THE RELATIONSHIP OF TWO WEB PROJECT ACTIVITIES TO MATH LEARNING AN ANALYSIS

Martin F. Gardiner, Ph.D.

I: INTRODUCTION

During the 1998-1999 school year two WEB Project activities were carried out with 7th grade students as part of the curriculum at the North Country Junior High School in Derby, Vt., located within the Orleans Essex School District in Northeastern Vermont in a rural area near the Canadian border.

This report is an analysis of the relationship of these two programs to math learning. The programs were presented at the school by a music teacher, Anne Hamilton, and by a visual arts teacher, Michelle Smyth. The entire research study has been carried out under the direction of Fern Tavalin, Ph.D. A research team consisting of the two teachers, Dr. Tavalin, and myself has met throughout the project to discuss all phases of the work.

We studied both programs. One program, which we will call here the MUSIC program, involved training in music composition with the aid of MIDI computer software. In previous pilot work at another site (Gardiner, 1998; Tavalin et al, 1998) we had observed initial evidence of relationship of performance in a WEB program similar to this to learning in math in 8th grade students in Barre, Vt. A theory of Mental Stretching (Gardiner et al, 1996; Gardiner, 1997,1998) under development building on observations of impact of music training on math learning in younger elementary students (Gardiner et al, 1996; Gardiner and Olson, 1999) can help to account for the Barre data as well. We undertook the present study in part to continue to examine the interrelationship of math learning to the WEB form of MIDI music training.

The second program, which we will call here the ART program, involved students in computer-based animation composition activities. It was thus similar in some ways to the MIDI type of program, though also different in some ways as well. We were interested if there would be any evidence of relationship to math learning similar to what was observed with the MIDI program. This is the first opportunity to study art and music programs such as this with similar analysis.

Research such as this is of interest to us not only for the possible practical value in relationship to math learning, but still more importantly for the insight it may give into more general issues of learning. If the models of Mental Stretching we are studying prove to be correct, they imply broad and overlapping brain involvement in at least some and perhaps many specific types of higher level mental activities and opportunities for cross-fertilization within learning (Tavalin et al, 1998; Gardiner, 1999). The WEB projects are demanding in the learning they require but appear to be engaging for many students of this age, a critical period of learning. These programs allow the students to stretch their capabilities for learning and also carry out creative activity within a domain that many appear to enjoy. Thus we seek to better understand how this opportunity may affect these students performance more broadly, focusing here with the further examination of mathematics because it is most closely related to our earlier work.

II: DESIGN

The school year is divided by the North Country school into quarters. A quarter of the students took the MUSIC program, another quarter the ART program each school quarter. The student groupings were arranged by the school, and then scheduled so that at the end of the 1st quarter the ART and MUSIC students switched. Thus at midyear, two groups of the students had had MUSIC and ART, and two other only started their programming in the second part of the year. Otherwise the students followed similar curriculum. We obtained baseline measures for all students at the beginning of the year, and then comparison measures in January, at the end of Quarter 2. We will call the group that got MUSIC and ART during the period we examined the TREATMENT GROUP (T) and the other half of the students, those who had not yet received the same programs, the CONTROL GROUP(C).

III: DATA

ACADEMIC PROGRESS

Our window on academic progress of each student builds on the nationally normed Stanford Achievement Tests, 7th Edition. These tests were chosen by the school and administered by them at the beginning of the year as their standard instrument to evaluate degree of academic progress of each student as he or she began the year. The results of this testing form our Baseline measures. At the end of January at our request the school administered the next level of Stanford Math testing for comparison. The major scales of this battery are listed in TABLE I.

This analysis will focus upon the major Math sub-scales derived from this testing. One, TOTAL MATH, measures the full spectrum of math performance. This is derived from two other scales: a PROBLEM SOLVING scale (abbreviated here PROBSLV) and a PROCEDURES scale (abbreviated here PROC.).

ASSESSMENT BY ART AND MUSIC TEACHERS

The teachers assessed and reported student performance by several scales:

* Overall Performance in Course (report card grade). This included the results from exams (not tabulated individually here) as well as all the other indicators summarized below. For the analysis four scales were derived from this information MUSIC GRADE (MUSGRD) AND ART GRADE (ARTGRD) report the overall performance as given, including plus and minuses of letter grades. MUSIC GRADEC (MUSGRDC) AND ART GRADEC (ARTGRDC) pool across specific A, B, C, and D categories, suppressing pluses and minuses. This is helpful for some of the statistical analysis where data subdivision is necessary.
* Approach to Work (3 pt scale)(Abbreviated here MUSWRK and ARTWRK respectively).

3 = (exceeds expectations)

on task and focused
makes decisions based on knowledge and skills
tries experiments

2 = (meets expectations)

usually on task
tries some experiments
decisions usually based on knowledge and skills

1 = ( )

rarely on task
contributes little
decisions rarely based on knowledge and concepts

* Respect for Others and Equipment (3 pt scale).(This is abbreviated here as MUSRSPINT and ARSRESPINT respectively).

3 = (exceeds)

always respects others' ideas
compromises and listens to others' ideas
treats equipment properly
interacts respectfully

2 = (meets)

usually respects others' ideas
compromises and listens to others' ideas
usually treats equipment properly
usually interacts respectfully

1 = ( )

* Quality of Product (3 pt scale)(Abbreviated here MUSPRD and ARTPRD respectively).

3 = (exceeds)

Composition

makes sense - uses patterns
treble and bass clef go together
proper notation

no wild skips
goes beyond expectation

Animation

smooth movement throughout
effective use of different camera shots
understandable and interesting storyline
clean or neat drawings

2 =

Composition

uses some patterns
may have some clashing notes
makes sense
meets expectations; completes assignment
proper notation

Animation

a few jumps or skips
uses different camera shots
understandable storyline

mostly neat and clean drawings

1 =

Composition

bass and treble clef don't go together
improper notation
doesn't make sense

Animation

many skips or jumps
no transitions or change in camera shots
confusing storyline
messy, sloppy drawings

* Reflection and Critique (3 pt scale) (This was only reported for music and is abbreviated here (MUSRFLCRT).

3 =

accurately describes the area being discussed
gives detailed examples, references, connections or responses to general insights
uses arts vocabulary

2 =

accurately describes the area being discussed
uses mix of arts vocabulary and general terms

1 =

gives general comments that could apply to other situations as well as the discussion

* Relationship between Articulation and Execution

3 = can articulate and execute concepts

2 = can articulate only

1 = can execute only

9 = cannot determine

IV: ANALYSIS

A: MISSING DATA

There were quite a few students in the school (approximately 40%) who either did not take or did not complete the January Testing. Some had left the school, some were Special Needs students who were either not asked to take the test or must be analyzed separately from the other students. Some for whatever reason did not finish their work on one or both sub-tests within the time limit.

There remained about 100 students with complete data that could be analyzed here: almost identical numbers in the T and C groups.

Some data is available from students with incomplete records, but will not be included here since the reasons for this incompleteness cannot be explored adequately at this time.

B: BASELINE ACADEMIC SCORES AND GENERAL GAINS IN MATH

(INSERT TABLE I AT OR NEAR THIS POINT)

As Table I shows, the average percentile scores for both the Controls and Treatment students for whom we have complete data were similar, though not identical. Though the differences did not reach statistical significance by T Statistic on any of the twenty major scales, the Treatment students entered the study somewhat stronger than Controls in almost every category, including math and science. The 50th percentile is the National Average for the standardization sample. By this measure, the students as a whole were slightly or somewhat below National Average in most areas, though again this was less true for Treatment students coming into the study than for Controls. This asymmetry between the groups complicates the analysis, as we will discuss. Note that still weaker students may not be included here because of problems with the January testing discussed earlier.

By January the relatively weak performance in math had come up considerably in both groups included here: to improve this performance may have been a goal of the school, but it is also possible that the exclusion of many students from testing or eliminating them due to incompleteness of their testing results in math is excluding the students who did not improve as much by January, and thus is seriously biasing the sample. This general gain in math again complicates the analysis, as we will discuss.

C: RELATIONSHIP OF ART AND MUSIC GRADES TO BASELINE ACADEMIC PERFORMANCE

(INSERT TABLE II AT OR NEAR THIS POINT)

As a conservative approach, Spearman Correlation Coefficients rather than the more usual Pearson Correlation Coefficients were used here to measure this relationship as they are based on rank order methods, and thus are less dependent on assumptions of normal distributions which are likely to be violated both to a known and also to an unknown degree by these data. The results of this analysis are given in Table II.

It is clear that at least in these programs the academically stronger students coming in tended to do better in both the art and music than the academically weaker students. The relationships of program grade to both music and art were typically highly significant, though not entirely identical. Thus the relationship of grade to baseline performance was stronger for music than art in most areas, including math, but interestingly almost identical for physical science.

The Mental Stretching theory (Gardiner, 1998, Gardiner, 1999) is focusing attention on specific types of mental capabilities that when built, developed, or 'stretched' in response to one type of learning need (for example to make progress in music), become increasingly available to facilitate learning in another area (e.g. Math) where a similar type of thinking may be needed. To the extent that such cross-fertilization occurs, and there is now some initial evidence for it (e.g. Gardiner et al, 1996; Gardiner and Olson, 1999; Tavalin et al, 1998), there is nothing in the theory to rule out cross-fertilization in either or both directions of application. Thus from the point of view of this theory it is plausible that previous academic 'stretching' influences performance in an arts program, and this may be especially true here where the programs integrate aesthetic and also practical considerations involving computer, and equipment.

D: RELATIVE GAINS IN MATH FOR TREATMENT AND CONTROLS : A RELATIONSHIP TO GRADE IN MUSIC

(INSERT TABLES IIIA-IIIC AT OR NEAR HERE)

To help address the asymmetries between the T and C groups noted earlier the data were analyzed in strata defined on the basis of the total math score at baseline. As can be seen from Table IIIA this helps to equalize the baseline conditions. Gains in Treatment and Control groups were compared using the T statistic, again a robust method insensitive to any violations from normality in these data. It can be seen that as already anticipated by Table I, the gains in both groups are quite similar, and not significantly different. This is true for the data as a whole, and also true when the data are subdivided by gender (Tables IIIB and IIIC).

That analysis, however, does not take into account performance in the art program. If a student receives an A in Music, for example, he or she has had to be performing differently in the course in some way, compared to a B student, and similarly for B vs C students. Thus if there is to be a relationship between the arts activity and gains in another area such as math, we would anticipate that it would be related to performance as measured by overall grade.

Tables IIIA-IIIC. shows that for the Music Program this expectation was indeed supported. Interestingly, the strong relationship is for the weakest students in math going into the program. In spite of the small sample sizes caused by the subdivisions by stratification and then further by grade, the gains with an A are significantly higher than for controls, and higher than for Bs, which in turn are higher than for Cs. Interestingly, in this group of students whose math performance was weakest going into the study, almost a third of these achieved As, and of the rest half received Bs. When genders are compared the effect in males is stronger than in Females, and only in males reaches statistical significance, though in Females the trends are similar and might well reach significance with a larger sample size. Note also that the proportion of males in this stratum who get As in Music is higher than for the females.

E: PARTIAL CORRELATION ANALYSIS OF RELATIONSHIP OF WEB PROGRAM PERFORMANCES TO GAINS IN MATH

(PLACE TABLES IVA-IVC AT OR NEAR HERE)

The results in the previous section have been explored more fully and powerfully with a Partial Correlation Analysis. This not only allows the correlation between measures of student performance in the WEB programs to be examined, but also adjustment for baseline differences to be carried out more powerfully than with the stratification analysis of the previous section. For the analysis now discussed, correlations have been corrected for baseline differences in both Math Procedures and Math Problem Solving, rather than relying on the total math score. And the analysis sets are much larger because the attrition due to stratification and subdivision is avoided.

It should be emphasized that this analysis applies only to the T group students. A comparison with controls by a similar method was also undertaken, as will be discussed in the next section.

For reasons discussed earlier, partial Spearman rather than the more usual partial Pearson correlations derived in standard Regression Analysis are examined here.

Table IVA shows the results of this analysis for all measures of music and art performance outlined earlier, and also for several further scales that reflect total performance in both programs. These new scales are derived as the sums of the other scales: since Spearman uses rank orders, more complicated methods to preserve scaling are not needed. Thus ArMusGrd sums the grades in Art and Music: the highest performance is with A in both, etc.

It is shown in this analysis that in these data within the T group Music Grade has a very powerful relationship to gains in Math Problem solving, but very much less relationship to gains in Math Procedures. From the last two columns we can see that changes in Math Grade is in turn related to at least some degree to all the other measures of Music performance examined here, but most weakly to the Music Product score. We will return to this below. Grade in Visual Arts shows weaker relationship though similar trend.

Very interesting strong predictors of gains in math turn out to be level of engagement in the activities. This is true for Music and also for Art. Interestingly, the quality of the product resulting has a much weaker relationship to math gains. This suggests that it may be the struggle rather than the accomplishment that may be more closely related to interactions with math. This appears to deserve further exploration. Indeed strong engagement in both music and art is an especially strong correlate of progress in math.

The importance of Reflection and Critique and also factors relating to cooperation and respect in these data are also of interest. There is support here for the notion that the enrichment of the learning setting where the student is encouraged to reflect, not simply to do, and the enrichment that can come from social interaction can both be related to the broader impact.

There are some gender differences, the males sometimes showing stronger effects than the females, but sometimes the females are stronger, for example in the importance of engagement in both music and art programs, there are often trends in the same direction but some differences. With this initial sample size it appears premature to draw strong conclusions about gender differences, but they seem worth pursuing further.

The emphasis on relationship to Math Problem Solving rather than computational capability per se is a strong general theme for many of the variables. This has implications for the developing Mental Stretching modeling and also for the evaluation of the impact of programs such as this.

F: MATCHING SAMPLE TEST AGAINST CONTROLS

(INSERT TABLE V AT THIS POINT OR NEARBY)

Table V presents further evidence that the calculations of the previous section do not in themselves create a spurious correlation somehow based upon the initial conditions uncorrected by the analysis corrections that are used, but rather are reflecting something real, i.e. something truly related to the experimental conditions, not a statistical artifact. With complex analyses, protection by what this author has previously termed a Randomization Test can help. The present calculation is an example of a Randomization test.

For this analysis all pairs of subjects were used that could be constructed, one from T group, the other from C group, where both Procedures and Problem Solving scores at baseline were similar to each other: specifically within 15 percentile points on both variables. These subjects were then linked in the following way. If a subject in the T group got an A in Music, the linked C group subject was assigned an A on a dummy variable called Pseudo Music Score. If Subject in T got a B, linked subject in C was assigned a Pseudo B, etc. If the gain in math is somehow a spurious artifact of method, then we should find as strong or at least a similar relationship of Pseudo Music Grade in Controls to math gain as is seen for grade in Music to math in the T group. But as shown in Table V, this does not happen, supporting that the results are not due to the statistics, but rather the actual situation.

G: GAINS IN MATH THINKING SKILLS

(INSERT TABLES VIA-C AT THIS LOCATION OR NEARBY)

Section E above examined gains in overall math performance in the two major sub-tests of the Stanford 9 Achievement Test, and in the total math capability score derived from these two sub-tests. In this section we will begin a further examination of math gains, now breaking them down in finer detail with the aid of further sub-scales derived from the Stanford 9.

As in Section E we will examine the relationship of gains in math capabilities to performance in the music and art WEB activities as assessed by the variables defined earlier. We will again use Spearman Partial Correlation Analysis to measure the relationships after correcting and adjusting for baseline math performance on both the Problem Solving and Procedures Math sub-tests of the Stanford 9.

In this section we will focus on gains in overall thinking skills required for and displayed by math capability at this level. The following math thinking scales were developed within the Stanford 9 as measures of Instructional Objectives, computed from math items.

1. Scales based on definitions of the National Council of Teachers of Mathematics (NCTM) (The publisher's definitions are quoted):
a) Mathematics as Problem Solving(MPS): 'Demonstrates the ability to apply strategies to solve a wide variety of problems'.

b)Mathematics as Reasoning(MR): 'Demonstrates the ability to apply inductive, deductive or spatial reasoning and make valid conjectures and draw valid conclusions'.

c)Mathematics as Communication(MComm): 'Demonstrates an understanding of the signs, symbols and terms of mathematics, and correctly interpret alternate representation of data'.

d)Mathematics as Connections(MConn): 'Demonstrates an understanding of the inter-relatedness of mathematical concepts, procedures and processes both among different mathematical topics and with other content areas'.

2. Measures of overall 'Thinking Skills': As distinguished from items testing for knowledge of basic math facts and information, according to the publisher 'items classified as (requiring) Thinking Skills measure the ability to analyze and synthesize information; to evaluate information in order to determine cause and effect, fact and opinion, relevant and irrelevant; and to interpolate and / or extrapolate beyond information in order to draw conclusions, make predictions, and hypothesize'.

Two sub-scales were examined made up of Thinking Skill items.

e) Total Score on 'Thinking Skill' items on the Math Concepts and Problem Solving Stanford 9 Test(ThMPS)

f)Total Score on 'Thinking Skill' items on the Math Procedures Stanford 9 Test(ThMProc)

The results for males and females combined are given in Table VIA. It can be seen that there are significant positive correlations between many measures of engagement in the music activity and gains on the NCTM scale of problem solving capability. The correlations are still stronger with the total set of Thinking Skill items on the Concepts and Problem Solving test now also reaching significance for several items that concern total performance on music and art activities. It should be noted that the quality of Musical Product item (MusPr) shows a stronger relationship to measures of problem solving and thinking than seen for the math scores as a whole (Table IV), reaching significance for the ThMPS scale. Thinking skills displayed in items drawn from the Math Procedures test (ThMProc) shown a pattern of relationship similar to that shown on Thinking Skill items drawn from the Concepts and Problem Solving test, but the correlations are weaker. In part this could be due to the fact that fewer Thinking Skill items are available from this sub-test, 15 as compared to 27 from Concepts and Problem Solving.

The results from Males alone are given in Table VIB, and Females alone in Table VIC. Some differences are apparent. Females show a pattern of correlations similar to that seen for Males and Females combined, but still stronger correlations with gains in Problem Solving and thinking skills than seen in the combined data set. Males, on the other hand, show less emphasis on gains in problem solving, and more on gains in Reasoning (the MR scale), something not at all apparent in the data from the females.

H: GAINS IN MATH CONCEPTS AND PROBLEM SOLVING BY TOPIC AREA

(INSERT TABLES VIIA-C and VIIA-C AT THIS LOCATION OR NEARBY)

We used further sub-scales to examine gains in math broken down by topic area. These were again scales made available by the test publisher. As there were relatively few items used on most of the sub-scales now discussed here, the results must be viewed with still greater caution than the results discussed up to this point. The relationship of gains in these math sub-scales to WEB art and music variables was again examined by partial correlation analysis as previously.

The following sub-scales were examined:

a) Measurement (Msmt): ‘Demonstrates an understanding of the principles of measurement’.

b) Estimation (Est): Determine the reasonableness of results and apply estimation in problem solving’.

c) Problem Solving Strategies (PSSrat): ‘Demonstrates an understanding of the process of solving conventional and non-routine problems’.

d) Number and Number Relationships (NumRel): ‘Represent and use numbers in equivalent forms in real-world and mathematical problems, and demonstrate number sense for whole numbers, fractions, decimals and integers’.

e) Patterns and Functions (PatFn) : ‘Identify numerical and geometric patterns and functions and use them to solve problems’.

g) Algebra (Algb) : ‘Demonstrate the ability to evaluate expressions and solve linear equations’.

h) Statistics (Stats): ‘Demonstrate an understanding of the relationships in data sets’.

i) Probability (Prob): ‘Demonstrate an understanding of the laws governing chance’.

j) Geometry (Geom): ‘Demonstrate an understanding of geometric principles’

Tables VIIA and VIIIA show the results of these calculations for males and females combined. Music grade and other aspects of music performance show some evidence of weak relationship to improvement in scales for Measurement, Estimation, Strategies for Problem Solving and stronger relationship to Patterns and Functions and capability in Probability, including a relationship to Music product which approaches significance. There is also evidence of negative correlations with the performance of the Number Systems and Number Theory items, indicating that gains in this sub-scale tended to take place in individuals who were weakest in the specific measures of arts performance indicated. We will discuss these negative correlations further later in the paper.

Tables VIIB,C and VIIIB,C examine correlations separately for males and females. There are some striking differences. The males show very strong relationship of music and total art music performance to gains in Patterns and Functions, Algebra and a significant correlation of assessment of Music Product to gains on the Probability scale. The females show a somewhat more evenly distributed patterning across topics, with again some evidence of relationship to gains in Probability, reaching significance of relationship to Art Grade.

V: DISCUSSION AND CONCLUSIONS

The statistical methods applied here have proved useful with the analysis of a small data set that presents difficulties that were discussed earlier. Such data sets are however quite representative of data available from relatively small studies such as this. Larger and better controlled data sets typically involve larger and much more costly studies, but results from smaller studies such as this can help to justify and if appropriate plan larger studies. The methods demonstrated here might well prove to be more generally useful perhaps especially for small studies such as this one. We are not aware of other examples of analyses similar to this in the literature. We plan to continue to explore these and related methods in our continuing work. The heavy use of ranking rather than averaging procedure are reminiscent of other powerful methods e.g. CART which is finding very interesting application (e.g. Walter et al, 1999).

In spite of the limitations of the data, our analysis provides fresh evidence of relationship between performance in these WEB project activities and gains in performance in academic areas not specific to the WEB project. In this work we have focused on detailed analysis of gains in math capability. The gains have been shown to most significantly involve improvements in general problem solving capabilities that go beyond specific content areas. There is also evidence that arts performance of the students was related to more general academic capabilities developed by the students before they entered these programs: our analysis has had to take into account these baseline effects and adjust the analysis of change accordingly.

Our results continue to support that there is cross fertilization of developed capability and learning in both directions between different areas of learning. The statistical methods have been helpful here to begin to unscramble this dialogue. The richness and complexity of the inter-relational topologies indicated seem difficult to account for in a single factor of capability, and also help support a model of at least some causality from the WEB-Arts related actives to more general mental skills which aid performance in Math. It is difficult to see how to account for the complex interrelationships shown here in ways that avoid some dialogue or cross fertilization of learning within the brain.

This study has focused upon gains in math for reasons discussed earlier. Very recent data in early elementary students (Gardiner and Olson, 1999) now suggests that gains in reading may well be related to different paths of mental stretching than gains in math. This is something we hope to also examine further with analyses such as this with regard to the WEB project.

These results reinforce the accumulating evidence that WEB project activities such as this can relate in useful ways to more general learning, and support the idea that strong engagement of the students in the WEB work and in working with each other are both important part of the possible more general value derived from these projects.

The differences between the boys and girls in these data would appear to be worthy of further investigation. In these data, for example, the gains in performance of the girl appear to involve more general problem solving skills, whereas in the boys there is more specific gains with regard to reasoning. This may point to differences in the ways individuals approach and learn from WEB tasks, focusing on different aspects of the WEB tasks, or looking at the tasks in different ways. It would be useful in continuing studies to broaden the range outcome variables beyond the math variables on which this study focused. In particular, changes in capabilities involving social behavior which in turn can at least in early elementary students influence academic learning by a different pathway of influence than that associated with math (Gardiner and Olson, 1999) deserve further investigation.

How to more clearly evaluate the importance and role of the aesthetic features of the projects remains a goal of continuing work. At the very least, it seems likely that the aesthetic features help to motivate student involvement, but the math gains are similar to ones seen in music programs without these technological features with 1st and second graders and where there is evidence of impact from the music training in part due to differential effects on different aspects of learning (Gardiner et al, 1996; Gardiner and Olson, 1999). Further studies of these or other students in the WEB project that examine a broader range of outcomes and with tools that study the arts activities in still greater detail can help to clarify this important question.