MIND, BRAIN, AND EDUCATION

The Effects of Arts Integration on Long-Term Retention of Academic Content

Mariale Hardiman[1], Luke Rinne[1], and Julia Yarmolinskaya[1]

ABSTRACT— Previous correlational and quasi-experimental studies of arts integration—the pedagogical practice of "teaching through the arts"—suggest its value for enhancing cognitive, academic, and social skills. This study reports the results of a small, preliminary classroom-based experiment that tested effects of arts integration on long-term retention of content. We designed matched arts-integrated (AI) and conventional science units in astronomy and ecology. Four randomized groups of 5th graders in one school completed one unit in the treatment (AI) condition and the other in the control (conventional) condition. To control for teacher effects, four teachers taught the same subject to different groups in each condition. We administered curriculum-based assessments before, immediately after, and 2 months after each unit to measure initial learning and retention. Results showed no differences in initial learning, but significantly better retention in the AI condition. Increases in retention were greatest for students at the lowest levels of reading achievement.

Scarce resources in public education and a focus on the "basics" have led to the well-documented narrowing of the curriculum, often resulting in a reduced role for the arts in schools (Mishook & Kornhaber, 2006). However, in-school artistic activity may be valuable not only "for arts' sake" but also because the arts may improve learning and student outcomes more broadly. A number of researchers have proposed that knowledge and skills gained uniquely through the arts correlate with success in other academic domains (Catterall, 2002; Deasy, 2002; Fiske, 1999; Spelke, 2008; Wandall, Dougherty, Ben-Shachar, Deutsch, & Tsang, 2008). Still others argue that the arts contribute to the development of more general thinking skills and dispositions that benefit performance (Hetland, Winner, Veenema, & Sheridan, 2007). Arts integration, defined here as the infusion of visual and performing arts activities into instruction in nonarts subjects, is also thought to enhance content learning (Burnaford, Brown, Doherty, & McLaughlin, 2007; Martin et al., 2013). Several quasi-experimental studies reported significant differences in academic outcomes in schools that instituted arts-integrated curriculum compared to control schools (Barry, 2010; Catterall, Dumais, & Hampden-Thompson, 2012; Phillips, Harper, Lee, & Boone, 2013; Scripp, Burnaford, Vazquez, Paradis, & Sienkiewicz, 2013). Yet, to our knowledge, no randomized controlled trials have tested the effectiveness of arts-integrated (AI) curricula. Such experiments are important because schools and/or teachers generally self-select AI instruction, raising the possibility that differences in student outcomes may be the result of selection bias. In this article, we report the results of a small, preliminary randomized experiment that investigated the effects of AI fifth-grade science units on long-term retention of content.

[1]Johns Hopkins University

Address correspondence to Mariale Hardiman, School of Education, Johns Hopkins University, 2800 N. Charles St., Baltimore, MD 21218; e-mail: mmhardiman@jhu.edu

Arts Integration and Long-Term Retention of Content

On the basis of the idea that rehearsal of information consolidates memories for long-term storage (Kandel, 2006), Hardiman (2003, 2010) posited that AI instruction improves retention by prompting students to rehearse content through the use of various visual and performing arts activities which may enhance student engagement (e.g., Smithrim & Upitis, 2005). Further, Rinne, Gregory, Yarmolinskaya, and Hardiman (2011) argue that the arts may engage learners in thinking about new information in ways that improve retention, for example through semantic elaboration (e.g., Craik & Tulving, 1975), generation of information from a cue (e.g., Slamecka & Graf, 1978), enactment (e.g., Mohr, Engelkamp, & Zimmer, 1989), oral production (MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010), "effort after meaning" (e.g., Zaromb &

Roediger, 2009), emotional arousal (e.g., Cahill & McGaugh, 1995), and pictorial representation (e.g., Paivio, 1971). The theory that arts integration improves retention of content is derived from the notion that the arts naturally take advantage of these strategies whereas conventional instruction typically does not.

This study contributes to the growing body of research on arts integration by testing the effects of AI curricula on academic outcomes through a preliminary study using a randomized experimental design. We developed matched AI and conventional versions of 5th grade science units in ecology and astronomy and tested their effects on retention of content in four fifth-grade classrooms. We hypothesized that the AI units would lead to better retention of content on a delayed posttest. We did not expect that the arts-focused curricula would facilitate the learning of more content over the course of the unit—rather, the prediction was that more of what students had learned would be retained.

METHOD

Participants

Our study was conducted at an urban elementary school that serves over 600 students (98.5% African American), largely from low-income families (83.5% are eligible for free and reduced-price meals [FARM]). The faculty (65.8% African American), are mostly classified as "highly qualified teachers" (85.3%) according to the school district. All participants in our study were African American (47% female; $M_{age} = 10.8$), and 14% were identified as students with disabilities (eligible for an individualized education plan [IEP]). All students in the fifth grade (97) at the school agreed to participate in the study.

Materials

We developed two science units in astronomy and ecology that spanned 3 weeks of instruction. For each unit, we designed a control version that utilized conventional forms of instruction and a treatment, arts-integrated (AI) version that incorporated a variety of forms of artistic activity (music, visual arts, and performance arts) into instruction. (Interested educators can contact us concerning access to these lessons.)

We first mapped the content to be taught over the course of both units and segmented this content into 15 days of instruction. We then designed conventional lesson plans for each day based on the 5E learning cycle model (Bybee et al., 1989): *Engage*, *Explore*, *Explain*, *Elaborate*, and *Evaluate*. One conventional instructional activity was designed for each of the five E's, and activity durations were estimated. We then compiled the AI units by designing and substituting in AI learning activities for approximately two-thirds of the standard activities. Each AI activity was designed to address the same science content as the standard activity it replaced and to take the same amount of time as its matched non-AI activity. In designing the AI activities, we controlled for a variety of other factors that could potentially influence retention. These included: the order in which information was presented or addressed, the "structure of the activity" (e.g., individual, pairs, small groups, or whole class), and the use of multimedia. In addition, we controlled for whether the activities were student-led or teacher-led and whether the activities produced student work, with the aim of ensuring that any student learning effects observed would not be simply attributable to "active learning" (e.g., Bonwell & Eison, 1991). Both AI and conventional curricula were designed to be active and engaging, differing only with respect to the integration of arts-based activities in place of conventional activities.

Arts-integrated substitutions were designed as follows:

1. Worksheets with written responses were replaced with activities in which students drew responses to the same prompts. This substitution was made to take advantage of the "picture superiority" memory effect (e.g., Paivio, 1971), in which pictures tend to be remembered better than words, even when verbal tests are administered.
2. Instead of silently reading informational texts, students in the AI condition silently read *stories* containing the same content. Here we aimed to improve recall by adding elaborative language (e.g., Craik & Tulving, 1975) and eliciting emotional responses (e.g., Cahill & McGaugh, 1995).
3. Instead of reading text passages out loud, we designed dramatic scripts that children were asked to perform.
4. Instead of verbal group presentations, students performed sketches or tableaux (two forms of dramatic play acting) that we hoped would take advantage of the enactment effect, which posits that carrying out an action produces better memory than simply reading or hearing about the action (Mohr et al., 1989).
5. For vocabulary development, in control lessons students wrote "elaborations" for each term by expanding on the definition or providing an example. In the AI condition, students created "doodles" (simple drawings) to elaborate the meaning of the new term, which we hoped would leverage the picture superiority effect (e.g., Paivio, 1971).
6. At the conclusion of each unit, one additional day (Day 16) was dedicated to review. We selected 10 matched pairs of activities from across each unit and instructed teachers to have students present the work they completed during the activity. In the AI unit, this day was framed as an "art exhibition day," and in the control unit, it was simply a "review day."

For examples, in one astronomy activity in the control condition, students traced the shapes of galaxies, wrote descriptions in notebooks, and shared them with the class.

In the matched AI activity, students drew shapes on posters and used dance movements to depict galaxy shapes in small performances. In another activity, students in the control condition described different planets' features by making Venn diagrams; in the matched AI activity, students drew advertising posters to sell real estate on different planets.

Measures

We developed multiple-choice curriculum-based assessments to measure student learning and retention of the science content. Assessments consisted of 25 multiple-choice questions as well as one "brief constructed response" (BCR short-answer) item. We created three versions of each curriculum-based assessment by changing question wording and/or modifying answer choices. These measures were deliberately designed to be difficult to avoid ceiling effects and make retention challenging. The order in which the three different forms were administered for the pretest, posttest, and delayed posttests was counterbalanced across participants.

Procedures

Prior to implementation, teachers received approximately 10 hr of training to review the content and activities in each unit. To mitigate potential bias, teachers were not informed of our theory regarding how arts integration might improve retention. Instead, they were informed only that the study aimed to test the effectiveness of arts-integration science curricula versus conventional curricula.

All fifth grade students in the school were randomly assigned to one of four equally sized groups for 1 hr per day of science instruction. To ensure that all participants received equivalent potential benefits from the treatment, students were taught one unit in the AI condition and the other in the control condition. After completing the first unit, each group then completed the second unit with a different teacher in the alternate condition (see Table 1). Pretests were administered the day before each unit began, posttests at the conclusion of each unit, and delayed posttests approximately 8 weeks after the posttest.

Each teacher taught one randomized group of students in one condition (treatment or control) and then taught the matched unit in the other condition to a second randomized group. Thus, each teacher taught one content area in both conditions, and the order of conditions was counterbalanced across teachers. This design allowed for the control of teacher effects. Fidelity-of-implementation observers were present in each classroom approximately 60% of the time to ensure that the units were taught as designed in terms of timing, content, and activities. Observers were instructed to note any deviation from the curriculum as well as any differences in instructional quality across conditions.

Table 1
Randomized Group Assignments

		Teacher 1 Astronomy	*Teacher 2 Astronomy*	*Teacher 3 Ecology*	*Teacher 4 Ecology*
Period 1	**Group**	A	B	C	D
	Condition	AI	Control	Control	AI
Period 2	**Group**	D	C	B	A
	Condition	Control	AI	AI	Control

RESULTS

Our analysis of student outcomes includes 82 students (of 97) who completed all assessments and for whom we received demographic information and standardized test scores from the district. We calculated each student's percent correct on each test, measuring *initial learning* in terms of the difference between the initial posttest (T_2) and pretest (T_1) scores (i.e., $T_2 - T_1$). Meanwhile, consistent with previous studies (Custers, 2010), *retention* equaled the delayed posttest (T_3) score as a proportion of the T_2 score (T_3/T_2). We identified three outlying high retention scores (two treatments, one control) whose scores fell more than three standard deviations above the mean. These were winsorized to the next highest value. Students completed the BCR (short-answer) item at the end of each test at a very low rate (<50%), so these responses were not analyzed.

We conducted repeated-measures ANOVAs of initial learning and retention. Factors in each model included condition order (AI vs. Control first), AI subject (Astronomy vs. Ecology), gender, FARM eligibility, IEP eligibility, and reading and math proficiency levels (Proficient/Advanced vs. Basic). To conserve statistical power given our relatively small sample size, factors that did not exhibit significant main effects or interactions with condition (AI vs. Control) were removed from our models.

None of the factors listed above was a significant predictor of initial learning, nor was there a significant effect of the treatment on initial learning (AI vs. Control), $F(1, 82) = 1.496$, $p = .225$. However, our analysis of retention revealed a significant main effect of AI, $F(1, 80) = 6.570$, $p = .012$, $\eta_p^2 = .076$, as well as a significant interaction between AI and reading proficiency level $F(1, 80) = 7.070$, $p = .009$, $\eta_p^2 = .081$. Mean retention was approximately .2 *SD* greater in the AI condition ($M = .912$) than in the control condition ($M = .866$)—an effect size that is "small" based on Cohen's (1988) benchmarks, but relatively typical of effective educational interventions (Hill, Bloom, Black, & Lipsey, 2008). The significant interaction indicates that students at a "basic" level of reading proficiency ($n = 16$) were the primary drivers of the observed main effect. Mean retention among these students was approximately .9 *SD* greater in the AI condition ($M = .976$) than the control

Table 2
Mean Test Scores (Percent Correct) by Test Time, Condition (AI vs. Control), and Reading Proficiency Level

Condition	*Reading proficiency level*	*Pretest*	*Posttest*	*Delayed posttest*	*Retention rate*
AI	Basic	29.0	44.0	39.0	97.6%
	Proficient/advanced	38.6	59.0	52.8	89.7%
Control	Basic	25.0	46.3	32.5	72.0%
	Proficient/advanced	38.9	61.3	54.1	90.2%

condition ($M = .720$). Interestingly, in the AI condition, poor readers exhibited *better* retention than their proficiently reading peers, though it is important to note that these students' long-term learning gains were smaller by about 4 percentage points, meaning they acquired less knowledge to begin with. See Table 2 for mean scores across the three test times.

DISCUSSION

Although there was no significant effect of arts integration on initial learning, arts integration curricula had a significant effect on retention. Participants learned roughly the same volume of science content regardless of how the units were taught, but scores on a delayed posttest showed that students retained what they learned significantly better when taught through AI instruction. These findings provide preliminary support for the theory that arts integration naturally leads students to interact with academic content in ways that promote long-term retention.

An important caveat to this finding is that it was the group of students at a "basic" level of proficiency who showed the biggest gains in retention. In many ways, this makes sense—AI instruction relies less on reading and writing than does conventional instruction, so students who struggle with reading and writing may benefit more from the opportunity to learn content through alternative, arts-based means. Thus, for students who are still developing literacy skills, arts integration may represent a useful means of improving learning. We emphasize that we are not endorsing reductions in reading and writing. If reading and writing activities were scaled back as part of AI instruction, the possibility of deleterious effects on reading and writing would need to be addressed, and educators would presumably need to compensate for this loss elsewhere in the curriculum.

The generalizability of our findings is limited by our relatively small sample of participants of the same age from a single school. In particular, our sample included 16 basic-level readers. Future research should attempt to replicate our results with a larger sample. Additionally, it is possible that results were biased by some teachers' natural enthusiasm for the use of the arts in instruction. Further, although teachers were told that the study aimed to compare the effectiveness of AI and conventional curricula, the fact that we were interested in this comparison suggests our hypothesis that the AI curricula may lead to better outcomes. However, the presence of observers in the classroom during a majority of instructional time helped to ensure that the lessons were taught as written.

Despite these limitations, we believe our study provides important preliminary evidence that many students, particularly struggling readers, may retain academic content better when instruction is integrated with the arts, particularly when mastery requires possession of considerable amounts of declarative knowledge. We hope that in the future, larger scale studies of arts integration that utilize randomized experimental designs will be conducted, and that these studies would be able to disentangle the underlying causes of the effects. Such efforts would provide for more effective measurement of the potential benefits of AI instruction, which may eventually lead to findings that can enhance educational practice and policy in important ways.

Acknowledgment—The authors would like to thank the Joseph P. Drown Foundation for support of this study.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix S1. Sample arts integrated/control lesson plans.

REFERENCES

Barry, N. H. (2010). *Oklahoma A+ Schools: What the research tells us 2002–2007. Volume three, quantitative measures*. Oklahoma A + Schools/University of Central Oklahoma. Retrieved from http://www.okaplus.org/storage/V3%20final.pdf

Bonwell, C. C., & Eison, J. A. (1991). *Active learning: Creating excitement in the classroom* (ASHE-ERIC Higher Education Report No. 1). Washington, DC: ERIC Clearinghouse on Higher Education, The George Washington University.

Burnaford, G., Brown, S., Doherty, J., & McLaughlin, H. J. (2007). *Arts integration frameworks, research and practice*. Washington, DC: Arts Education Partnership.

Bybee, R. W., Buchwald, C. E., Crissman, S., Heil, D. R., Kuerbis, P. J., Matsumoto, C., & McInerney, J. D. (1989). *Science and technology education for the elementary years: Frameworks for curriculum and instruction*. Washington, DC: The National Center for Improving Instruction.

Cahill, L., & McGaugh, J. L. (1995). A novel demonstration of enhanced memory associated with emotional arousal. *Consciousness and Cognition*, *4*, 410–421. doi:10.1006/ccog.1995.1048

Catterall, J. S. (2002). The arts and the transfer of learning. In R. J. Deasy (Ed.), *Critical links: Learning in the arts and student academic and social development* (pp. 151–157). Washington, DC: Arts Education Partnership.

Catterall, J. S., Dumais, S. A., & Hampden-Thompson, G. (2012). *The arts and achievement in at-risk youth: Findings from four longitudinal studies*. Washington, DC: National Endowment for the Arts. Retrieved http://arts.gov/file/2684

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268–294. doi:10.1037/0096-3445.104.3.268

Custers, E. (2010). Long-term retention of basic science knowledge: A review study. *Advances in Health Sciences Education*, *15*, 109–128.

Deasy, R. J. (Ed.). (2002). *Critical links: Learning in the arts and student academic and social development*. Washington, DC: Arts Education Partnership.

Fiske, E. (Ed.) (1999). *Champions of change: The impact of the arts on learning*. Washington, DC: Arts Education Partnership and President's Committee on the Arts and Humanities.

Hardiman, M. M. (2003). *Connecting brain research with effective teaching: The brain-targeted teaching model*. Lanham, MD: Rowman & Littlefield Education.

Hardiman, M. M. (2010). The creative-artistic brain. In D. Sousa (Ed.), *Mind, brain, and education: Neuroscience implications for the classroom* (pp. 226–246). Bloomington, IN: Solution Tree Press.

Hetland, L., Winner, E., Veenema, S., & Sheridan, K. (2007). *Studio thinking: The real benefits of arts education*. New York, NY: Teachers College Press.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172–177.

Kandel, E. R. (2006). *In search of memory: The emergence of a new science of mind*. New York, NY: W. W. Norton.

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 671–685. doi:10.1037/a0018785

Martin, A. J., Mansour, M., Anderson, M., Gibson, R., Liem, G. A. D., & Sudmalis, D. (2013). The role of arts participation in students' academic and nonacademic outcomes: A longitudinal study of school, home, and community factors. *Journal of Educational Psychology*, *105*, 709–727. doi:10.1037/a0032795

Mishook, J. J., & Kornhaber, M. L. (2006). Arts integration in an era of accountability. *Arts Education Policy Review*, *107*(4), 3–11.

Mohr, G., Engelkamp, J., & Zimmer, H. D. (1989). Recall and recognition of self-performed acts. *Psychological Research*, *51*(4), 181–187.

Paivio, A. (1971). *Imagery and verbal processes*. New York, NY: Holt, Rinehart, & Winston.

Phillips, J., Harper, J., Lee, K., & Boone, E. (2013). *Arts integration and the Mississippi Arts Commission's Whole Schools Initiative.* Retrieved from http://www.mswholeschools.org

Rinne, L., Gregory, E., Yarmolinskaya, J., & Hardiman, M. (2011). Why arts integration improves long-term retention of content. *Mind, Brain, and Education*, *5*(2), 89–96. doi:10.1111/j.1751-228X.2011.01114.x

Scripp, L., Burnaford, G. Vazquez, O. Paradis, L., & Sienkiewicz, F. (2013). *Partnerships in arts integration research final reports*. Washington, DC: Arts Education Partnership. Retrieved http://www.artsedsearch.org

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 592–604. doi:10.1037/0278-7393.4.6.592

Smithrim, K., & Upitis, R. (2005). Learning through the arts: Lessons of engagement. *Canadian Journal of Education*, *28*(1), 109–127.

Spelke, E. (2008). Effects of music instruction on developing cognitive systems at the foundations of mathematics and science. In C. Asbury & B. Rich (Eds.), *Learning, arts, and the brain: The Dana consortium report on arts and cognition* (pp. 17–49). New York, NY: The Dana Foundation.

Wandall, B., Dougherty, R. F., Ben-Shachar, M., Deutsch, G. K., & Tsang, J. (2008). Training in the arts, reading and brain imaging. In C. Asbury & B. Rich (Eds.), *Learning, arts, and the brain: The Dana consortium report on arts and cognition* (pp. 51–59). New York, NY: The Dana Foundation.

Zaromb, F. M., & Roediger, H. L. (2009). The effects of "effort after meaning" on recall: Differences in within-and between-subjects designs. *Memory and Cognition*, *37*, 447–463.