

Æsthetic Thought, Critical Thinking and Transfer

ABIGAIL C. HOUSEN

Visual Understanding in Education

OVERVIEW OF STUDY

About 10 years ago, my colleagues and I set out to explore the relationship between our Visual Thinking Strategies curriculum, (VTS), and the development and transfer of critical thinking skills. VTS is an art-viewing program originally designed to develop aesthetic understanding: the range of thoughts and feelings that occur when looking at art.

Since the curriculum's early use, anecdotal reports from teachers suggested an unexpected result: students were using VTS thinking strategies in other subjects.

In 1993, we had the opportunity to design and implement a longitudinal study to test not only for the effect of our curriculum in stimulating aesthetic growth, but also to look for evidence that it develops critical thinking, and its transfer. We chose Byron, Minnesota, a farming community eighty-five miles from Minneapolis as our implementation test site. The Byron Study consisted of a five-year partnership between the Byron School District, the Minneapolis Institute of Arts, and Visual Understanding in Education (VUE).¹ VUE co-founder and VTS curriculum co-author, Philip Yenawine, senior research associate, Karin deSantis, and I worked closely with site director Catherine Egenberger in this complex, longitudinal project. In a controlled experimental design, we implemented VTS curricula and observed experimental and control students in each of two age groups (beginning with 2nd and 4th graders).²

The results of our five-year study supported our hypothesis that our curriculum accelerates aesthetic growth. Moreover, we found evidence that VTS causes the growth of critical thinking and

enables its transfer to other contexts and content. Ultimately, our results serve not only as a window into the kinds of thinking and learning that occur when elementary age students respond to works of art over an extended period, but also how learning in the arts can enable students to move beyond the interpretation of images.

Visual Thinking Strategies

VTS is a sequential curriculum that includes in-school teacher-facilitated discussions about art, yearly art museum visits, and teacher training.³ The curriculum grows out of my theory and research about aesthetic development, and is designed to match images and questions to the aesthetic developmental needs and naturally occurring capacities of beginner viewers (Housen, 1983). Our lessons are designed to be carried out in an environment of group discovery. The classroom teacher poses a sequence of open-ended questions about a series of carefully selected images of art works. Discussions unfold in which students puzzle and construct meaning about the art works. The learners are given a lot of 'time on task,' and have ample opportunity to build meaning one way and then another. They are also exposed to the thinking of their peers, which can accelerate shifts in their own thinking.⁴

Critical Thinking and Transfer

There are numerous ways in which VTS methods and materials would be theoretically likely to produce growth in evidentiary reasoning. Notably, the curriculum's three foundation questions prompt beginner viewers to reason about evidence by asking 'What is going on here?', 'What do you see that makes you say that?', and 'What more can you find?' Responding to the first question is relatively effortless for students since it invites them to begin with their natural inclination to record their observations or tell a story⁵ (see Stages I & II in Appendix); but it also invites them to re-examine the image carefully, mining for deeper meaning. The second question, 'What do you see that makes you say that?' challenges students by calling for learners to cite evidence for their interpretative comments. And, the third, 'What more can you find?' encourages students to keep on searching.

These questions promote extended, careful and intricate observations. They focus learners, allow choice, require learners to be active, call for reflection, invite many kinds of responses as well as change in responses, allow group participation, and elicit responses which provide a source of information and learning for further discussion. Directed towards carefully chosen art images, the questions create a kind of 'critical thinking studio' in which learners observe carefully, evaluate, synthesize, justify and speculate – habits of mind which have a long history in education and which we find central to aesthetic growth and critical thinking.⁶

Dewey, in particular, championed this complex type of thinking, advocating that the role of education was to develop "...effective habits of discriminating tested beliefs from mere assertions, guesses, and opinions [...] sincere, and open-minded preference for conclusions that are properly grounded, and [...] methods of inquiry and reasoning appropriate to the various problems that present themselves." (Dewey, 1910/1997, p. 28). Such thinking bears an elusive relationship to subject matter. On the one hand, critical thinking cannot be developed in a vacuum and needs a subject matter as a medium for its exercise and development. On the other hand, critical thinking transcends the subject matter in which it initially develops. Jerome Bruner posited that transfer involved the capacity to go beyond the information given, calling attention to the fact that transfer extends thinking beyond where such thinking was first learned (Bruner, 1973). While critical thinking takes root in one area, it has the potential to blossom in others. One could even argue that transfer is a predictable attribute of critical thinking. Critical thinking may not be critical thinking unless it shows signs of transfer. (Anderson, Reder, & Simon, 1997; Bransford & Schwartz, 1999; Cobb & Bowers, 1999; Dewey, 1910/1997; Kirshner & Whitson, 1998; Kuhn, 1986, 1999-a; Piaget, 1970; Salomon, 1993; Vygotsky, 1987).

Study Design and Measures

Testing a transfer hypothesis is especially difficult. Our approach eight years ago was to rely upon an extended longitudinal study, one comprised of ten data collections (every Fall and Spring) combined with on-going observations, in a controlled experimental design. Using this approach as a base, we found a way to assess multiple forms of transfer to provide evidence of sequential

patterns of change, thereby avoiding the typical difficulties of correlational studies. It is commonly understood that establishing a significant association between a and b in one observation period does not make it possible to definitively answer the question of which variable causes the other. By contrast, through using multiple measures over multiple observation periods, it is possible to observe a pattern of statistically significant shifts over time; changes that make it possible to identify a causal relationship.

We selected two schools from matching neighboring towns, one we designated experimental and the other control. The entire experimental school received VTS. The schools were matched in terms of SES, race, mobility, and state test scores. From Fall 1993 to Spring 1998, beginning with 2nd and 4th graders, we collected pre and post data from the same subjects: fifty-two randomly selected students from the experimental school, and forty-seven randomly selected subjects from the control school.⁷ We, also, collected data from twelve experimental teachers and the control school art teacher.

Our study and measurement procedures included full descriptions of treatment and outcomes: our treatment was VTS curriculum, as well as VTS teacher training. Our measures, included below in Table 1, summarize the various streams of information we collected throughout the study to support and explain growth in aesthetic thought, critical thinking, and the transfer of critical thinking.

Measures of Aesthetic Understanding, Critical Thinking, and Transfer

Perplexing obstacles confront researchers in the design of an appropriate measure for both critical thinking and the transfer of critical thinking strategies beyond the original learning context. Our concern was that general observation methods – multiple choice instruments, logs, questionnaires, and observation forms – would neither capture the elusive phenomenon of critical thinking, nor yield data that would measure its growth or transfer. Rather, we concluded that accurately and sensitively measuring the transfer of critical thinking had distinct and inherent requirements.

Foremost among them were: non-obtrusive observations that do not distort the phenomenon itself, content that is respondent-structured, and multiple lines of evidence from which to draw conclusions.

To capture the types of behaviors we were attempting to study, we used two principal instruments: the Aesthetic Development Interview (ADI), which I designed as part of my early research (Housen, 1983); and the Material Object Interview (MOI), which was designed specifically for this study. The MOI was designed along the lines of the ADI, and differs primarily in one way: in the ADI, learners puzzle out loud about art; in the MOI, learners puzzle about non-art objects. The non-art objects used in the MOI were unusual and unfamiliar objects taken from material culture or the natural world. Our objects included a Danish coin, a trilobite fossil, calipers, a mortar and pestle, an anemometer, and a candle-snuffer.

The ADI and MOI — both non-obtrusive, non-directive, stream-of-consciousness interviews — are in fact, monologues which effectively sample a person's thinking process. They are open and sensitive enough to capture ephemeral phenomena, create minimum 'disturbance', allow the learner to express any kind of idea in any way desired, and they can be used repeatedly without systematically biasing results.⁸ Each monologue is tape-recorded, transcribed and analyzed by breaking it into 'thought units,' independent ideas, often only a few words long. An example of two thought units is: "I think the trees are black / And some of them are tall."

Methodology

In the ADI, thought units are coded in two ways. First, we code for Aesthetic Stage. We do this by classifying thought units using categories in the Aesthetic Development Coding Manual (Housen, 1983). Using this framework — direct samplings of aesthetic responses in process — it is possible to give an Aesthetic Stage score to the overall thought patterns of our subjects (Housen, 1983). (See Appendix for detailed Stage descriptions). Moreover, since the ADI method is extremely sensitive to change, it allows us to

Table 1. Byron, Minnesota Study: Research Instruments

Measures	Description
Aesthetic Development Interview (ADI)	An open-ended monologue documenting a student's thoughts and reasoning skills as s/he responds to a work of art. ADIs are non-directed monologues, delivered one-on-one to a researcher.
Material Object Interview (MOI)	A non-directed monologue focusing on an artifact such as a mortar and pestle or a set of calipers. Delivered one-on-one to a researcher.
Demographic questionnaire	Questionnaire read or given to each student.
Art and Museum Biographies	Questionnaires which detail the student's personal history regarding art and museums.
Art, Museum, Artist Questions	Open-ended questions included in the questionnaire asking students about art related content.
Writing samples	School and VTS assignments, executed in school, at home and during museum visits. (These include writing from Byron student journals, which Byron students begin in kindergarten and continue until the completion of elementary school.)
Teacher logs	Teachers were asked to keep a log of their observations of student behavior with particular focus on student thinking and learning behaviors. They also observe their own teaching styles.
Teacher Trainer and Site Coordinator Notes	Written Observations and notes from debriefings, including a post-study debriefing of the experimental students and their parents.
Videotapes	Some VTS classroom lessons and videotaping of experimental teacher training sessions were videotaped for analysis.
Student Exit interviews	Debriefing interviews.

Constructs measured	Scoring method	Collection timing
Aesthetic Stage level; Context Transfer.	Ogive and clinical scores.	Collected bi-annually.
Content Transfer.	Frequency, Category and Attribute Classifications.	Collected bi-annually.
Demographic variables.	Frequency, Category and Attribute Classifications.	Collected bi-annually.
Variables related to background experience with art.	Frequency, Category, and Attribute Classifications.	Collected bi-annually.
Variables related to background experience with art.	Frequency, Category, and Attribute Classifications and Clinical analysis.	Collected bi-annually. Analysis not included in this paper.
Content Transfer.	Frequency, Category, and Attribute Classifications and Clinical analysis.	Collected throughout the year from experimental groups. Analysis not included in this paper.
Teacher performance and perception of students.	Frequency, Category, and Attribute Classifications and Clinical analysis.	Ongoing. Collected Annually. Analysis not included in this paper.
Implementation, Teacher and student performance, and Administrative Feedback.	Clinical analysis.	Ongoing. Collected Annually.
Variables related to transfer. Not used in this study.	Frequency, Category, Attribute Classifications and Clinical analysis.	Selected groups. Not included in the analyses of this paper.
Transfer.	Frequency, Category, Attribute Classifications and Clinical analysis.	Final interview of experimental groups.

observe small changes of half Stages. Second, ADIs are independently coded for critical thinking counts (described below).

MOI monologues are coded solely for critical thinking counts. Like ADIs, MOI monologues contain a rich array of thoughts, including a variety of observations, interpretations, clarifications, supported observations and speculations. For this paper we isolated two categories for analysis: supported observations and speculations. We see these variables as good examples of contemporary theory which characterizes critical thinking as “the art of thinking about your thinking” (Paul, 1990, p. 23). Critical thinking scores, then, are the sum of counts of Supported Observations (thoughts that entail an assertion that is grounded in evidence) and Speculations (thoughts about a possible meaning or outcome based on evidence). The following student response illustrates a Supported Observation: “*I think they welded the handle onto it because it kind of looks like the metal kind of melted and then hardened.*” The student’s hypothesis that the handle is welded is based on her observations of visual properties of the metal, or its ‘melted-then-hardened’ quality. The next response illustrates a Speculation: “*And there’s a little slot for it to lay on in the side. Probably to crush stuff, too.*” Here, the student offers a conjecture about the slot, which allows the object to lay on its side. In the analysis of both monologues, we established a high level of inter-rater reliability.⁹

Transfer of Critical Thinking

We distinguished between two kinds of transfer of critical thinking within our study: transfer of context and transfer of content.

Transfer I: Context Transfer

Here, we looked for critical thinking strategies in a social setting different from the one in which such thinking was learned. The learning context of VTS was the social setting of the classroom that included teacher-facilitated group discussions, open-ended questions, paraphrasing and linking of student responses. By

contrast, in the ADI context, the individual subjects talked out loud by themselves about what they saw in a work of art. The ADI context was different from the student's learning environment in several significant respects: the subject was alone, no teacher was prompting for critical thinking, no peers were modeling the behavior, nor adding comments that might stimulate thought, and there was no group expectation to support evidentiary reasoning. The only constant between the two settings was the topic under examination, namely art.

Transfer II: Content Transfer

The second kind of transfer we looked for was the exercise of critical thinking applied not only to a different social context but also within a different subject domain, or content. We call this second transfer, simply, Content Transfer.

Would VTS-trained students apply critical thinking strategies to a non-art object of attention? In the MOI, we have both a shift of context, from classroom discussions to individual monologues, and of content, from a work of art to an object from material culture or the natural world. While Content Transfer is logically related to Context Transfer, we believe the two are meaningfully distinct.

HYPOTHESES

The study entailed five hypotheses.

Hypothesis 1: Context Transfer. The first hypothesis is that critical thinking transfers across social context – in this case, from classroom dialogues to individual monologue. To test this hypothesis, we coded ADIs for instances of critical thinking.

Hypothesis 2: Content Transfer'. The second hypothesis is that critical thinking shows transfer not only across social context, but across content as well. In this case, the transfer is from classroom dialogues to monologue, and from art to a non-art object. To test this hypothesis, we coded MOIs for instances of critical thinking.

Hypothesis 3: Sequence effects. The third hypothesis is that transfer scores show sequence effects. This means that we would not expect to see both types of transfer occurring immediately and simultaneously. Rather, significant differences in mean score between experimental and control groups for each type of transfer would occur at distinct times. To test this hypothesis, we examined incidences of critical thinking counts in ADIs and MOIs over time.

Hypothesis 4: Developmental Effects. The fourth hypothesis is that Content Transfer scores increase significantly by developmental level. To test this hypothesis, we coded ADIs for Aesthetic Stage level and MOIs for critical thinking counts to assess the impact of developmental level on critical thinking.

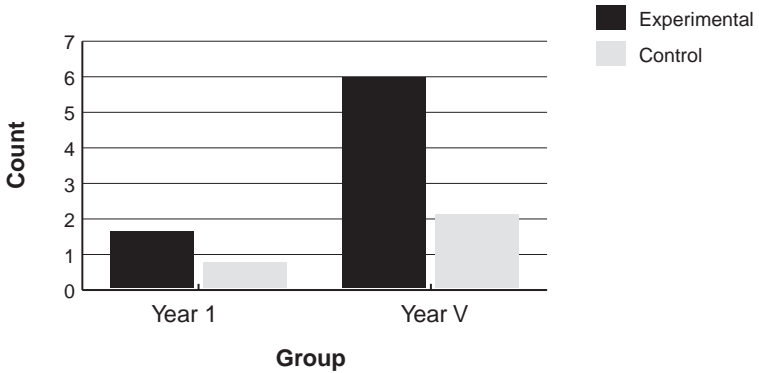
Hypothesis 5: Developmental Growth The fifth hypothesis is that exposure to VTS over time causes students to achieve higher Aesthetic Stages. Average Stage scores derived from the ADIs for each of the ten observation periods were compared for experimental and control groups to determine whether the experimental groups' average scores increased at a higher rate than the controls'.

Findings

Below we will review the findings by hypothesis.

Transfer Across Social Context (Hypothesis 1). Our hypothesis was that critical thinking would show transfer across social context. Figure 1 shows an increase in the differences in mean scores in critical thinking found in the ADI for experimental and control groups from Year I to V. While both groups show a gain over that period, the experimental group did show significantly more Context transfer than the control group in Year V.¹⁰ (Year I: $F= 1.647$; $df = 1,60$; $p <.2043$; Year V: $F=15.234$; $df =1,60$; $p<.0002$).

Figure 1: Mean Context Transfer Scores (ADI) by Group and by Year (Byron Study Year V, Spring)



Transfer Across Content (*Hypothesis 2*). The second hypothesis was that critical thinking would not only show transfer across social context but also across the content to which it was applied. Over time, the experimental groups were expected to show higher scores than control groups when applying Supported Observations and Speculation to non-art objects (MOIs). A simple test of Content Transfer was to combine both age groups and compare mean critical thinking scores of the experimental group to that of the control group. Figure 2 shows that the mean critical thinking score of the experimental group at the end of Study Year V was more than twice that of the control group ($F= 6.409$; $df =1,84$; $p<.01$).

Figure 2: Mean Content Transfer Scores (MOI) by Group (Byron Study Year V, Spring)

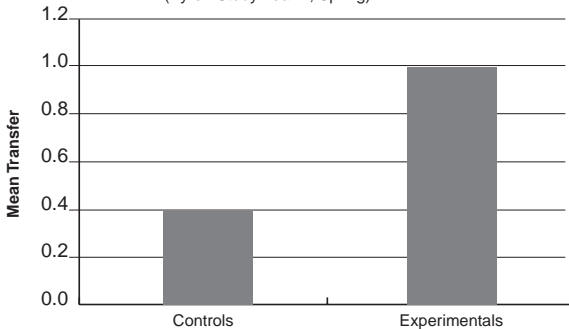


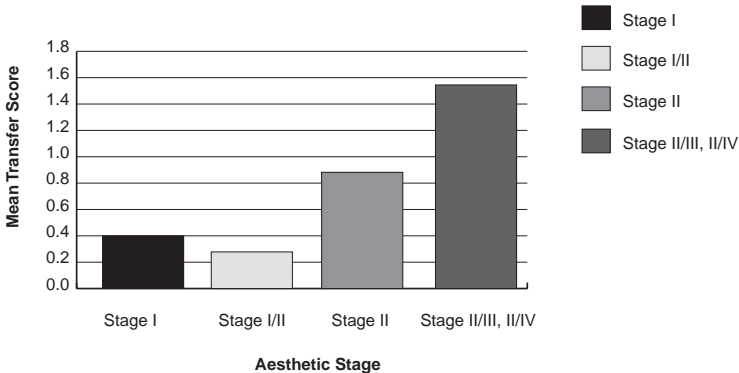
Table 2. Sequence of Transfer (Byron Study Years I to V)

		Year I		Year II		Year III		Year IV		Year V	
Young		Score	p<	Score	p<	Score	p<	Score	p<	Score	p<
Context	Experimental	0.783	.011*	1.545	.020*	2.100	.0006***	2.091	.24	7.480	.002***
	Control	0.000		0.190		0.217		1.682		2.050	
Content	Experimental	0.000	.856	0.403	.107	0.450	.08~	0.762	.880	1.350	.024*
	Control	0.050		0.190		0.174		1.455		0.619	
Older											
Context	Experimental	1.590	.378	1.304	.087~	1.864	.018*	2.818	.026*	4.435	.004***
	Control	1.230		0.750		0.682		1.621		2.091	
Content	Experimental	0.261	.284	0.273	.341	0.600	.620	1.227	.173	1.419	.047*
	Control	0.095		0.208		0.682		0.828		0.251	

Legend: Stronger patterns have more asterisks (* = $p < .05$; ** = $p < .01$; *** = $p < .005$)

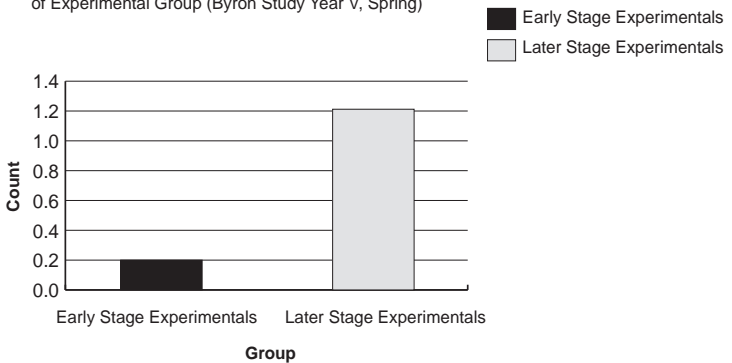
Sequence Effects (Hypothesis 3). The third hypothesis is that transfer scores would show sequence effects. Table 2 presents the sequence pattern for both kinds of transfer. Significant differences in Context Transfer scores begin to be measurable about one and one-half years into the project (Spring of Year II), and continue to appear with general consistency to the end of the study. By contrast, Content Transfer differences remained non-significant well into the later years of the study, and well past the point at which the experimental group was showing significant gains in Context Transfer. This finding strongly confirms that there was a sequence to the emergence of transfer, with Context Transfer appearing earlier than Content Transfer.

Figure 3: Mean Content Transfer Scores(MOI) by Aesthetic Stage of Experimental and Control Groups Combined (Byron Study Year V, Spring)



These are the general patterns regardless of treatment. But what is the impact of VTS on Content Transfer in the experimental group, taken alone, when we take developmental level into account? To examine this question in a second analysis, we combined the younger and older experimental groups. Then we divided this group into two: those at 'early' Stages (Aesthetic Stages I and I/II) and those at 'later' Stages (Aesthetic Stages II and above). Figure 4 shows mean Content Transfer scores at the end of Study Year V. The 'later' Aesthetic Stage group showed a mean that was roughly six times greater than that of the 'early' Aesthetic Stage group ($F=4.95$; $df=1,41$; $p<.032$).

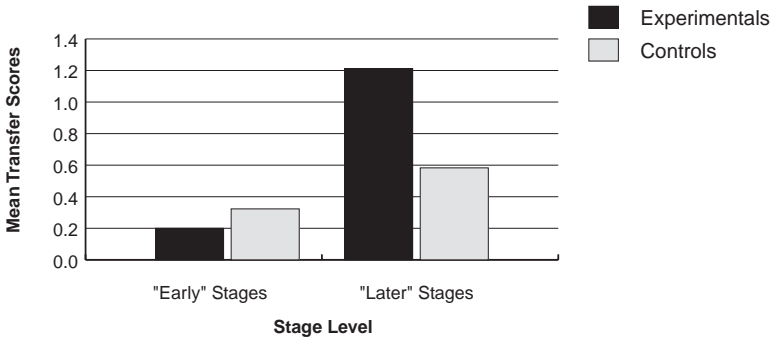
Figure 4: Mean Content Transfer Scores (MOI) by "Early" and "Later Aesthetic Stage of Experimental Group (Byron Study Year V, Spring)



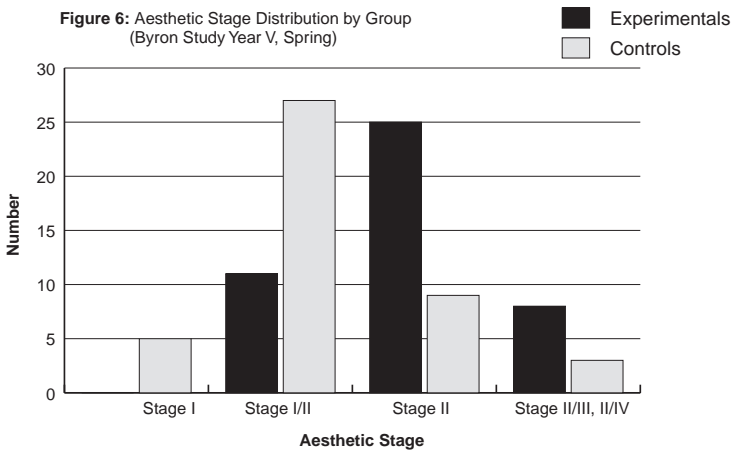
In a third analysis, we again looked at the interaction of developmental stage and critical thinking, this time taking both experimental and control groups into account.

We compared mean scores of the experimental and control groups at 'early' and 'later' Aesthetic Stages. (Figure 5). Again, we see the pattern that critical thinking scores in MOIs increase at Stage II, but do so more sharply with students exposed to VTS. VTS experimental group at Stage II and higher shows twice the mean level of Content Transfer of critical thinking and these differences are significant. These patterns suggest that gain in developmental level for the experimental group appears to interact with Content Transfer ($F= 4.78$; $df=3,81$; $p<.004$).

Figure 5: Mean Content Transfer Scores (MOI) by "Early" and "Later" Aesthetic Stage and Group (Byron Study Year V, Spring)



Change in Aesthetic Stage (*Hypothesis 5*). Our final hypothesis was that VTS causes an increased rate of Aesthetic Development. One way to understand the cumulative impact of VTS on Aesthetic Stage is to look at the distribution of the experimental and control groups by Stage at the end of Study Year V. In Figure 6, which compares the number of subjects at each Stage for the experimental and control groups, we see that that by the end of Study Year V, the experimental group has a much higher distribution than the control group. This difference is strongly significant ($t=4.70$, $df=86$, $p<.0001$). It is worth noting that while the control group subjects are concentrated at Stage I/II, the bulk of the experimental group is within the range of Stage II. In short, the experimental group has grown above the threshold at which we have already seen that Supported Observations and Speculation are more likely to flourish. This fact will become important in the following section when we discuss how the transfer process appears to operate.



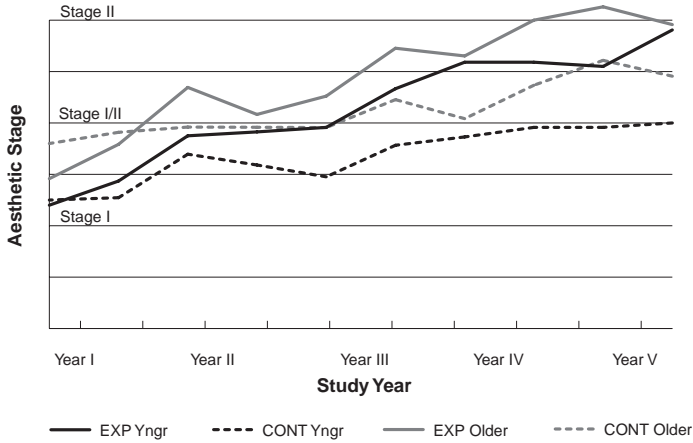
Our final figure (Figure 7) compares aesthetic development mean Stages of the experimental and control groups for all five years of the study. One way to understand the cumulative impact of VTS on Aesthetic Stage is to look at the increasing difference in mean Aesthetic Stage by groups. Both groups began with similar mean Stage scores (control group being slightly higher) but the experimental groups gain is significantly greater than the controls.¹¹

Findings Epilogue

Researchers sometimes encounter the unanticipated. From time to time, unplanned events can lead to unexpected findings. In our case, we had not one, but five unplanned aspects in the study. Since many of these circumstances actually undermined the likelihood that the experimental group would outperform the control group, these factors can actually be seen as adding strength to our argument.

The first three unanticipated conditions had to do with sample and implementation. First, the older experimental group happened to have more special needs students and fewer gifted students than the older control group. Second, in our Year II data analysis, we discovered that the fifth grade teachers (of the older experimental

Figure 7: Mean Aesthetic Stage by Group (Byron Study Year I-V)



	Year I	Year II	Year III	Year IV	Year V					
Exp. Grp / Yngr	1.2	1.435	1.875	1.913	1.957	2.333	2.591	2.591	2.55	2.905
Cont Grp / Yngr	1.25	1.273	1.696	1.591	1.476	1.783	1.864	1.957	1.957	2.
Exp. Grp / Older	1.458	1.792	2.346	2.083	2.261	2.727	2.652	3.	3.13	2.975
Cont Grp / Older	1.8	1.909	1.96	1.958	1.955	2.227	2.042	2.367	2.609	2.455

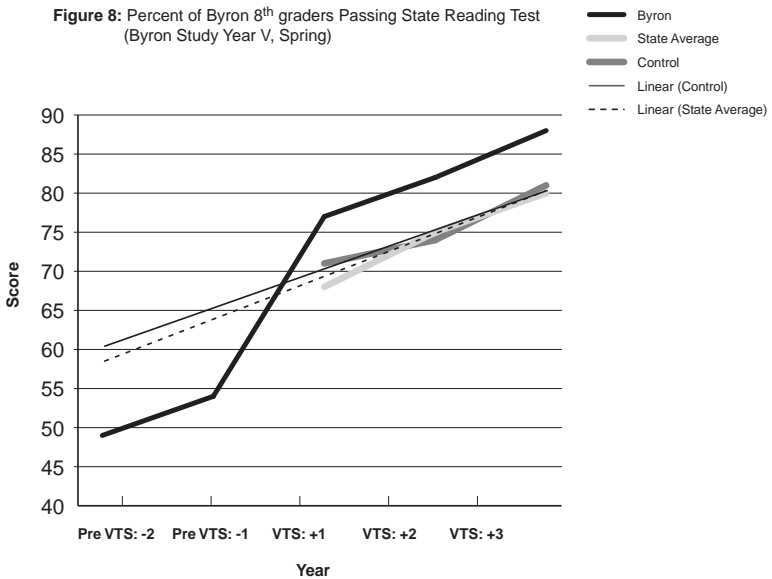
group) stopped teaching VTS midway through Year II implementation.¹² Third, more than a year into the study, we discovered that the controls had been receiving an alternative treatment: art classes taught by a practitioner in Discipline Based Art Education. This art teacher, who was in her fourth year of teaching at the control school, was in fact an award-winning practitioner of Discipline Based Art Education. (In 1998, she won the State’s Art Teacher of the Year Award.) By contrast, the Byron School, which did not have any art teacher for several years prior to our program, went through a series of four art teachers during our study.

The other two unplanned factors were independent evaluations of VTS outcomes: the first, Minnesota achievement test scores, suggesting *general transfer*; the second, a Project Zero Study of two VTS questions, providing an *independent replication of findings*.

General transfer. In the spring of 1996, Minnesota began requiring eighth graders to take a state test of achievement. Every community in the state was ranked annually, by percentile, on the basis of the scores of its eighth graders. The test was first introduced in 1996. In 1997, only 54 percent of Byron 8th graders passed in reading. In 1998, when Byron 8th graders who had been participating in VTS since 1993 took the reading test, 77 percent tested at or above the state’s passing score. (The state score was 68 percent; the control school score was 71 percent.) This gain of 23 points for Byron (1997-1998) was 2.5 times the state average gain of nine points, placing the school in the top 8% for gains.

This trend continued. By 1999-2000, 88 percent of the Byron eighth grade students (our younger experimental group) was at or above that year’s passing score. (The state average for that year was 80 percent; the control was 81 percent.) The Byron school principal, teachers and school board members believed that the school district’s participation in the five-year pilot program of VTS, contributed significantly to Byron’s placement in the top 8 percent of Minnesota schools.¹³

Figure 8: Percent of Byron 8th graders Passing State Reading Test (Byron Study Year V, Spring)



Independent replication of findings.

In the early 1990's, Philip Yenawine and I (with Amelia Arenas) designed a forerunner of VTS, which we called Visual Thinking Curriculum (VTC). This curriculum was also based on my work on Aesthetic Stages from the early 1980's, and on our question-based strategy to promote aesthetic development. The Museum of Modern Art began using the VTC in the early 1990's. Completely independent of the Bryon study, MoMA commissioned Harvard's Project Zero to analyze the effects of the VTC in an unpublished study.

Looking at our first two core questions but lacking a longitudinal design and control pre-tests, the Harvard researchers used a completely different study design, measurement methods and population sample, and yet arrived at conclusions essentially consistent with our study. They found that: "The VTC tends to contribute to a modest but significant increase in students' evidential reasoning skills when [students] are interpreting the meaning of a work of art. These skills also appear to transfer to interpreting the meaning of a non-art image..." (Tishman, MacGillivray, & Palmer, 1999, p. 63).

It is noteworthy in the social sciences when two groups independently study two distinct educational programs – ones, however, based on the same underlying research models and pedagogical assumptions – and, using different research methods, arrive at similar conclusions based on statistical evidence.¹⁴

Discussion

VTS and Transfer. At the most basic level, VTS appears to be teaching critical thinking skills that transfer across social context. By Year II, we saw that experimental students use significantly more Supported Observations and Speculations in their ADIs. Furthermore, when the content is shifted from art to artifact, the experimental groups outperform the control groups.

In short, all the hypotheses we had in this study were confirmed by statistically significant differences. It seems possible to develop

critical thinking strategies (as defined here) with VTS, and such strategies appear to transfer across social context and content area.¹⁵

Correlational or Causal Association?

There is an on-going theoretical and statistical discussion about the inherent dilemmas in proving causality. It is generally understood that causality cannot be proven only disproven. The best, then, that any social scientist can do is to make carefully structured

observations, ones that maximally filter out the possibilities of faulty inferences. With that in mind, accepting the limits of what observations are practically possible in real world settings, we have invoked a wide range of observation methods to maximize the likelihood of detecting the *absence* of causality. Nonetheless, we believe that the results of our study are more than mere findings of correlation. Because of the way we have constructed our observations, there are several reasons for concluding that our data are consistent with a pattern of causality.¹⁶

Identification of more than one kind of transfer. We have attempted to show that there is more than one kind of transfer. We developed separate ways to measure each kind. Each kind of transfer appears to be distinct because the same subjects in the experimental group in the early years of the study showed a significant increase in one type of transfer (Context Transfer) but not the other (Content Transfer). However, by the end of the study, the significant increase in both types of transfer in the experimental group provides strong evidence that VTS was the joint cause.

Sequencing effects. Correlational association is typically hard to distinguish from causal association when one has observations from only one point in time. With longitudinal observations, however, it becomes possible to establish the ordering of effects, especially if one can measure more than one kind of change.

Causal patterns become most compelling in a sequence of observations in which, during an early time period, there is a treatment but no immediate outcome, followed by a subsequent period in which the outcome *a* appears, followed by a period in

which outcome *a* and *b* are both present. We not only had such a sequence, but also found such a pattern. By executing ten, 6-month observation periods, with repeated and diverse measures at each point, we created the necessary conditions to be able to observe a 'causal chain of change.' The fact that there are plausible lags between treatment and the onset of outcome effects such as Context Transfer and Content Transfer, respectively, further reinforces the likelihood of causality.

Context and Content Transfer may be usefully distinctive ideas because they are cumulative, Content Transfer subsuming Context Transfer in that the latter is a necessary but not sufficient condition for the former. As such, the two constructs can act as a 'sequential snapshot' of the transfer process itself, as we have seen in this study. First one occurs, then another. Furthermore we have identified other necessary conditions for Content Transfer, namely developmental level (see below).

Evidence of the process of transfer. It is more likely that the data are evidence of a causal process if they also contain compelling signs of how that causal process may actually operate. Below, we will argue that the data here give a strong basis for understanding the process by which VTS has its impact.

Difficulty in arguing 'reverse causality'. Assertions that findings are only correlational in nature and not causal, directly imply that there is a completely different way to explain the relation of observations in our study. Often this entails arguing for a reverse pattern of causality: *a* does not cause *b*, rather *b* causes *a*. Or, alternatively, both *a* and *b* have a joint hidden cause.

In our case, we entered the study with variety of inter-related hypotheses. The fact that virtually all of them were confirmed suggests that there is theoretical coherence in our arguments. The tenet underlying all of them is that VTS causes the growth of evidentiary reasoning and enables its transfer to other contexts and content. It is not unreasonable for transfer to appear in a new arena with some lag after the repertoire was developed in an original area. One might expect to see first see transfer to a domain that is more similar to the original one in which the repertoire was developed, and later to observe it applied to a domain that is more

distant and dissimilar. If the findings here are merely correlational in nature, what is the corresponding plausible 'reverse causality' argument? The timing of treatment of measured changes establishes an ordering that is hard to argue can be reversed: VTS preceded Context Transfer, which preceded Content Transfer.

Furthermore, there was no evidence in our data that a hidden 'joint cause' was operating. Other typical ecological variables were matched as carefully as possible in our selection of a control community, including factors such as SES, parental background, teacher quality and school climate. On-site coordinators and teacher data kept us apprised of on-going changes in program, school and community activities. In addition, we looked at art and museum biographies of each student (collected bi-annually) to further reduce the likelihood that the observed sequence of changes in the experimental group was the result of correlation with other external causal factors. Indeed, through these other data sources, we identified three extraneous factors discussed above, all of which could have easily have undermined and reduced the significant findings in our study.

The Process of Transfer: Two paths of causality

The evidence we have reviewed points to a complex pattern of causality in the process of transfer of critical thinking. There are two pathways of apparent influence. First, VTS has a direct effect, by giving students the opportunity to practice evidentiary reasoning in a supportive environment, with scaffolds (VTS questions, teacher facilitation, peer discussions) and with high amounts of time on task. Thus, the first effect might be termed the 'Performance Effect.' VTS appeared to increase the frequency of Supported Observations and Speculations in the experimental group by providing an environment that supported the 'performance' of reasoning based upon evidence. Thus, VTS gave both practice and support for the exercise of critical thinking strategies.¹⁷

The second causal pathway might be called a 'Developmental Effect.' We have seen that VTS causes learners in the experimental group to move to higher Aesthetic Stages in a more accelerated way. Furthermore, VTS process does not seem to have the same

outcome for all learners. There appears to be a critical developmental threshold, at Aesthetic Stage II. Subjects at this level or higher display markedly increased Supported Observations and Speculations. Increases in developmental level, by shifting both the propensity and capacity to form certain classes of ideas, may alter the likelihood of the sustained appearance of other phenomena such as Content Transfer.

In our study, achieving Stage II of Aesthetic Development appears to be a necessary but not a sufficient condition for the flourishing of Content Transfer.¹⁸ Thus, VTS appears to create its impact by both raising Aesthetic Stage in the experimental group, and by making those at any Aesthetic Stage, but especially at higher Stages, more capable of exercising evidentiary thinking. Our data show clearly that critical thinking scores in our elementary school population increase sharply by developmental level, an effect that is made much more pronounced by exposure to VTS.

Conclusion

We have made much of methodology, in the sense of demonstrating causality of transfer, by using proper experimental method. But our methodological emphasis does not highlight several other aspects of the study that contribute fundamentally to its achieved outcomes. The first is the type of measurement. This study has shown the value of using an open method of sampling thought as it occurs. Without this stream-of-consciousness method, sensitive enough to capture the phenomenon being studied, and the coding procedure that discriminates separate units of thought, it might well have proven impossible to test the transfer hypothesis, even with a perfect experimental design.

The second aspect of this study addresses the developmental nature of the treatment – namely, the methodology of designing a curriculum from a developmental point of view.

While it is not the purpose of this study to go into the design techniques underlying VTS, we would be remiss not to point out that this vital area is often overlooked in methodological discussions. VTS looks deceptively simple, but the sequencing of questions and images is based on the developmental analysis of

thousands of samples of viewers' reasoning about art, collected over a period of two decades. The VTS process accesses the power of art by focusing attention at a level that is *accessible* and *provocative* to the learner in an appropriate sequence over time. This design is not arbitrary and is continually refined on the basis of on-going research findings.

And, finally, we come to curricular content. Here, the power of art cannot be minimized. Critical thinking cannot be developed in a vacuum; it needs a subject to act upon as an object of thought. For this reason, it is commonly believed that critical thinking is difficult to elicit until the late elementary years when learners have mastered some level of understanding of a domain of knowledge and/or reading. Our evidence, however, suggests that critical thinking and its transfer can be developed as early as the second grade. It seems clear that before some children can read a book, they are quite capable of 'reading' a work of art for meaning and evidence.

It is easy to overlook all the subtle, yet significant ways that art can act as an ideal medium for the development of critical capacities when it is framed with the right kind of pedagogical process. Good, well-selected art has several remarkable attributes that allow children to immediately exercise their thoughts. First encounters with art do not take years of background preparation. A well-chosen work of art is a self-contained world. It has all the information one needs to begin to interpret it. And its presence is a challenge to make new meaning.

- *Art is accessible.* Art can speak to all viewers, allowing them to enter its space early and easily. Children can 'read' a picture long before they can read print.
- *Art touches timeless issues.* Art can take the viewer as deep as the viewer has the capacity to go.
- *Art is compelling.* Seeing a work's meaning change as interpretations grow can rivet attention.
- *Art is ambiguous.* Art has more than one 'right' interpretation. Its crafting contains carefully shaped clues. Its ambiguity invites speculation.
- *Art viewing unfolds.* The more one looks, the more one sees. The interpretative possibilities in art keep unfolding.

Each new viewing of a work of art is a new episode, an invitation to begin the spiral of meaning making all over again. With each new moment of noticing, the work invites the viewer deeper into its world. With each step, new connections are made, the learner grasps more and is drawn another step forward in understanding. When exposed to new works, the viewer repeatedly experiences that moment when something he might have turned away from, becomes something he is curious about and has the capacity to see and to understand.

Our Byron study convinced us that reasoning about art is an effective way to pursue one of education's most elusive goals: the development of critical thinking. That said, we believe that art should not have to be justified as a means to other ends. It is through art that we have experiences we would not have in other ways. Art tells us who we are; it helps us understand what it is to be human. Yet, in a world that increasingly prepares students to pass standardized tests, there is little time left in the classroom to go beyond what schools currently see as 'the basics.' Despite these disturbing trends, there are strong arguments for art – in all its manifestations – to remain a part of the educational landscape. If an art program can develop the capacity to respond to art, while supporting critical educational goals, then that is fortuitous for the art program as well as for the students.

References

- Anderson, J. R., Reder, L. M., & Simon, H. A. (1997, January/February). Rejoinder: Situative Versus Cognitive Perspectives: Form Versus Substance. *Educational Researcher*, 26 (1), 18-21.
- Arnheim, R. (1966). *Art and Visual Perception: A Psychology of the Creative Eye*. Berkeley, CA: University of California Press.
- Arnheim, R. (1969). *Visual Thinking*. Berkeley, CA: University of California Press.
- Baldwin, J. M. (1975). *Thought and Things: A Study of the Development and Meaning of Thought or Generic Logic* (Vols. III and IV). New York: Arno Press.
- Bransford, J., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, 24, 61-100.

- Bruner, J. (1960). *The Process of Education*. Cambridge: Harvard University Press.
- Bruner, J. (1966). *Toward a Theory of Instruction*. Cambridge: The Belknap Press of Harvard University Press.
- Bruner, J. (1973). *Beyond the Information Given*. New York: W.W. Norton & Company, Inc.
- Bruner, J. (1973). *The Relevance of Education*. New York: W.W. Norton & Company, Inc.
- Bruner, J. (1996). *The Culture of Education*. Cambridge, MA: Harvard University Press.
- Burton, J., Horowitz, R., & Abeles, H. (1999, April). *Learning In and Through the Arts: The Issue of Transfer*. Paper presented at the National Conference of the American Educational Research Association, New York, NY.
- Catterall, J. S. (1998, July). Does Experience in the Arts Boost Academic Achievement?: A Response to Eisner. *Art Education*, 51 (4), 6-11.
- Chomsky, N. (1972). *Language and Mind*. New York: Harcourt, Brace, Jovanovich.
- Cobb, P., & Bowers, J. (1999, March). Cognitive and Situated Learning Perspectives in Theory and Practice. *Educational Researcher*, 28 (2), 4-15.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation: Design & Analysis issues for Field Settings*. Boston: Houghton Mifflin Company.
- Cook, T. D., (1999, March). *Considering the Major Arguments Against Random Assignment: An Analysis of the Intellectual Cultural Surrounding Evaluation in American Schools of Education*. Paper presented at the Harvard Faculty Seminar on Experiments in Education, Cambridge, MA.
- Cronbach L. J. (1987). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- Darby, J. T., & Catterall, J. S. (1994). The Fourth R: The Arts and Learning. *Teachers College Record*, 96 (2), 299-328.
- Dewey, J. (1910/1997). *How We Think*. New York: Dover Publications, Inc.
- Dewey, J. (1934). *Art as Experience*. New York: Capricorn Books.
- Duckworth, E. (1996). *'The Having of Wonderful Ideas' and Other Essays on Teaching and Learning*. New York: Teachers College Press.
- Eisner, E. (1998, January). Does Experience in the Arts Boost Academic Achievement? *Art Education*, 51 (1), 7-15.

- Eisner, E. (July 1998). A Response to Catterall. *Art Education*, 51 (4), 12.
- Fosnot, C. T. (Ed.). (1996). *Constructivism: Theory, Perspectives, and Practice*. New York: Teachers College Press.
- Housen, A. (1983). *The Eye of the Beholder: Measuring Aesthetic Development*. Unpublished doctoral dissertation, Harvard Graduate School of Education, Cambridge.
- Housen, A. (1992). Validating a Measure of Aesthetic Development for Museums and Schools. *ILVS (International Laboratory for Visitor Studies) Review: A Journal of Visitor Behavior*, 2 (2), 213-237.
- Housen, A. (2000, March). *Aesthetic Thought: Assessment, Growth, and Transfer*. Paper presented at the annual conference of the National Art Education Association, Los Angeles, CA.
- Housen, A. (2000-2001). Voices of Viewers: Iterative Research, Theory, and Practice. *Arts and Learning*, 17 (1), 1-21.
- Housen, A. (2001, March). *Methods for Assessing Transfer From an Art-Viewing Program*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Kirshner, D., & Whitson, J. A. (1998, November). Obstacles to Understanding Cognition As Situation. *Educational Researcher*, 27 (8), 22-28.
- Kohlberg, L., Mayer, R. (1972, November). Development as the Aim of Education. *Harvard Educational Review*, 42 (4), 449-496.
- Kuhn, D. (1986). Education for Thinking. *Teachers College Record*, 87 (4), 495-512.
- Kuhn, D. (Ed.). (1990). Developmental Perspectives on Teaching and Learning Thinking Skills. *Contributions to Human Development*, 21, 1-134. Basel: Karger.
- Kuhn, D. (1991). *The Skills of Argument*. New York, NY: Cambridge University Press.
- Kuhn, D. (1997). Constraints or Guideposts? Developmental Psychology and Science Education. *Review of Educational Research*, 67 (1), 141-150.
- Kuhn, D. (1999, March). A Developmental Model of Critical Thinking. *Educational Researcher*, 28 (2), 16-26, 46.
- Kuhn, D. (1999). Metacognitive Development. In Lawrence Balter & Catherine S. Tamis-LeMonda (Eds.), *Child Psychology: A Handbook of Contemporary Issues* (pp. 259-286). New York, NY: Psychology Press.

- Kuhn, D., Garcia-Mila, M., Zohar, A., & Andersen, C. (1995). Strategies of Knowledge Acquisition. *Monographs of the Society for Research in Child Development*, 60 (4), Serial No. 245.
- Kuhn, T. (1970). *The Structure of Scientific Revolutions* (2nd ed.). Chicago: University of Chicago Press.
- Loevinger, J., & Wessler, R. (1970). *Measuring Ego Development*. San Francisco: Jossey-Bass.
- Loevinger, J., & Wessler, R. (1976). *Ego Development: Conceptions and Theories*. San Francisco: Jossey-Bass.
- Nave, B., Miech, EIJ., & Mosteller, F. (1999). *A rare design: The role of field trials in evaluating school practices*. Paper presented at the American Academy of Arts and Sciences, Harvard University.
- Paul, R., *Critical thinking*. Rohnert Park, CA: Center for Critical Thinking and Moral Critique, Sonoma State University.
- Piaget, J. (1926). *The Language and Thought of the Child*. New York: Harcourt Brace.
- Piaget, J. (1928). *Judgment and Reasoning in the Child*. New York: Harcourt Brace.
- Piaget, J. (1965). *The Child's Conception of the World*. New Jersey: Littlefield Adams, and Co.
- Piaget, J. (1970). *Structuralism*. New York: Basic Books.
- Salomon, G. (ed.) (1993). *Distributed Cognitions: Psychological and Educational Considerations*. Cambridge: Cambridge University Press.
- Tishman, S., MacCillivary, D., & Palmer, P. (1999). *Investigating the Educational Impact and Potential of the Museum of Modern Art's Visual Thinking Curriculum: Final Report*. Unpublished manuscript, Harvard Project Zero, Harvard University.
- Vygotsky, L. S. (1978). In M. Cole, V. John-Steiner, S. Scribner & E. Soubberman (Eds.), *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1987). *The Collected Works of L. S. Vygotsky: Problems of General Psychology*. (Vol. 1). Including the volume *Thinking and Speech*. (N. Minick, Trans., Ed.). New York: Plenum.
- Winner, E., & Hetland, L. (2000). The Arts in Education: Evaluating the Evidence for a Causal Link. *The Journal of Aesthetic Education*, 34 (3-4), 3-10.

APPENDIX

Abigail Housen's Stages of Aesthetic Development

Stage I

Accountive viewers are storytellers. Using their senses, memories and personal associations, they make concrete observations about a work of art that are woven into a narrative. Here, judgments are based on what is known and what is liked. Emotions color their comments, as viewers seem to enter the work of art and become part of its unfolding narrative.

Stage II

Constructive viewers set about actively building a framework for looking at works of art, using the most logical and accessible tools: their own perceptions, their knowledge of the natural world, and the values of their social, moral and conventional world. If the work does not look the way it is "supposed to"—if craft, skill, technique, hard work, utility, and function are not evident, or if the subject seems inappropriate—then these viewers judge the work to be "weird," lacking, or of no value. Their sense of what is realistic is the standard often applied to determine value. As emotions begin to go underground, these viewers begin to distance themselves from the work of art.

Stage III

Classifying viewers adopt the analytical and critical stance of the art historian. They want to identify the work as to place, school, style, time and provenance. They decode the work using their library of facts and figures, which they are ready and eager to expand. This viewer believes that properly categorized, the work of art's meaning and message can be explained and rationalized.

Stage IV

Interpretive viewers seek a personal encounter with a work of art. Exploring the work, letting its meaning slowly unfold, they appreciate subtleties of line and shape and color. Now, critical skills are put in the service of feelings and intuitions as these viewers let underlying meanings of the work—what it symbolizes—emerge. Each new encounter with a work of art presents a chance for new comparisons, insights and experiences. Knowing that the work of art's identity and value are subject to reinterpretation, these viewers see their own processes subject to chance and change.

Stage V

Re-creative viewers, having a long history of viewing and reflecting about works of art, now "willingly suspend disbelief." A familiar painting is like an old friend who is known intimately, yet full of surprise, deserving attention on a daily level but also existing on an elevated plane. As in all-important friendships, time is a key ingredient, allowing Stage V viewers to know the ecology of a work—its time, its history, its questions, its travels, its intricacies. Drawing on their own history with one work in particular, and with viewing in general, these viewers combine personal contemplation with views that broadly encompass universal concerns. Here, memory infuses the landscape of the painting, intricately combining the personal and the universal.

In addition to the above five Stages, transitional Stages occur when patterns of thinking from two Stages co-exist. For example, in Stage I/II, idiosyncratic Stage I behaviors occur simultaneously with more object-centered Stage II behaviors.

Significant to understanding Aesthetic Development is that, while growth is related to age, it is not determined by it. In other words, a person of any age with no experience with art will necessarily be in Stage I. An adult will not be at a higher Stage than a child simply by virtue of age or education. Exposure to art over time is the only way to develop aesthetic understanding, and without this time and exposure, Aesthetic Development does not occur.

Over the course of these studies, I have found that most interviewees are beginner viewers, ranging from Stages I to II, or Stage II/III (the transition Stage between Stages II and III). Even among frequent museum-goers, there are relatively few people who have had sufficient interaction with art to have developed beyond the understandings of Stage II/III.

© A. Housen

Acknowledgements

Our deepest appreciation goes to Gwen Amstutz, Diane Jaquith, Kathryn Johnson, the Minneapolis Institute of Arts, Jana Odette, Karen Roos, Wendy Shannon, Anne Suchomel, Kris Wishy, and all of the teachers and parent volunteers in Byron and Kasson. Thank you to the individuals and institutions who provided financial support, including the Andy Warhol Foundation for the Visual Arts, Byron Public Schools, Institute for Museum and Library Services, McKnight Foundation, Minnesota State Arts Board, and the National Endowment for the Arts. And, especially, I want to thank Harry M. Lasker who was there at the beginning, told me to follow my data, and whose voice is always with me.

Author's Notes

- ¹ Visual Understanding in Education is a non-profit educational organization which focuses on research and education in art. For more information, see the VUE web site at <http://www.vue.org/intromat.html>.
- ² The Year I curriculum was comprised of five, forty-five minute lessons; Years II-V were comprised of ten, forty-five minute lessons per year.
- ³ For practical information regarding VTS training and curricular materials, see the VUE web site at <http://www.vue.org/intromat.html>.
- ⁴ For a more complete discussion of aesthetic stage and VTS curriculum, see "Voices of Viewers: Iterative Research, Theory, and Practice" (Housen, 2000-2001).
- ⁵ See beginner viewer behavior described in Stages I & II, Appendix.
- ⁶ Many educational theorists, especially constructivists, learning theorists, and cognitive developmentalists, advocate learning activities which include this

kind of observing, responding and thinking. (Arnheim, 1966, 1969; Bruner, 1960, 1966, 1973, 1996; Dewey, 1910/1997, 1934; Duckworth, 1996; Fosnot, 1996; Kohlberg, 1972; Kuhn, 1990; Vygotsky, 1978, 1987).

⁷ Our Year I pre test subjects numbered 102, and our year V post test subjects numbered 90; during that period we lost 3 control and 9 experimental students.

⁸ We begin the ADIs and MOIs by asking subjects to talk out loud about the different things that go through their heads while looking at a reproduction of a work of art or at a material object, such as what they are looking at and where their eyes are going. In the late 1970's, I began using such "thinking-alouds" as a way to trace the thoughts of viewers while looking at works of art. My initial impetus came from the prevailing interest in, first, the interconnection between language, thought and perception, and, second, new ways of how to gather information which could lead to new understandings of such interconnections. The writings of Rudolf Arnheim, James Mark Baldwin, Jerome Bruner, Noam Chomsky, Lawrence Kohlberg, Jane Loevinger, Jean Piaget, and Lev Vygotsky were essential to my thinking at the time.

⁹ In a sample of cases from this study, inter-rater reliability for coding of ADIs for Aesthetic Stage within a half-stage score was $r = .961$. In a random sample of 233 thought units drawn from MOIs, inter-rater reliability was $r = .855$. For a fuller discussion of coding Aesthetic Stage see Housen, 1983, pp. 96-99; and Housen, 1992, p. 231.

¹⁰ For a more in-depth discussion of the performance of students by grade level see Housen, 2001.

¹¹ The finding is consistent with our other art viewing studies which show a mean rate of growth in Aesthetic Stage in elementary age viewers of at most a half Stage per year. (See Directory of Studies, under VUE Research on the VUE website at: http://www.vue.org/res_directory.html. Also see Housen, 1992, pp. 232-234 for dynamics of change and stage ceilings).

¹² The following year, sixth grade teachers made up for this lapse.

¹³ Bruce Johnson, a Byron 4th grade teacher, noted that the use of VTS produced more descriptive and detailed writing and a more sophisticated vocabulary. Janis Tanner, Byron School Board member, affirmed that "[t]his year's 8th grade students are the first ones to take the Basic Skills test after having been in our VTS arts program. For the past five years, these students have learned critical thinking skills that they are able to then apply to other areas. I attribute some of our increase in reading skills to this program. I feel there

is a direct relationship between our participation in this program and our increase in reading test scores.” Karen Roos, Byron Elementary Principal, stated: “The VTS has allowed students to think at a higher level because [in VTS discussions] we continue to ask the question ‘Why?’ This was transferred to the test, I think, because the test also asks [...] the reasons and rationale ‘Why?’”

¹⁴ The Harvard Project Zero report concluded that having found gains in the VTC assessment activity and the transfer activity, [which] are “... *roughly parallel* [...] *It is possible to view this as evidence of the cross-domain power of the core VTC methodology. It appears to teach an integrated set of skills that cohere naturally in the minds of students.... According to [students] VTC methodology can be used in other areas to help you figure things out, explain things, get information and understand things. [emphasis not added] These comments suggest that students see the VTC questions as addressing thinking and learning challenges that commonly occur in their lives.*” (Tishman, MacCillivary, & Palmer, 1999, p. 64).

¹⁵ For recent studies looking at art education transfer effects, see: Burton, Horowitz & Abeles, 1999; Catterall, 1998; Darby & Catterall 1994; Eisner, 1998 January, 1998 July; Winner & Hetland, 2000.

¹⁶ Much has been written about the prioritizing of random assignment in experimental research designs. Although some researchers claim that the inclusion of random assignment rules out all threats to internal validity, and therefore, is necessary in order to claim causal results, Cronbach writes that “Randomization may be achieved at the expense of relevance. But relevance is surely the sine qua non in evaluation” (Cronbach, 1987, p. 114). For another discussion of this issue see, Kuhn (1970). Cook, also, maintains that “random assignment does not deserve any special privilege since it entails trade-offs not worth making” (Cook, 1999, p. 29). He lists experimental conditions which are not conducive to random assignment: when the treatment is of longer duration; when extensive retraining of teachers is required; when new patterns of coordination among school staff are required; and when communication between students receiving treatment is possible (Cook, 1999). The frequency of these conditions in educational field settings must contribute to the fact that random assignment is estimated to have been used in less than one percent of studies of educational effects (Nave, B., Miech, EIJ., & Mosteller, F., 1999). We did not use random assignment at Byron. Our view is that the prerequisites of random assignment run so counter to the typical organization of a school, that a school willing to accept such requirements is likely to be extremely atypical. Consequently, inferences

drawn from such a school is not likely to be informative or representative of school environments in general.

- ¹⁷ Some theorists contend that transfer is far less likely to be achieved if one does not explicitly teach for transfer (Tishman, et al., 1999). With VTS we avoid this approach, since it could be seen as 'teaching to the test.' Instead, our study is a more open assessment of transfer, without explicitly advocating transfer to students or teachers.
- ¹⁸ This type of phenomenon has been identified by a number of observers of developmental growth and change including Jean Piaget and Jane Loevinger. That said, while one might expect to see critical thinking more frequently displayed, as well as more differentiated, at higher stages of development, Kuhn notes that shifts in reasoning are gradual and overlap with the use of less complex forms of reasoning (Kuhn, 1986, 1999-a). Further, Kuhn has noted that empirical data supporting expectations of higher order thinking occurring more frequently at higher levels of development has not been readily available (Kuhn, 1999-a).

Abigail Housen is founder and principal partner in Visual Understanding in Education (V.U.E.). In the mid 1970s she began to develop a theory and method for understanding aesthetic development which has become a basis for her work in museum programs and school-based curricula. She received her Ed.M. (1976) C.A.S. (1979) and Ed.D. (1983) from Harvard University Graduate School of Education. From 1979-1994 she directed the Graduate Program in Art Education at Massachusetts College of Art where she also was a Professor.

