traditional subscore of correspondence based on the T pairs, the possible range being 0 - 10.

The mean modern and traditional subscores of correspondence for Ss in the type-order levels of AP, JP, and PJ were used in analyzing Hypothesis VI.

Test Reliability

The issue of reliability was considered especially important because of the recency of test construction. The nature of the test was such that there were no "right" nor "wrong" responses and reliability could not be determined through conventional means. The procedure followed in the pilot study (Mendelson, 1973) was to use the Binomial Probability Model to examine the probability of obtained correspondence scores of given magnitudes for like decisions occurring through chance alone.

In the pilot study, only the type-order levels of JP and PJ were used. Ten of the 35 pairs of visuals were shown twice rather than once in the second series so as to examine the consistency of the second series choices, i.e., the degree to which second series choices were maintained over a short time span. One point was scored for each pair of the ten pairs in which the same visual was chosen for both viewings in the second series. Perfect agreement would have yielded a consistency score of ten; complete lack of agreement

would have yielded a score of zero.

The mean consistency scores for the four groups tested were extremely high: 8.7647, 9.4615, 9.3182, and 9.8333. Use of the Binomial Probability Model showed that the probability of a score of eight or more occurring by chance alone was .04. The probability of a score of nine or more occurring by chance alone was .01. These probability levels indicated high non-chance consistency of decisions and were taken as indicators of short-term test-retest reliability.

Summary

The Art Decision Relationship Test used in this study measured the degrees of relationship among paired decisions about works of art. Test items were 35 pairs of paintings in contrast, traditional, and modern categories in the form of projected 35 millimeter color slide reproductions. The same pairs of visuals were group administered in two consecutive series while examinees chose one of each pair according to a decision basis specified for each series. Three types of decisions -- judgment (J), preference (P), and ambiguous choice (A) -- in six different pair combinations of type-order were specified. These levels of type-order were: AJ, AP, JP, PJ, AA, and JJ.

The standardized testing procedure included a packet of instructions, answer sheet, and post-experimental questionnaire for each examinee. The test yielded a score of correspondence, which reflected the number of pairs in which

the same visual of each pair was chosen for both series, and subscores of correspondence for modern and traditional pairs. Preference scores were also obtained for examinees in the type-order levels of AP, JP, and PJ. Findings of an earlier pilot study (Mendelson, 1973) supported the test's ability to achieve consistent scores in a short-term test-retest situation.

CHAPTER IV: DESIGN OF THE STUDY

In this chapter discussion is presented concerning the college student samples, the design for data collection, the testable hypotheses, and the analysis models used in this study.

Samples

Samples were drawn from populations of art-naive (AN) and art-sophisticated (AS) college students attending classes at Michigan State University during the spring and summer sessions of 1973. Both random sampling and volunteer sampling from these populations were ruled out due to the inaccessibility of students on these bases. Practical considerations led to the use of intact art and non-art classes as the sources of subjects. Non-art classes were drawn from the College of Education where faculty cooperation was forthcoming. AS Ss were drawn from art history and art studio classes given in the Department of Art.

Since some classes did not have their full complement of enrolled students at testing, sample representativeness was somewhat suspect. Concerns that outcome measures might be biased by a volunteer artifact were allayed by Kruglanski's (1973, p. 350) data showing no significant volunteer status main effect on research outcomes. Volunteers apparently do not differ consistently from non-volunteers. It also appeared as if attendance at testing might have been motivated more by student conscientiousness and curiosity

than by any special interest in taking an art test.

While all attending students were tested, standards for sample admission were established against which to assess the art background of examinees. Background data were obtained in postexperimental self-reports (see Appendix J). Any examinee not qualified for inclusion in either sample was eliminated from the study prior to data analysis.

Examinees had to meet the following requirements to be included in the AN sample:

1. Major and cognate fields of study in areas other than art.
2. Limited or no formal training in art.
3. Little or no time spent in art-related activities.

To insure meeting requirements two and three above, examinees were not admitted to the AN sample if they met any three of the following conditions: (1) three or more college art courses taken during term of testing; (2) three or more college art courses taken prior to term of testing; (3) more than one hour a week spent in art-related activities outside of course time; (4) three or more high school art courses taken; and (5) high interest expressed in art.

The AS sample requirements were:

1. Major or cognate field of study in art.
2. Extensive formal training in art.
3. Much time spent in art-related activities.

Five classes containing 105 examinees generally expected to meet AN requirements, and five other classes containing 72 examinees generally expected to meet AS

requirements, were tested. Eleven of the 105 were disqualified for not meeting art-level criteria; and nine others of that group were excluded for procedural reasons, such as incomplete participation. In the AS group, eight were disqualified due to insufficient art background and one for procedural reasons. This brought the AN sample number to 85 and the AS number to 63.

A balanced research design was sought in order to examine the hypotheses of interest with a powerful post hoc procedure.⁷ Therefore, excess numbers per design cell were identified and extra examinees belonging to these cells were deleted at random. Twenty-five AN examinees and three AS examinees were so deleted bringing the number in each sample to 60. A deletion of 25 may seem excessive except that the AN group was intentionally overtested to insure meeting art-level qualifications and retention of the 25 would have resulted in a non-proportional design.

The characteristics of the AN and AS samples are given in Table 1.

It is evident from Table 1 that wide differences in art background and commitment to art separated the two samples. AN Ss were education majors whose art training, art knowledge, and contact with art were limited. They were individuals who did not actively pursue the visual arts. In contrast, AS Ss were characterized by an extensive art

⁷See Chapter IV analysis section.

Table 1

Characteristics of the Art-Naive (AN)
and Art-Sophisticated (AS) Samples

Characteristic	AN Sample (N=60)		AS Sample (N=60)	
	No.	p	No.	p
<u>Educational Level</u>				
Sophomore	1	.017	2	.033
Junior	2	.033	24	.400
Senior	27	.450	25	.417
Graduate	29	.483	9	.150
Other	1	.017	-	.000
<u>Sex</u>				
Female	44	.733	43	.716
Male	16	.267	16	.267
Not Given	-	.000	1	.017
<u>Number of Art Courses Enrolled in During Term of Testing</u>				
0	34	.567	-	.000
1-2	26	.433	22	.367
3 or more	-	.000	38	.633
<u>Number of Art Courses Taken Prior to Term of Testing</u>				
0	28	.467	-	.000
1-2	31	.517	3	.050
3-9	1	.017	10	.167
10 or more	-	.000	47	.783
<u>Hours Per Week Spent in Art Activities Outside of Class</u>				
Less than 1	46	.767	4	.067
1-3	8	.133	7	.117
More than 3	5	.083	49	.816
Not given	1	.017	-	.000
<u>Number of High School Art Courses Taken</u>				
0	42	.700	7	.117
1-2	16	.267	21	.350
3 or more	2	.033	31	.516
Not given	-	.000	1	.017
<u>Level of Art Interest</u>				
Low interest	11	.183	-	.000
Moderate interest	36	.600	4	.067
High interest	13	.217	55	.916
Not given	-	.000	1	.017

background and an active art interest. It was anticipated that any differences in testing responses found to exist between the two samples could reasonably be explained by differences in art training and art knowledge rather than by such differences as educational level or age.

The major areas of study of the AN Ss are reported in Table 2.

As can be seen from Table 2, the subject areas of study heavily represented were elementary education, social sciences education, and special education.

Of the 60 AS Ss, 54 (90%) were art majors and six (10%) were art minors. The emphasis areas of the 54 majors were: art practice (such as sculpture and graphics), $n = 49$, of which 9 were being certified to teach; and art history, $n = 5$.

Design

A two-way fixed-effects analysis of variance (ANOVA) design served as the framework for testing Hypotheses I - V (for statements of hypotheses, see Chapter IV, section on testable hypotheses). The two independent variables incorporated into this design were: art-level and type-order of decisions about works of art. Two art-levels were observed: art-naive and art-sophisticated. Examinees were admitted to these levels on the basis of art background and art training criteria.

The variable type-order specified the two test decisions that were paired and the order in which they were

Table 2
Major Areas of Study of Art-Naive Subjects

Area of Study	Number	p
Elementary Education	23	.383
Social Sciences Education	10	.167
Special Education	8	.133
Mathematics Education	6	.100
Educational Administration	1	.017
Educational Psychology	1	.017
Curriculum	1	.017
No Major Area Given	10	.167

N = 60

rad

Mat

ment

tion

betw

fina

an n

art h

the i

made. The six levels of type-order were:

1. AJ = ambiguous, judgment
2. AP = ambiguous, preference
3. JP = judgment, preference
4. PJ = preference, judgment
5. AA = ambiguous, ambiguous
6. JJ = judgment, judgment

The cells to which Ss were assigned can be seen in the Matrix in Figure 1.

		Type-Order					
		AJ	AP	JP	PJ	AA	JJ
Art-Level	AN						
	AS						

Figure 1

Two-Way Fixed Effects Analysis of Variance Design

The testing procedure allowed for the random assignment of examinees to levels of type-order. This randomization provided good assurances that initial differences between groups within the same art-level did not exist. The final selection of Ss using art-level criteria resulted in an n of ten for each of the 12 cells of the ANOVA design.

Within each art-level Ss were homogeneous in terms of art background and art training. The unit of analysis was the individual subject; the dependent measure was the score

of correspondence. No experimental treatment in the form of training was employed. The treatment to which Ss were assigned consisted of the differential response system established for levels of type-order.

Hypothesis VI (for the hypothesis statement, see Chapter IV, section on testable hypotheses) was examined using a one-way analysis of variance of a difference score for each subgroup of interest. The unit of analysis was the individual subject; the dependent measure was the difference between the modern correspondence score (based on the ten modern test items) and the traditional correspondence score (based on the ten traditional test items).

Four subgroups were identified:

1. ANT: art-naive with preference for traditional art.
2. ANM: art-naive with preference for modern art.
3. AST: art-sophisticated with preference for traditional art.
4. ASM: art-sophisticated with preference for modern art.

Ss in these subgroups were determined by examining preference scores based on the 15 contrast test items. The possible range of preference scores was 0 - 15. High scores indicated preference for modern art; low scores a preference for traditional art. Preference scores were available for only those Ss in the type-order levels of AP, JP, and PJ. Preference scores were thus available for 30 AN and 30 AS

Ss.⁸ Ss in the top and bottom 27% of each of these preference distributions were identified, resulting in the assignment of eight Ss to each of the subgroups: ANT, ANM, AST, and ASM. The preference scores for these subgroups are given in Table 3.

The M and T groups were meaningfully different in their preference biases from one another as evident in the mean preference scores reported in Table 3.

Testable Hypotheses

The first five hypotheses of this study were examined through an analysis of variance (ANOVA) model. The sixth and last hypothesis was examined through ANOVA of a difference score. The hypotheses as stated in Chapter I referred to the relationships among different pairs of decisions used by college students in making selections of works of art. These relationships were operationalized using the independent variable type-order and through scores of correspondence obtained on the Art Decision Relationship Test (ADRT). These operational restrictions are general qualifiers to the following hypotheses.

Hypotheses I - IV The ANOVA null hypothesis that subsumes Hypotheses I - IV is: No effect of type-order on the

⁸It was interesting to note from visual inspection that the AN distribution of preference scores was markedly lower than the AS distribution, the AN subsample of 30 showing, in general, more of a preference for traditional art than did the AS subsample. This difference is reflected in the mean preference scores reported in Table 3 to follow.

Table 3
Frequency Distribution of Preference Scores
For Bipolar Preference Subgroups

Score	Subgroups		Score	Subgroups	
	ANM	ASM		ANT	AST
15	-	1	7	-	-
14	-	1	6	-	5
13	1	-	5	-	-
12	2	3	4	-	2
11	1	3	3	-	1
10	1	-	2	3	-
9	1	-	1	2	-
8	2	-	0	3	-
Mean Score	10.38	12.25		1.00	5.13

ANM and ASM = art-naive and art-sophisticated modern preference subgroups, respectively.

ANT and AST = art-naive and art-sophisticated traditional preference subgroups, respectively.

n = 8

mag

col

Sym

leg

Alt

Rest

Rest

magnitude of correspondence between the art decisions of college students will be found.

Symbolically: $H_0: \beta_{AJ} = \beta_{AP} = \beta_{JP} = \beta_{PJ} = \beta_{AA} = \beta_{JJ} = 0.$

Legend: β refers to type-order (column) effects. The β subscripts refer to levels of type-order.

Alternate Hypothesis I: Art judgment decisions about works of art are not equivalent to personal preference nor to ambiguous choice decisions about works of art.

Restatement of Alternate Hypothesis I: The magnitude of correspondence between paired judgments will exceed that of judgment paired with either preference or ambiguous choice.

Restatement of Alternate Hypothesis I into its components:

Ia: The magnitude of correspondence between paired judgments will exceed that of judgment paired with preference when judgment precedes preference.

Symbolically: $H_{1a}: \beta_{JJ} > \beta_{JP}.$

Ib: The magnitude of correspondence between paired judgments will exceed that of judgment paired with preference when preference precedes judgment.

Symbolically: $H_{1b}: \beta_{JJ} > \beta_{PJ}.$

Ic: The magnitude of correspondence between paired judgments will exceed that of judgment paired with ambiguous choice.

Symbolically: $H_{1c}: \beta_{JJ} > \beta_{AJ}$.

Alternate Hypothesis II: Ambiguous choice decisions about works of art are more similar to personal preference decisions than they are to aesthetic judgment decisions.

Restatement of Alternate Hypothesis II: The magnitude of correspondence between ambiguous choice and personal preference will exceed that of ambiguous choice and judgment.

Symbolically: $H_2: \beta_{AP} > \beta_{AJ}$.

Alternate Hypothesis III: The degree of agreement between judgment and preference is affected by the order in which these decisions are made.

Restatement of Alternate Hypothesis III: The magnitude of correspondence between judgment and preference when judgment precedes preference is not the same as when preference precedes judgment.

Symbolically: $H_3: \beta_{JP} \neq \beta_{PJ}$.

Alternate Hypothesis IV: Aesthetic judgment decisions are held with greater consistency than are ambiguous choice decisions.

Restatement of Alternate Hypothesis IV: The magnitude of correspondence between paired judgments will exceed that of paired ambiguous choices.

Symbolically: $H_4: \beta_{JJ} > \beta_{AA}$.

Hypothesis V Hypothesis V as stated in Chapter I dealt with the effect of college students' art level on the relationships among different pairs of decisions on which selections of works of art were based. The ANOVA null hypothesis for Hypothesis V is: No effect of college students' art level on the magnitude of correspondence between paired decisions will be found.

Symbolically: $H_0: \alpha_{AN} = \alpha_{AS} = 0.$

Legend: α refers to art-level (row) effects. The α subscripts refer to art-naive and art-sophisticated levels.

Alternate Hypothesis V: Art background has an effect on the degree of agreement between decisions about works of art.

Restatement of Alternate Hypothesis V: The magnitude of correspondence between paired decisions of the art-naive population will not equal that of the art-sophisticated population.

Symbolically: $\alpha_{AN} \neq \alpha_{AS}.$

Hypothesis VI Hypothesis VI dealt with the effect of preference bias on the relationships among types of decisions. Preference bias was quantified by the preference score determined from the 15 C test pairs. The dependent measure for the hypothesis was the difference between the modern and the traditional subscores of correspondence. It should be noted that these subscores were derived independently from the preference score.

of

be

Sy-

Le-

Alt

Rest

Symbo

Analy

a two

total

The A

prefer
decisi

sugges
fruitf

The null hypothesis for Hypothesis VI is: No effect of art preference bias on the magnitude of correspondence between paired decisions will be found.

Symbolically: $H_0: u_M - u_T = 0$ for each subgroup.

Legend: u refers to the mean subscore of correspondence.

The u subscripts refer to modern test items (M) and traditional test items (T). The subgroups were: ANT, ANM, AST, and ASM.

Alternate Hypothesis VI: Preference bias has an effect on the degree of agreement between decisions about works of art.

Restatement of Alternate Hypothesis VI: The magnitude of correspondence between unlike decisions⁹ made about art that represents a preference bias will not be equal to the correspondence between unlike decisions made about art opposed to the bias.

Symbolically: $u_M - u_T \neq 0$ for each subgroup.

Analyses

Hypotheses I - V were tested within the framework of a two-way fixed-effects univariate ANOVA model using the total test score of correspondence as the dependent measure.¹⁰ The ANOVA null hypotheses of no main effects and no

⁹The restriction of unlike-decisions occurred because preference scores were unavailable for groups tested in like-decision levels.

¹⁰The earlier pilot study (Mendelson, 1973) strongly suggested that multivariate analysis of subscores would not be fruitful.

interaction effect of the independent variables art-level and type-order were examined through the F test. The alpha level for rejection of the null hypotheses was set at .05.

The assumptions of ANOVA are:

1. Normality: sampling from a normal population distribution of the ability.
2. Homoscedasticity of variance: equality of variance of observations for each group in the design.
3. Independence of errors between and within groups: the probability that an observation has a particular value must not depend on the value of other observations.

The F test is robust to the violations of normality and homogeneity especially with large and equal sample sizes (Boneau, 1971, pp. 312-314, 320). This study's design met both these conditions. The assumption of independence for the units of analysis (individual subjects) was met by standardizing the testing conditions, by the individual nature of the response task, and by the lack of any group training.

The usual ANOVA procedure tests all null hypotheses and examines the results for interaction first. A significant interaction effect particularly if disordinal rules out further consideration of the main effects involved in the interaction. Significant main effects not so involved may be examined using either the Tukey or Scheffé post hoc procedure.

These follow-up procedures to ANOVA identify which group means are contributing to the nonnull main effect by generating confidence intervals for the differences between means. Each interval has a $1-\alpha$ probability of capturing the true value of the difference between population means. The contrasted means are taken as different for any interval not containing zero, the direction of the difference being suggested by the sample means (Glendening, 1973, pp. 1-2).

The Tukey post hoc procedure (T-method) is more powerful than Scheffé for finding differences in pair comparisons. Since Hypotheses I - IV dealt with the differences between group means when taken in pairs, Tukey was chosen over Scheffé. The T-method requirement of homoscedasticity of variance was satisfied by having equal cell size. The alpha level controlled for the entire set of contrasts and was set at .05.

Hypothesis VI concerning the effect of preference bias was examined using a one-way ANOVA of a difference score for each of the four subgroups of interest. Since each within group difference was of interest rather than between group differences this model was more appropriate than a two-factor ANOVA using art-level and preference bias as independent variables.

The difference score (d) for each $\underline{S}$ in each of the subgroups ANT, ANM, AST, and ASM was derived, where $d = X_M - X_T$; X_M = modern subscore of correspondence; and X_T = traditional subscore of correspondence. The mean difference score ($\bar{d}$) was determined for each subgroup. ANOVA tested

the null hypothesis that $\bar{d}$ is equal to zero for each subgroup.

Summary

Ss in this study were students enrolled in education and art courses at Michigan State University in the spring and summer of 1973. Ss were admitted to art-naive and art-sophisticated categories on the basis of art background and art training criteria. AN Ss were characterized by limited or no art training. AS Ss were art majors and minors trained in and actively involved with art.

Ss were randomly assigned to six treatment levels of type-order: AJ, AP, JP, PJ, AA, and JJ. The variable type-order specified the paired test decisions that were assigned in the Art Decision Relationship Test and the order in which decisions were made for the two series of administered test visuals. The test yielded a score of correspondence which reflected the number of pairs in which the same visual of each pair was chosen for both series.

Group mean scores of correspondence were examined through analysis of variance to see if differences due to art-level and type-order between groups existed. The Tukey post hoc procedure was chosen to test for differences in pair comparisons.

Analysis of variance and Tukey procedure were applied to the following hypotheses:

judg

choi

choi

and

same

judg

dec

the

of t

pref

scor

scor

corr

gate

This

corr

repr

resp

to t

I: The magnitude of correspondence between paired judgments will exceed that of judgment paired with

- a. preference when judgment precedes preference.
- b. preference when preference precedes judgment.
- c. ambiguous choice.

II: The magnitude of correspondence between ambiguous choice and personal preference will exceed that of ambiguous choice and judgment.

III: The magnitude of correspondence between judgment and preference when judgment precedes preference is not the same as when preference precedes judgment.

IV: The magnitude of correspondence between paired judgments will exceed that of paired ambiguous choices.

V: The magnitude of correspondence between paired decisions of the art-naive population will not equal that of the art-sophisticated population.

Additionally, subgroups of subjects at both extremes of the bipolar continuum of modern versus traditional art preference were identified. An analysis of a difference score, which examined the difference between the modern subscore of correspondence and the traditional subscore of correspondence, was performed for each subgroup to investigate the effect of preference bias on decision relationships. This analysis examined Hypothesis VI: The magnitude of correspondence between unlike decisions made about art that represents a preference bias will not be equal to the correspondence between unlike decisions made about art opposed to the bias.

CHAPTER V: ANALYSES OF RESULTS

The results of data analysis are report in this chapter. The two-way ANOVA results are presented first and are followed by the analysis of results of the specific hypotheses of this study. The interpretation of results is then given.

Two-Way Analysis of Variance

The group mean correspondence scores of the two-way ANOVA design and their standard deviations are given in Table 4.

The three null hypotheses of the two-way ANOVA design were:

1. No type-order effect:

$$\beta_{AJ} = \beta_{AP} = \beta_{JP} = \beta_{PJ} = \beta_{AA} = \beta_{JJ} = 0.$$

2. No art-level effect:

$$\alpha_{AN} = \alpha_{AS} = 0.$$

3. No interaction effect:

$$\text{all } (\alpha\beta) = 0.$$

The ANOVA table of values is given in Table 5.

Findings based on the F-ratio values reported in Table 5 resulted in:

1. Rejection of the null hypothesis of no type-order effect. A type-order effect, significant at P less than .0001, was found.

2. Failure to reject the null hypothesis of no art-level effect.

Table 4

ANOVA Group Mean Correspondence Scores
and Standard Deviations

Art- Level	Type-Order						Marginal Means
	AJ	AP	JP	PJ	AA	JJ	
AN	$\bar{X}$ 25.60	29.80	27.20	24.10	32.40	31.90	28.50
	s 4.40	3.55	3.43	5.88	2.76	2.13	
AS	$\bar{X}$ 23.80	27.90	27.00	23.30	31.30	31.10	27.40
	s 4.37	3.35	6.57	2.54	2.67	3.96	
Marginal Means	24.70	28.85	27.10	23.70	31.85	31.50	27.95

$\bar{X}$ = Mean; s = Standard Deviation
n = 10; N = 120

Table 5
Two-Way ANOVA Table of Values

Source	<u>df</u>	Mean Squares	<u>F</u>	P Less Than
Type-Order	5	231.8800	14.4141	.001*
Art-Level	1	36.3000	2.2565	.1360
Interaction	5	2.1200	.1318	.9848
Error	108	16.0870		

* = Significant

N = 120

3. Failure to reject the null hypothesis of no interaction effect.

Analyses of Hypotheses I - IV

Since the null hypothesis of no interaction was not rejected, and since a significant main effect was found for the variable type-order, the T-method was used to locate significant pairwise mean differences. The data used in the T-method comparisons are given in Table 6.

The generic form of the T-method null hypothesis is $\Psi = 0$, where Ψ (psi) is any contrast of population means such that the weights sum to zero. Given the data in Table 6, the null hypothesis that $\Psi = 0$ could be rejected for any pairwise type-order comparison in which the interval $\Psi \pm 3.6855$ did not contain zero.

The null hypotheses of the pairwise contrasts of interest and the decisions concerning rejection reached through T-method analysis were:

1. $H_0: \Psi_1 = \beta_{JJ} - \beta_{JP} = 0$ was rejected.
Finding: $\beta_{JJ} > \beta_{JP}$.
2. $H_0: \Psi_2 = \beta_{JJ} - \beta_{PJ} = 0$ was rejected.
Finding: $\beta_{JJ} > \beta_{PJ}$.
3. $H_0: \Psi_3 = \beta_{JJ} - \beta_{AJ} = 0$ was rejected.
Finding: $\beta_{JJ} > \beta_{AJ}$.

Research Hypothesis I was fully supported by the rejection of the three null hypotheses above. The general conclusion was that the magnitude of correspondence between the paired judgments of AN and AS college students as measured by the Art Decision Relationship Test (ADRT) exceeds that of

Table 6
T-Method Table of Values

Source	Mean Square	q Value	$q\sqrt{MS_w/n}$	P Level
Type- Order	16.0870	4.1094	3.6855	.05

N = 20

Numerator df = 6

Denominator df = 108

judgment paired with either preference or ambiguous choice.

4. $H_0: \Psi_4 = \beta_{AP} - \beta_{AJ} = 0$ was rejected.

Finding: $\beta_{AP} > \beta_{AJ}$.

Conclusion: Research Hypothesis II was supported.

The magnitude of correspondence between ambiguous choice and preference as measured in AN and AS college students by the ADRT exceeds the magnitude of correspondence between ambiguous choice and judgment.

5. $H_0: \Psi_5 = \beta_{JP} - \beta_{PJ} = 0$ was not rejected.

Finding: No statistical difference was found between β_{JP} and β_{PJ} .

Conclusion: Research Hypothesis III was not supported.

The magnitude of correspondence between judgment and preference as measured in AN and AS college students by the ADRT when judgment precedes preference is statistically the same as when preference precedes judgment.

6. $H_0: \Psi_6 = \beta_{JJ} - \beta_{AA} = 0$ was not rejected.

Finding: No statistical difference was found between β_{JJ} and β_{AA} .

Conclusion: Research Hypothesis IV was not supported.

The magnitude of correspondence between the paired judgments of AN and AS college students as measured by the ADRT is statistically the same as that of paired ambiguous choices.

Analysis of Hypothesis V

Hypothesis V was examined directly through the ANOVA analysis reported earlier in this chapter.

Finding: $H_0: \alpha_{AN} - \alpha_{AS} = 0$ was not rejected. No statistical difference was found between α_{AN} and α_{AS} .

Conclusion: Research Hypothesis V was not supported. The magnitude of correspondence between paired decisions as measured by the ADRT is statistically the same in AN and AS college student populations.

Analysis of Hypothesis VI

To examine the effect of preference bias, a one-way ANOVA of a difference score ($\bar{d}$) at the .01 alpha level was performed for each of the subgroups ANT, ANM, AST, and ASM. This analysis tested the null hypothesis that the magnitude of correspondence between unlike decisions made about art that represents a preference bias will equal that of art opposed to the bias. The mean scores and F -ratio values for the tests of this hypothesis are given in Table 7.

As can be seen in Table 7, the difference score was not significantly different from zero for any of the subgroups.

Finding: $H_0: u_M - u_T = 0$ for each subgroup was not rejected; the difference score was not significantly different from zero for any of the subgroups: ANT, ANM, AST, and ASM.

Conclusion: Research Hypothesis VI was not supported. The magnitude of correspondence between unlike decisions made about art that represents a preference bias as measured by the ADRT is the same as that about art opposed to the bias in AN and AS college students at the extremes of the bipolar preference continuum.

Table 7

Mean Scores and $\bar{F}$ -Ratio Values for Tests of Difference
Scores for Bipolar Preference Subgroups

Sub-Group	$\bar{X}_M$	$\bar{X}_T$	$\bar{d} = \bar{X}_M - \bar{X}_T$	$\bar{F}$ Value	P Less Than
ANT	7.500	7.250	.250	.4667	.5165
ANM	7.875	7.375	.500	1.1667	.3159
AST	8.625	7.500	1.125	3.7550	.0939
ASM	7.000	7.875	-.875	1.8743	.2133

$\bar{X}_M$ and $\bar{X}_T$ = mean modern and traditional subscores, respectively.
numerator $df = 1$; denominator $df = 7$

Interpretation of Results

This study focused on the differences between relationships among types of decisions used as the basis for making selections of works of art. The major issue that guided this study and that was reflected in Hypothesis I was whether aesthetic judgment is equivalent to either of the surrogate decisions personal preference and ambiguous choice. The unequivalence of judgment to either preference or ambiguous choice that had been predicted was confirmed by the results of this study. In this as in other conclusions, the restriction of the hypotheses to the defined populations and measure must be recognized.

The issue that was reflected in Hypothesis II was whether ambiguous choice is more similar to personal preference than to aesthetic judgment. The prediction of greater similarity to personal preference was supported statistically.

Hypothesis III examined the degree of relationship between judgment and preference as a function of the order in which these decisions were made. The prediction that the reactive effect of judgment on preference would be different from the reactive effect of preference on judgment resulted from an earlier finding (Mendelson, 1973) of a significant interaction between order of occurrence and art-level as summarized in Figure 2.

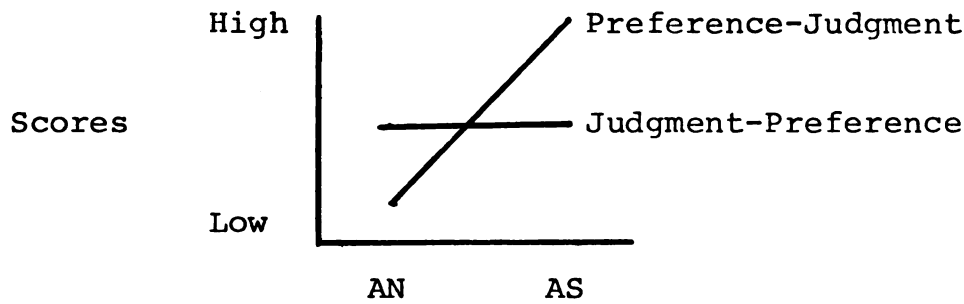


Figure 2

Pilot Study Interaction of Art-Level and Order of Judgment and Preference

An inadequacy of the cited study was that initial equivalence between groups within art levels had not been assured. The findings of the current study did not substantiate the interaction nor find any differential reactive effect.

Hypothesis IV was the last to examine the effect of type-order and stated that judgment decisions are held more consistently than ambiguous choice decisions. This prediction was not supported, the conclusion being that aesthetic judgment and ambiguous choice are equally consistent over a short time period.

No support was found for Hypothesis V which stated that art background has an effect on the degree of agreement between decisions about works of art. It thus appears as if the relationships between the specified art decisions are the same in art-naive and art-sophisticated college student populations. Judgment and preference do not appear to be more equivalent for art-naive than for art-sophisticated college students.

Hypothesis VI which stated that an extreme position on the bipolar continuum of modern versus traditional art preference affects the degree of agreement between decisions about works of art was not supported. Thus the degree of agreement between preference and judgment, and between preference and ambiguous choice, appears not to be affected by the work of art's consonance with the individual's preference bias as defined in terms of modern versus traditional art.

As reflected by the relatively high correspondence scores for all groups, it seems as if judgment, preference, and ambiguous choice are similar but not synonymous kinds of decisions. A chance level correspondence score would have been 17.50. The lowest observed mean score was 23.30. However, due to the proximity in time of the two sets of decisions, a carry-over memory effect of prior decisions might have spuriously inflated correspondence scores. Since this uncontrolled for variable had an equal chance of affecting all levels of type-order it was not seen as a threat to the internal validity of the design.

Although the art-naive and art-sophisticated populations were restricted to college students, it is reasonable to conclude from the findings that tests of aesthetic judgment that suggest the criteria of ambiguous choice to examinees are not accurately measuring aesthetic judgment and that their validity could be increased by specifying the criteria of criticism and evaluation.

Chi-Square Analysis of Questionnaire Responses

Findings parallel to the type-order effect and lack of art-level effect occurred in how Ss perceived their test choices and the test instructions as reported in questionnaires subsequent to testing. The questions and possible responses were:

1. Were the choices you made in the 1st series different from the choices made in the 2nd series?
a) No different b) Somewhat different c) Clearly different
2. Were the decision-bases specified for the two series different from one another?
a) No different b) Somewhat different c) Clearly different

The frequency distributions for responses to questions 1 and 2 are given in Table 8. The frequencies for each question appear first for art-level categories and then for type-order categories.

Chi-square tests of independence between variables were performed for each of the four frequency distributions reported in Table 8. The chi-square null hypothesis is that variables are independent of one another. The data for these chi-square tests are reported in Table 9.

The conclusions reached through chi-square analysis were that art-level was independent of the responses to both Questions 1 and 2, whereas type-order and responses to both questions were dependent.

Summary

Hypotheses I - VI were investigated through two-way ANOVA, the Tukey post hoc procedure, and analysis of a difference score. The two-way ANOVA of the effects of

Table 8
Frequencies of Responses to Choice and Instruction
Questions of Questionnaire

Art- Level	Choice Question (1)			Instructions Question (2)				
	Responses			Art- Level	Responses			Mar- ginal
	a	b	c		a	b	c	
AN	19	40	1	AN	22	29	9	60
AS	14	46	0	AS	21	31	8	60
Mar- ginal	33	86	1	Mar- ginal	43	60	17	120
Type- Order	Responses			Type- Order	Responses			Mar- ginal
	a	b	c		a	b	c	
AJ	2	18	0	AJ	1	16	3	20
AP	3	17	0	AP	4	14	2	20
JP	2	18	0	JP	1	14	5	20
PJ	4	15	1	PJ	3	11	6	20
AA	9	11	0	AA	16	4	0	20
JJ	13	7	0	JJ	18	1	1	20
Mar- ginal	33	86	1	Mar- ginal	43	60	17	120

Table 9
Chi-Square Tests of Independence Data

Variables	<u>df</u>	Chi-Square	Decision
Art-Level and Responses to Question 1	2	2.18	Fail to Reject Null
Type-Order and Responses to Question 1	10	30.35*	Reject Null
Art-Level and Responses to Question 2	2	.96	Fail to Reject Null
Type-Order and Responses to Question 2	10	69.76*	Reject Null

* = Significant at P less than .001

type-order and art-level disclosed no significant interaction effect and no significant art-level main effect. A type-order main effect significant at the .0001 level was identified. The findings and conclusions of all analyses in relation to the hypotheses are reported in Table 10.

Within the limitations established for this study, the conclusions reached were that:

1. Aesthetic judgment is not equivalent to personal preference nor to ambiguous choice.
2. Ambiguous choice is more similar to personal preference than it is to aesthetic judgment.
3. The order in which judgment and preference decisions are made does not affect the magnitude of correspondence between them.
4. Aesthetic judgment and ambiguous choice are held with equal consistency.
5. Art-level has no effect on the magnitude of correspondence between decisions.
6. Preference bias has no effect on the magnitude of correspondence between judgment and preference nor between ambiguous choice and preference.

Table 10
Summary of Findings and Conclusions

Hypothesis	Null Hypothesis	Analysis	Decision Re Null Rejection	Conclusion
Ia	$B_{JJ} - B_{JP} = 0$	Tukey	Reject at .05	$B_{JJ} > B_{JP}$
Ib	$B_{JJ} - B_{PJ} = 0$	Tukey	Reject at .05	$B_{JJ} > B_{PJ}$
Ic	$B_{JJ} - B_{AJ} = 0$	Tukey	Reject at .05	$B_{JJ} > B_{AJ}$
II	$B_{AP} - B_{AJ} = 0$	Tukey	Reject at .05	$B_{AP} > B_{AJ}$
III	$B_{JP} - B_{PJ} = 0$	Tukey	Fail to reject	$B_{JP} = B_{PJ}$
IV	$B_{JJ} - B_{AA} = 0$	Tukey	Fail to reject	$B_{JJ} = B_{AA}$
V	$\alpha_{AN} = \alpha_{AS} = 0$	ANOVA	Fail to reject	$\alpha_{AN} = \alpha_{AS} = 0$
VI	$u_M - u_T = 0$ for each group: ANT, ANM, AST, & ASM	ANOVA of a difference score	Fail to reject	$u_M - u_T = 0$ for each group

CHAPTER VI: SUMMARY AND CONCLUSIONS

In this chapter a synopsis of the study is presented and followed by a discussion of the results of investigation. The chapter concludes with the implications for research and for instruction in art education.

Summary

The ability to perceive visual artistic quality is measured through tests of aesthetic judgment using sets of visuals and a comparative choice procedure. An examinee's level of aesthetic judgment is determined by the degree of correspondence between his choices and those established through art expert consensus.

The way in which aesthetic judgment is operationalized in standardized tests of aesthetic judgment appeared ambiguous in that the suggested decision criteria are artistic quality, pleasure, and satisfaction. The equivalence of aesthetic judgment and personal preference for art appeared to be an underlying assumption of this decision basis. The distinction between aesthetic judgment and personal preference posited by aestheticians was seen as a challenge to this assumption and therefore to the construct validity of these tests.

This study was designed primarily to investigate the congruence between empirically obtained data and the philosophical premise of the unequivalence of aesthetic judgment and personal preference for art. The method used was to

compare the relationships among types of decisions that governed selections of one of each pair of a set of paired paintings. Three such decisions were examined: aesthetic judgment (J) defined as a critical evaluation of a work of art and based on the criteria of artistic quality and aesthetic value; personal preference (P) based on the criteria of appeal, pleasure, and satisfaction; and ambiguous choice (A) based on the criteria of the two aforementioned decision types and representing the decision specified in standardized tests of aesthetic judgment.

The main purpose of this study was to compare the magnitudes of correspondence between these decisions when taken in the six type-order combinations of AJ, AP, JP, PJ, AA, and JJ to see if differences between combinations existed. Other purposes that guided this study were to examine the effects of art background and preference bias on the magnitudes of correspondence between paired decisions. Art background was represented by the levels of art-naivety and art-sophistication. Preference bias was defined by the polarity of modern versus traditional art preference.

The research hypotheses of this study were:

Hypothesis I: Aesthetic judgment decisions about works of art are not equivalent to personal preference nor to ambiguous choice decisions about works of art.

Hypothesis II: Ambiguous choice decisions about works of art are more similar to personal preference decisions than they are to aesthetic judgment decisions.

Hypothesis III: The degree of agreement between judgment and preference is affected by the order in which these decisions are made.

Hypothesis IV: Aesthetic judgment decisions are held with greater consistency than are ambiguous choice decisions.

Hypothesis V: Art background has an effect on the degree of agreement between decisions about works of art.

Hypothesis VI: Preference bias has an effect on the degree of agreement between decisions about works of art.

The Art Decision Relationship Test (ADRT) constructed for the preceding pilot study (Mendelson, 1973) was modified for use in this study. The test is composed of 35 items, each a pair of projected 35 millimeter color slide reproductions of paintings in modern (M), traditional (T), and contrast (C) categories. The test task is that of comparative choice. Two randomized series of the same 35 pairs are administered with a type of decision specified for each series. Selections of one of each pair made in the first series are compared with selections made in the second series, yielding a total test correspondence score and painting category subscores that reflect the extent of agreement between pairs of decisions.

The ADRT was group administered to 177 students enrolled in education and art classes at Michigan State University in the spring and summer of 1973. Examinees were randomly assigned to the six testing levels of type-order but were admitted to the art-naive (AN) or art-sophisticated (AS)

sample only if their art background met established criteria. In final form, the AN and AS samples each contained 60 Ss evenly distributed in the six type-order levels. AN Ss were characterized by limited or no art training whereas AS Ss were art majors or art minors trained in and actively involved in the visual arts.

A two-way fixed-effects univariate analysis of variance (ANOVA) balanced design was used to compare decision relationships. Two-way ANOVA was thus applied to Hypotheses I - V by testing the main and interaction effects of art-level and type-order. The unit of analysis was the individual subject; the dependent measure was the score of correspondence.

A one-way ANOVA of a difference score for each preference bias subgroup was used to test Hypothesis VI. The four preference subgroups were ANT, ANM, AST, and ASM, where T and M refer to preference for traditional art and preference for modern art, respectively. These subgroups each contained eight Ss tested in the type-order levels of AP, JP, and PJ whose preference scores placed them at an extreme on the bipolar preference continuum. The ANOVA dependent measure was the difference between the modern subscore and the traditional subscore ($u_M - u_T$).

Conclusions

F test analysis disclosed no significant interaction effect and no significant art-level main effect (α). A

type-order main effect (β) significant at the .0001 level was identified. The Tukey post hoc procedure at the .05 level was used to examine type-order comparisons of interest to identify which type-order effects were contributing to the nonnull condition. The effect of preference bias was examined through ANOVA of a difference score at the .01 level for each of the subgroups ANT, ANM, AST, and ASM.

Within the qualifications imposed by the established populations and measure, the results of these analyses were:

1. Aesthetic judgment is not equivalent to personal preference nor to ambiguous choice: $\beta_{JJ} > \beta_{JP}$; $\beta_{JJ} > \beta_{PJ}$; and $\beta_{JJ} > \beta_{AJ}$. Research Hypothesis I was supported.

2. Ambiguous choice is more similar to personal preference than it is to aesthetic judgment: $\beta_{AP} > \beta_{AJ}$. Research Hypothesis II was supported.

3. The order in which judgment and preference decisions are made does not affect the magnitude of correspondence between them: $\beta_{JP} = \beta_{PJ}$. Research Hypothesis III was not supported.

4. Aesthetic judgment and ambiguous choice are held with equal consistency: $\beta_{JJ} = \beta_{AA}$. Research Hypothesis IV was not supported.

5. Art-level has no effect on the magnitude of correspondence between decisions: $\alpha_{AN} = \alpha_{AS} = 0$. Research Hypothesis V was not supported.

6. Preference bias has no effect on the magnitude of correspondence between judgment and preference nor between

ambiguous choice and preference: $u_M - u_T = 0$ for each subgroup. Research Hypothesis VI was not supported.

Discussion

The results of this study pertain to type-order, art-level, and preference bias and are discussed in that order in this section.

Type-Order The conclusion that aesthetic judgment is not equivalent to personal preference nor to ambiguous choice calls into question the validity of tests which operationalize aesthetic judgment through the use of the surrogate decision ambiguous choice. Given the same subject populations, there is little reason to believe that substituting a test format different from pair comparisons of projected visuals would significantly alter the obtained differences in relationships. This dimension of test format was not seen as a threat to the external validity of the study.

However, generalizability might be limited by having restricted works of art to paintings, and further to paintings of established merit. Decision relationships might not be the same were art forms other than painting used as test items, or were items to display a wide range of aesthetic excellence from the very mundane to "masterpiece" quality. The effects of diversity of form and aesthetic quality among visuals have yet to be examined.

The conclusion that ambiguous choice is more similar to personal preference than to aesthetic judgment indicated

that ambiguous choice is more a substitute for preference than for judgment and further disconfirmed the validity of aesthetic judgment tests. The purpose of these tests being to measure a perceptual capability thought to be innately determined, it seems strongly indicated that in suggesting the criteria of personal appeal these tests are encouraging variables different from aesthetic judgment to influence test scores and thus are introducing unnecessary error variance.

By encouraging decision bases other than critical objective ones to influence scores, it appears as if tests are not measuring the capability of aesthetic judgment as accurately as they might. If the interest is in measuring the ability to recognize such qualities as aesthetic organization, unity, and harmony, examinees should be made aware of the critical nature of the task so that they may apply the appropriate criteria.

Art-Level The absence of a significant art-level main effect cannot be attributed to lack of design precision. The statistical test for type-order found a significant effect despite having a smaller sample size and thus less power for finding differences than had the test for art-level effect. The lack of difference in decision relationships between art-levels must be attributed to a genuine similarity between art-level groups as measured by the ADRT. Chi-square tests of how subjects perceived their choices and the task further confirmed the equivalence of decision

relationships of the AN and AS samples. This lack of art-level difference might not be substantiated given test items representing a wide range of aesthetic excellence.

As to generalizability, the findings of this study are not applicable to college students in art-levels other than those examined because the variable art-level was a fixed factor. How much the restriction of the art-naive population to education majors limits generalizability is speculative.

Preference Bias Several weaknesses inherent in the procedure to examine preference bias effect might have accounted for the lack of effect found. Each group being composed of only eight Ss was relatively small. Each of the two compared scores was based on only ten items and might not have been a reliable enough measure. The general lack of reliability of a difference score (see Mehrens & Lehmann, 1973, pp. 117-118) further jeopardized the dependability of the difference measure. For these reasons the conclusion of no preference bias effect on the relationships between decisions must be held cautiously.

Implications for Future Research

Several tangential research directions that seem promising in light of the knowledge gained through this investigation are presented below.

The ADRT was constructed primarily to examine the issues and appeared satisfactory for this purpose. Its

continued use would allow for direct comparisons with other populations of interest and is recommended. Cross-validation of the study is strongly recommended using samples that extend the age, art background, and educational levels of subjects. Non-art college students, if used, should be drawn from fields of study other than, or in addition to, education.

Other existing measures could be modified for use in addressing the same issues and would lend validity to this study's findings were results replicated. Were investigations using different measures to reach anomalous conclusions, the design of this study would need to be reexamined for weaknesses as yet unidentified. For example, a test in which items varied from very poor to outstanding examples of art might more readily reveal any art-level influence not disclosed by the ADRT.

Research following the lines of this study would be justified in eliminating either the JP or PJ level of type-order. However, correspondence between judgment and preference was somewhat greater for both AN and AS groups when judgment preceded preference. This difference was not statistically significant but if replicated might identify a directional trend.

The decisions within each type-order level were made contiguously and the effect of time on decision relationships is unknown. It would be valuable to examine the stability of relationships longitudinally, at the least in

terms of intervals of several months. This interest could be combined with an investigation of the effects of aesthetics training on aesthetic judgment at any educational level. Whether a pair comparison or rank order test procedure is followed, examining training effects would require that test items be prejudged for aesthetic quality prior to test administration.

A more specific measure of the relationships between decisions than a scale unit dependent score of correspondence could be achieved through correlation. One approach would require that Ss choose one of each pair of a repeated series of paired paintings in M, T, and C categories. Pairs would differ from those employed in this study in that each pair would have a poor and a good example of art as prejudged using the interval sorting technique often used in constructing attitude scales (see Edwards, 1957, pp. 83-98). The two decision scores would each reflect agreement with expert consensus and would allow for the computation of correlations between any pair of decisions. A correlation coefficient of itself would be a useful indicator of the magnitude of decision correspondence.

Additionally, continued research into the effects of preference bias with a more precise design and more reliable contributing measures than those of this study is clearly indicated.

Implications for Instruction in Art Education

Wilson (1971, pp. 512-517) outlined the cognitive and affective objectives common to art education programs from the intermediate elementary grades through high school. These objectives result from the union of content areas and student behaviors. The content areas are: media, tools, and forming processes; visual structure; subject matter; art form; cultural context; and art theory and criticism. The student behaviors are: perception, comprehension, analysis, evaluation, appreciation, and production.

The development of the critical ability of aesthetic judgment may be seen as one important culmination of these educational objectives. Wilson (1971, p. 520) wrote:

Evaluation in art education is primarily concerned with making reasoned critical judgments about the aesthetic quality and values of works of art. The making of such judgments is generally considered to depend upon an intimate knowledge of the field of art, sensitive perception, analysis, and the application of reliable aesthetic criteria. Although it is possible to apply reliable aesthetic criteria without being conscious of those being used, many would hold that there are distinct advantages to having a knowledge of the range of criteria which might be used in making a critical judgment, since knowing alternative criteria frees one to apply various standards rationally as they apply to the aesthetic situation.

It is logical to assume that the critical objective criteria of which Wilson spoke are applied in the aesthetic evaluations made by individuals representative of the art-sophisticated population to which this study generalized. By contrast, art-naive individuals' appraisal of art can only be seen as less conscious and less critical.

That there was no significant difference between art-naive and art-sophisticated groups in the art decision relationships measured in this study was surprising in view of the above assumptions and those stated earlier. Counter to expectation, aesthetic judgment and personal preference did not appear to be any more differentiated for art-sophisticated individuals than for art-naive individuals. Art-sophisticated individuals seemed no more able to separate their subjective reactions from their objective evaluations than could art-naive individuals. One may interpret this finding by concluding either that the two groups make the distinction equally as well or equally as poorly, depending upon one's perspective.

From the latter viewpoint, it may be that differences among modes of reaction to works of art have not been a focus of art programs at any educational level and therefore students with extensive art training are not any better equipped to sort emotional from cognitive reactions. An indication that this might indeed be the case was suggested by Feldman's (1967, p. 478) faulting of art instruction for emphasizing interpretation and explication and too rarely stressing the expression of value judgments.

One would hope that art education programs are capable of developing aesthetic judgment and further that they might clarify for all students, both those who will in adulthood interact with art peripherally as consumers and those for whom art's contribution to life will be substantial,

distinctions between objective and subjective value. Smith (1968, p. 14) posited that one educated in aesthetics must go beyond personal pleasure in reacting to art to form judgments and supportive reasons for those judgments.

Perhaps it is not so much that the goals of increased competency in formulating aesthetic judgments and recognizing the bases of one's own preference mode have been approached through inadequate instructional practices in the past, but that the goals have not really been functional ones. With more attention focusing on the behavioral goals of art education, this situation may be in the process of being corrected.

Other than instructional emphasis on the full range of objective procedures and subjective criteria, as presented in Chapter I, supplemental means to these goals were provided by Sibley. Sibley (1966, pp. 336-338) suggested that instruction in aesthetics be approached through the use of the following assessment techniques:

1. Attending to non-aesthetic features.
2. Attending to aesthetic features and qualities.
3. Pointing out the interrelatedness of aesthetic and non-aesthetic features.
4. Using metaphors and similies to describe qualities.
5. Contrasting and comparing works of art.
6. Reiterating analysis.
7. Using nonverbal behavior to supplement description and analysis.

Smith (1971, p. 483) pointed out that the use of these kinds of procedures constitutes a set of heuristic devises with which to approach art in the hopes of gaining

some meaningful understanding. The first three techniques in particular could be effective in elucidating the differences between objective and subjective reactions, the component parts of each of these processes, and their sets of underlying standards.

The development of critical competency and conscious awareness of reaction components is within the realm of realization. In a discussion about testing for evaluative competencies, Wilson (1971, p. 545) presented sample items each of which required a different criterion for evaluation, such as the work's integration, its transmission of cultural aspects, and its ability to elicit an intense experience. Wilson has found that even elementary grade students can learn to apply different criteria when making evaluations. It follows that the learning of other distinctions among reactions is also possible.

APPENDICES

Cont

Pair

C1

C2

C3

C4

C5

C6

C7

C8

C9

C10

C11

C12

C13

C14

C15

C16

APPENDIX A

Test Item Pairs of Paintings

Contrast Pairs

Pair		Artist	Nation- ality	Century	Title of Painting
C1	a	Gainsborough	English	18	The Morning Walk: W. Hallett and Wife
	b	Beckmann	German	20	Double Portrait
C2	a	Murillo	Spanish	17	Madonna and Child
	b	Picasso	Spanish	20	Maternity
C3	a	Valazquez	Spanish	17	Toilet of Venus
	b	Picasso	Spanish	20	Girl Before a Mirror
C4	a	VanEyck	Flemish	15	The Marriage of Arnolfini
	b	Chagall	Russian	20	Portrait with Wineglass
C5	a	Chardin	French	18	Basket of Peaches
	b	Braque	French	20	Still Life with Fruit
C6	a	Turner	English	19	Junction of Thames and Medway
	b	Derain	French	20	The Pool of London
C7	a	ElGreco	Spanish	16	View of Toledo
	b	VanGogh	Dutch	19	The Starry Night
C8	a	Lorrain	French	17	Landscape: Marriage of Isaac and Rebecca
	b	Sloan	American	20	South Beach Bathers
C9	a	daMessina	Italian	15	Portrait of a Young Man
	b	Gaugin	French	19	Self Portrait
C10	a	Courbet	French	18	Apples and Pomegranates
	b	Wesselman	American	20	Still Life #36
C11	a	Goya	Spanish	18	Portrait of Don Manuel
	b	Evergood	American	20	Lily and the Sparrows
C12	a	Titian	Italian	16	Venus of Urbino
	b	Modigliani	Italian	20	Reclining Nude (Le grand nu)
C13	a	Vermeer	Dutch	17	Young Woman with Water Jug
	b	Braque	French	20	Woman with a Mandolin
C14	a	Correggio	Italian	16	Leda and the Swan
	b	Matisse	French	20	Landscape with Figures
C15	a	Constable	English	19	Dell at Helmingham
	b	Vlaminck	French	20	Landscape with Red Trees
C16	a	Ingre	French	19	Portrait of Mme de Senonnes
	b	DeKooning	American	20	Seated Woman

APPENDIX A
(Continued)

Pair		Artist	Nation- ality	Century	Title of Painting
M1	a	Bonnard	French	20	Nude Before Mirror
	b	Vuillard	French	20	The Sunny Room
M2	a	Kirchner	German	20	The Street
	b	Nolde	German	20	The Large Gardner
M3	a	Dali	Spanish	20	Mae West
	b	deChirico	Italian	20	Philosopher's Conquest
M4	a	Carra	Italian	20	Rhythm of Objects
	b	Severini	Italian	20	Blue Dancer
M5	a	Gorky	American	20	Orators
	b	Miro	Spanish	20	Figure
M6	a	Delaunay	French	20	Eiffel Tower
	b	Marin	American	20	Movement: Boats and Objects
M7	a	Motherwell	American	20	The Voyage
	b	Dove	American	20	Cows in Pasture
M8	a	Rothko	American	20	#24
	b	Baziotes	American	20	Pompei
M9	a	O'Keeffe	American	20	Cow's Skull with Calico Rose
	b	Sheeler	American	20	Barn
M10	a	Sutherland	English	20	Variation on a Theme II
	b	Bacon	English	20	Fragment of a Crucifixion

Traditional Pairs

T1	a	Bellini	Italian	15	Feast of the Gods
	b	Signorelli	Italian	15	Eunostos of Tanagra
T2	a	Raphael	Italian	16	St. George and the Dragon
	b	Tintoretto	Italian	16	Susanna and Elders
T3	a	Bosch	Flemish	15	Arrival of the Magi
	b	Weyden, VanDer	Flemish	15	Man with Turban
T4	a	Boucher	French	18	Four Seasons: Winter
	b	Fragonard	French	18	The Music Lesson
T5	a	Hogarth	English	18	Heads of Six Servants
	b	Reynolds	English	18	Lady Elizabeth Delme and Children
T6	a	Cranach (Elder)	German	16	The Stag Hunt
	b	Holbein (Younger)	German	16	The Ambassadors
T7	a	Brueghel, P.	Flemish	16	Return of the Herd
	b	Brueghel, J.	Flemish	16	The Village Street
T8	a	Hals	Dutch	17	Laughing Cavalier
	b	Rembrandt	Dutch	17	Artist in His Studio
T9	a	Guardi	Italian	18	View of the Rialto
	b	Tiepolo	Italian	18	Fresco at Villa Valmarana
T10	a	Gericault	French	19	The White Horse
	b	Corot	French	19	Old Beech Tree

Pair

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36

APPENDIX B

ORDER OF SLIDE PRESENTATION FOR SERIES ONE AND TWO

Series 1				Series 2			
Pair #	Pair	Position Left Right		Pair #	Pair	Position Left Right	
1	C 6	a	b	1	C14	b	a
2	T 6	a	b	2	T 5	b	a
3	M 1	b	a	3	T 8	b	a
4	C 9	a	b	4	M 4	a	b
5	M10	b	a	5	T 3	b	a
6	C 5	b	a	6	M 2	b	a
7	M 2	a	b	7	C15	b	a
8	T 4	b	a	8	T10	a	b
9	C16	b	a	9	C12	b	a
10	T 8	a	b	10	M 7	b	a
11	M 4	a	b	11	C 8	b	a
12	T 2	b	a	12	M10	a	b
13	C11	b	a	13	T 6	b	a
14	T 7	a	b	14	M 9	a	b
15	M 5	a	b	15	C10	a	b
16	T 3	a	b	16	M 6	a	b
17	C14	b	a	17	C13	b	a
18	T10	a	b	18	C16	b	a
19	T 5	b	a	19	T 1	b	a
20	C 3	b	a	20	C 5	a	b
21	C 8	a	b	21	M 3	a	b
22	T 1	b	a	22	C11	a	b
23	C12	b	a	23	C 7	a	b
24	M 8	b	a	24	T 4	b	a
25	M 3	b	a	25	T 9	a	b
26	C 4	a	b	26	T 2	a	b
27	C 7	a	b	27	C 4	b	a
28	C15	b	a	28	T 7	a	b
29	M 6	b	a	29	C 3	a	b
30	M 7	b	a	30	M 5	b	a
31	C10	a	b	31	C 2	b	a
32	C 1	b	a	32	C 1	a	b
33	T 9	b	a	33	M 8	b	a
34	M 9	a	b	34	C 9	a	b
35	C 2	a	b	35	M 1	a	b
36	C13	a	b				

PLE

sho
a p

cho
fie
its
wit
ber
row

|

Disr

choi
choi

Do no
that

is ac

APPENDIX C

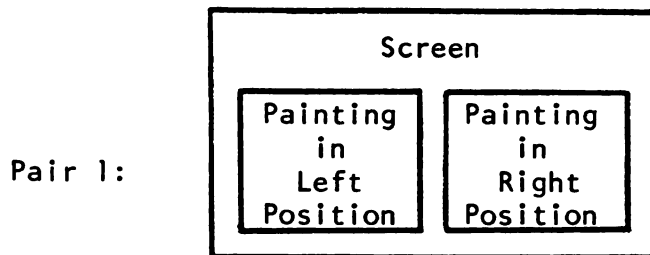
TEST PACKET TOP PAGE: DIRECTIONS FOR ALL EXAMINEES

PLEASE READ ONLY THIS PAGE. DO NOT TURN TO OTHER PAGES UNTIL INSTRUCTED TO DO SO.

You will see two series of paintings. These paintings will be shown in pairs. The two of each pair will be shown simultaneously for a period of 20 seconds per pair.

Your task is to compare the two paintings of each pair and to choose one according to a certain type of decision that will be specified. The pairs will be numbered consecutively. As each pair is shown, its number will be called. As you make each choice, mark your choice with the pencil provided on the answer sheet in the space for that number. The numbers on the answer sheet follow across the answer sheet in rows.

This is the way to mark your choices:



Answer Sheet		
TF	TF	
1 ☐ ABCDE	2 ABCDE	3...
If you choose the painting in the left position blacken space A		

Answer Sheet		
TF	TF	
1 ☐ ABCDE	2 ABCDE	3...
If you choose the painting in the right position blacken space B		

Disregard spaces C, D, and E.

It is expected that the choices you make will differ from the choices made by others in the room. There are no predetermined "correct" choices. Please make the choices that you think are right.

Make a choice for each pair. Blacken the appropriate space heavily. Do not omit any pair. Make your choices according to the type of decision that will be specified for each series.

The first series you will see contains 36 pairs of paintings. Pair 1 is actually a practice pair to familiarize you with the task.

Type

ing
pare
to y
are

ings

PLEA

Inst

ings
and
posi

are
assie
qual

ing c
ings
judgm
are a

ings
pair

PLEASE

APPENDIX D

LEVEL 1 INSTRUCTIONS: AMBIGUOUS CHOICE - JUDGMENT

Type of Decision for This First Series: Comparative Choice

For this series, take the role of an observer deciding which painting of each pair shown is more pleasing, artistic, or satisfying. Compare the two paintings of each pair and choose the one that, according to your judgment or feelings, is better. Use the criteria that you think are appropriate for making your choices.

You will need to make a choice even if you feel that the two paintings are very close in value.

PLEASE DO NOT TURN TO OTHER PAGES UNTIL INSTRUCTED TO DO SO.

Instructions for the Second Series:

This second series you will see contains the same 35 pairs of paintings, but in a different order. You will again choose one of each pair and mark your choice on the answer sheet according to its left or right position on the screen.

The choices you will be making about the pairs in this series, however, are to be based on a more specific type of decision. The decision-basis assigned to this series is: critical, objective judgment of artistic quality.

For this series, take the role of an art critic deciding which painting of each pair shown is the better work of art. Compare the two paintings of each pair and choose the one that, according to your objective judgment, is of greater artistic merit. Use the standards that you think are appropriate as the criteria for making your critical evaluations.

You will need to make a choice even if you think that the two paintings are very close in aesthetic value. Choose the one painting of each pair that, in your judgment, is better.

PLEASE DO NOT TURN TO OTHER PAGES UNTIL INSTRUCTED TO DO SO.

APPENDIX H

LEVEL 5 INSTRUCTIONS: AMBIGUOUS CHOICE-AMBIGUOUS CHOICE

Type of Decision for This First Series: Comparative Choice

For this series, take the role of an observer deciding which painting of each pair shown is more pleasing, artistic, or satisfying. Compare the two paintings of each pair and choose the one that, according to your judgment or feelings, is better. Use the criteria that you think are appropriate for making your choices.

You will need to make a choice even if you feel that the two paintings are very close in value.

PLEASE DO NOT TURN TO OTHER PAGES UNTIL INSTRUCTED TO DO SO.

Instructions for the Second Series:

This second series you will see contains the same 35 pairs of paintings, but in a different order. You will again choose one of each pair and mark your choice on the answer sheet according to its left or right position on the screen.

Please make the same kind of choice for this series as you did for the first. Use the same decision-basis: comparative choice. Decide which painting of each pair shown is more pleasing, artistic, or satisfying.

PLEASE DO NOT TURN TO OTHER PAGES UNTIL INSTRUCTED TO DO SO.

APPENDIX I

LEVEL 6 INSTRUCTIONS: JUDGMENT-JUDGMENT

Type of Decision for This First Series: Critical, Objective Judgment of Artistic Quality

For this series, take the role of an art critic deciding which painting of each pair shown is the better work of art. Compare the two paintings of each pair and choose the one that, according to your objective judgment, is of greater artistic merit. Use the standards that you think are appropriate as the criteria for making your critical evaluations.

You will need to make a choice even if you think that the two paintings are very close in aesthetic value. Choose the one painting of each pair that, in your judgment, is better.

PLEASE DO NOT TURN TO OTHER PAGES UNTIL INSTRUCTED TO DO SO.

Instructions for the Second Series:

This second series you will see contains the same 35 pairs of paintings, but in a different order. You will again choose one of each pair and mark your choice on the answer sheet according to its left or right position on the screen.

Please make the same kind of choice for this series as you did for the first. Use the same decision-basis: critical, objective judgment of artistic quality. Decide which painting of each pair shown is of greater artistic merit.

PLEASE DO NOT TURN TO OTHER PAGES UNTIL INSTRUCTED TO DO SO.

APPENDIX J

Questionnaire

Name _____ Student I.D. # _____

College Level (check one):

Freshman (fewer than 40 credits)	_____	Graduate	_____
Sophomore (40-84 credits) . .	_____	Other	_____
Junior (85-129 credits) . . .	_____		
Senior (130 or more credits)	_____		

Sex:

Male _____ Female _____

College Major (such as Science Ed., Art, English): _____

College Minor: _____

Number of college art courses (such as studio, art history, art education methods) you are taking this semester:

0 _____ 1-2 _____ 3 or more _____

Number of college art courses you have taken, not including this semester:

0 _____ 1-2 _____ 3-9 _____ 10 or more _____

Average number of hours you now spend per week in art or art-related activities (such as art production, museum visits), not including actual art class meeting time:

Less than 1 hour _____ 1-3 hours _____ More than 3 hours _____

Number of high school art courses you have taken:

0 _____ 1-2 _____ 3 or more _____

Your level of interest in art is:

No interest _____ Low interest _____ Moderate interest _____
High interest _____

Were the choices you made in the 1st series different from the choices made in the 2nd series?

No different _____ Somewhat different _____ Clearly different _____

Were the decision-bases specified for the two series different from one another?

No different _____ Somewhat different _____ Clearly different _____

If you would like to receive a copy of the study's findings indicate mailing address:

Street

City

State

Zip

APPENDIX K

STATEMENT READ TO EXAMINEES PRIOR TO TESTING

A research study is being conducted to examine how people react to art in the hopes of extending knowledge in this field. Your help is being requested for this study. Your participation will involve seeing two series of color slides of paintings produced during the 15th through 20th centuries. The paintings will be presented in pairs for comparison.

You will be asked to indicate your reactions to the paintings simply by marking choices on an answer sheet. Although you'll be asked to put your name on the answer sheet, the information will be used only in terms of group data and not on an individual basis.

The total time involved is about 45 minutes. Your cooperation in this study will be extremely helpful. If you are interested, you can receive the results after the data have been analyzed.

You'll each be given a marking pencil and a packet of material. Because the testing procedure needs to be standardized, please observe the following instructions:

Keep the packet as arranged. Read only the top page of the packet. Do not turn to other pages until instructed to do so. It is essential that you not discuss test items or your reactions to them with others in the room.

APPENDIX L

SEQUENCE OF TESTING

The sequence of the testing session was as follows:

1. Prior to testing, an introductory statement that briefly described the study was read to examinees. This statement appears in Appendix K. At the conclusion of the statement anyone who had taken the test previously was asked to identify himself. No one in any of the groups had taken the test before.

2. Pencils and packets were passed out and extra pencils were positioned in the room. Examinees were asked and were given time to read the top instruction page of the packet. The following statement was read to examinees:

To summarize the instructions, you will choose one painting of each pair as the pair is shown and indicate your choice by blackening the appropriate space on the answer sheet. If the chosen painting is in the left position on the screen, blacken space A. If it is in the right position, blacken space B.

3. Procedural questions were asked for and answered.

The following statement was read to examinees:

Will you turn now to the second page of the packet, which is titled "Type of Decision for the First Series," and read that page.

4. Examinees were given time to read the instructions for the first series. The following statement was read to examinees:

Please slip out the answer sheet, place it on top of the packet, and write just your name where indicated. Then turn the answer sheet to its upright position and check the way the items are numbered.

APPENDIX L
(Continued)

5. After an appropriate period, the following statement was made:

We'll now see the first series. The lights will be dimmed but you will still be able to see and mark your answer sheet.

6. (Hereafter lights were dimmed and raised as needed.) The first series of slides was administered. The following statement was read to examinees:

Please turn to and read the page in the packet titled "Instructions for the Second Series."

7. Examinees were given time to read the instructions for the second series. The following statement was read:

The same pairs of paintings will be shown, but in a different order. Again, choose one of each pair according to the specified decision and mark your choice on the answer sheet starting with Item 37. We'll now see the second series.

8. The second series of slides was administered.

9. The need for keeping the testing procedure confidential from subsequent examinees was explained. Examinees were asked to fill out their questionnaires and were thanked for their participation. Pencils, packets, and answer sheets were collected.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Barron, Frank. Complexity-Simplicity as a Personality Dimension. Journal of Abnormal and Social Psychology, 1953, 48, 163-172.
- Barron, Frank, & Welsh, S. Artistic Perception as a Possible Factor in Personality Style. Journal of Psychology, 1952, 33, 199-203.
- Beardsley, M. C. Aesthetics and Criticisms. In R. A. Smith (Ed.), Aesthetics and Criticism in Art Education, pp. 305-314. Chicago: Rand McNally, 1966.
- Birkhoff, G. D. Aesthetic Measure. Cambridge, Massachusetts: Harvard University Press, 1933.
- Boneau, C. A. The Effects of Violations of Assumptions Underlying the t Test. In J. A. Steger, Readings in Statistics for the Behavioral Scientist. New York: Holt, Rinehart & Winston, 1971.
- Bulley, M. H. An Enquiry as to the Aesthetic Judgments of Children. British Journal of Educational Psychology, 1934, 4, 162-182.
- Bullough, Edward. Aesthetics. Stanford: Stanford University Press, 1957.
- Buros, Oscar. (Ed.) Seventh Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1972.
- Burt, Cyril. The Psychology of Art. In Cyril Burt (Ed.), How The Mind Works, pp. 267-310. New York: Appleton-Century Co., 1934.
- Canaday, John. Mainstreams of Modern Art. New York: Simon Schuster, 1959.
- Child, I. L. Personal Preference as an Expression of Esthetic Sensitivity. Journal of Personality, 1962, 30, 496-512.
- Child, I. L. Personality Correlates of Esthetic Judgment in College Students. Journal of Personality, 1965, 33, 476-511.
- Child, I. L., Ford, C. S., & Prothro, E. T. Some Trans-cultural Comparisons of Aesthetic Judgment. Journal of Social Psychology, 1966, 68, 19-26.

- Child, I. L., & Iwao, Sumiko. Comparisons of Esthetic Judgments by American Experts and by Japanese Potters. Journal of Social Psychology, 1966, 68, 27-33.
- Child, I. L., & Iwao, Sumiko. Personality and Esthetic Sensitivity. Journal of Personality and Social Psychology, 1968, 8, 308-312.
- Dewar, Heather. A Comparison of Tests of Artistic Appreciation. British Journal of Educational Psychology, 1938, 8, 29-49.
- Dewey, John. Art As Experience. New York: Minton, Balch & Co., 1934.
- Edwards, A. L. Techniques of Attitude Scale Construction. New York: Appleton-Century-Crofts, 1957.
- Eysenck, H. J. The General Factor in Aesthetic Judgments. British Journal of Psychology, 1940, 31, 94-102.
- Eysenck, H. J. Type Factors in Aesthetic Judgments. British Journal of Psychology, 1941, 31, 262-270.
- Eysenck, H. J. Sense and Nonsense in Psychology. Bungay, Suffolk, Great Britain: Richard Clay & Co., 1957.
- Fechner, Gustav. Elements of Psychophysics, 1860. Translated by H. E. Adler. New York: Holt, Rinehart & Winston, 1966.
- Feldman, E. B. Art as Image and Idea. Englewood Cliffs, New Jersey: Prentice-Hall, 1967.
- French, J. W., & Michael, W. B. Standards for Educational and Psychological Tests and Manuals. American Psychological Association, 1966.
- Glendening, Linda. POSTHOC: A Fortran IV Program for Generating Confidence Intervals Using Either Tukey or Scheffé Multiple Comparison Procedures. Occasional Paper No. 20, Office of Research Consultation, College of Education, East Lansing: Michigan State University, 1973.
- Gotshalk, D. W. Art Criticism. In R. A. Smith (Ed.), Aesthetics and Criticism in Art Education, pp. 343-363.
- Graves, Maitland. Design Judgment Test and Manual. New York: Psychological Corp., 1948.

- Kruglanski, A. W. Much Ado About the "Volunteer Artifacts." Journal of Personality and Social Psychology, 1973, 28, 348-354.
- McElroy, W. A. Aesthetic Appreciation of Aborigines of Arnhem Land: A Comparative and Experimental Study. Oceania, 1952, 23, 81-94.
- Mehrens, W. A. & Lehmann, I. J. Measurement and Evaluation in Education and Psychology. New York: Holt, Rinehart & Winston, 1973.
- Meier, N. C. The Meier Art Tests: I. Art Judgment and Examiner's Manual. Iowa City: State University of Iowa, 1942.
- Meier, N. C. The Meier Art Tests: II. Aesthetic Perception and Preliminary Manual. Iowa City: The University of Iowa, 1967.
- Meier, N. C., & Seashore, C. E. The Meier-Seashore Art Judgment Test and Examiner's Manual. Iowa City: State University of Iowa, 1928.
- Mendelson, Marilyn. The Relationship Between Aesthetic Judgment and Personal Preference for Works of Art. Unpublished Pilot Study, 1973.
- Norris, E. L., & Goodwin, Barbara. (Eds.) National Assessment of Educational Progress: Art Objectives. Ann Arbor, Michigan: National Assessment Office, 1971.
- Pickford, R. W. Factorial Studies of Aesthetic Judgment. In A. A. Roback (Ed.), Present-Day Psychology, pp. 913-929. New York: Philosophical Library, 1955.
- Pyron, Bernard. Rejection of Avant-Garde Art and the Need for Simple Order. Journal of Psychology, 1966, 63, 159-178.
- Rosen, J. C. The Barron-Welsh Art Scale as a Predictor of Originality and Level of Ability Among Artists. Journal of Applied Psychology, 1955, 39, 366-367.
- Schutz, H. A. Review. In Oscar Buros (Ed.), Fourth Mental Measurements Yearbook, pp. 223-225. Highland Park, New Jersey: Gryphon Press, 1953.
- Sibley, Frank. Aesthetic Concepts. In R. A. Smith (Ed.), Aesthetics and Criticism in Art Education, pp. 332-342. Chicago: Rand McNally, 1966.

- Sloan, Patricia. Art and Art Token in Tests Involving Art Evaluation. Psychological Record, 1970, 20, 191-196.
- Smith, R. A. Aesthetic Criticism: The Method of Aesthetic Education. Studies in Art Education, 1968, 9, 12-31.
- Smith, R. A. Aesthetic Criticism: The Method of Aesthetic Education. In R. A. Smith (Ed.), Aesthetics and Problems of Education, pp. 473-484. Urbana, University of Illinois Press, 1971.
- Thomas, R. M. A Rationale for Measurement in the Visual Arts. Educational and Psychological Measurement, 1965, 25, 163-189.
- Ushenko, A. P. Dynamics of Art. Bloomington, Indiana: Indiana University Press, 1953.
- Weitz, Morris. The Role of Theory in Aesthetics. In R. A. Smith (Ed.), Aesthetics and Criticism in Art Education, pp. 84-94. Chicago: Rand McNally, 1966.
- Wilson, B. G. Evaluation of Learning in Art Education. In B. S. Bloom, J. T. Hastings, & G. F. Madaus, Handbook on Formative and Summative Evaluation of Student Learning, pp. 499-558. New York: McGraw-Hill Book Co., 1971.

MICHIGAN STATE UNIV. LIBRARIES



31293104711019