Document généré le 23 avr. 2024 05:30

McGill Journal of Education
Revue des sciences de l'éducation de McGill

Curriculum Integration: Opportunities to Maximize Assessment *as*, *of*, and *for* Learning

L'intégration des divers éléments du curriculum : des opportunités pour maximiser l'évaluation *comme*, *de*, et *pour* l'apprentissage

Sheryl MacMath, John Wallace et Xiaohong Chi

Volume 44, numéro 3, automne 2009

URI : https://id.erudit.org/iderudit/039949ar
DOI : https://doi.org/10.7202/039949ar

Aller au sommaire du numéro

Éditeur(s)

Faculty of Education, McGill University

ISSN

0024-9033 (imprimé)
1916-0666 (numérique)

Découvrir la revue

Citer cet article

MacMath, S., Wallace, J. & Chi, X. (2009). Curriculum Integration: Opportunities to Maximize Assessment *as*, *of*, and *for* Learning. *McGill Journal of Education / Revue des sciences de l'éducation de McGill*, *44*(3), 451–465. https://doi.org/10.7202/039949ar

Résumé de l'article

L'intégration des divers éléments des programmes, s'articulant autour de tâches et projets multidisciplinaires s'ancrant dans la « vraie vie », rend possible des pratiques d'évaluation authentiques. Cependant, tenter d'intégrer des mécanismes d'évaluation comme mode d'apprentissage, de et pour l'apprentissage peut se révéler un défi. Pour rédiger cet article, les auteurs se sont basés sur des données issues de deux études de cas mixtes (n=52, n=27) ayant suivi les apprentissages de jeunes étudiants du secondaire au sein de groupe distincts. Ils ont analysé les méthodes d'évaluations préconisées par les enseignants et les relations existant entre l'évaluation et les apprentissages des étudiants. Les résultats et leurs implications pour les pratiques enseignantes sont détaillés.

Copyright © Faculty of Education, McGill University, 2010

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

https://apropos.erudit.org/fr/usagers/politique-dutilisation/

Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche.

https://www.erudit.org/fr/

CURRICULUM INTEGRATION: OPPORTUNITIES TO MAXIMIZE ASSESSMENT *AS, OF,* AND *FOR* LEARNING

SHERYL MACMATH, JOHN WALLACE, XIAOHONG CHI
Ontario Institute for Studies in Education, University of Toronto

ABSTRACT. Curriculum integration, focusing on multidisciplinary tasks/projects relevant to the real-world, lends itself to authentic assessment practices. However, attempting to incorporate assessment *of, for*, and *as* learning can be challenging. Using data from two mixed method case studies (*n*=52, *n*=27) which tracked middle school student learning throughout separate integrated units, we analyzed the types of assessments used by teachers and the relationship of assessment to student learning. Results and implications for teacher practice are explored.

L'INTÉGRATION DES DIVERS ÉLÉMENTS DU CURRICULUM : DES OPPORTUNITÉS POUR MAXIMISER L'ÉVALUATION *COMME, DE,* ET *POUR* L'APPRENTISSAGE[1]

RÉSUMÉ. L'intégration des divers éléments des programmes, s'articulant autour de tâches et projets multidisciplinaires s'ancrant dans la « vraie vie », rend possible des pratiques d'évaluation authentiques. Cependant, tenter d'intégrer des mécanismes d'évaluation comme mode d'apprentissage, de et pour l'apprentissage peut se révéler un défi. Pour rédiger cet article, les auteurs se sont basés sur des données issues de deux études de cas mixtes (n=52, n=27) ayant suivi les apprentissages de jeunes étudiants du secondaire au sein de groupe distincts. Ils ont analysé les méthodes d'évaluations préconisées par les enseignants et les relations existant entre l'évaluation et les apprentissages des étudiants. Les résultats et leurs implications pour les pratiques enseignantes sont détaillés.

Terms such as integration, interdisciplinary, multidisciplinary, and transdisciplinary are related topics that are often viewed as interchangeable by teachers and researchers even though they represent different concepts (Adler & Flihan, 1997; Drake, 1998; Relan & Kimpston, 1993). This has resulted in some confusion regarding these concepts, their application in the classroom, and research findings in relation to student learning. Consequently, research is currently underway attempting to better define the concepts (see Applebee, Adler, & Flihan, 2007) and link these definitions to teacher practice. Unfortunately, this research is not focused on the effects of these practices on student learning,

an area lacking in empirical research (Berlin & Lee, 2005; Hargreaves, Earl, & Ryan, 1996; Hargreaves & Moore, 2000). To address this gap, our current research project focuses on tracking student learning in relation to specific teacher practices during integrated units. This paper focuses specifically on the relationship between student learning, integration, and assessment.

1. How do teachers implementing an integrated unit assess their students?
2. How effective are these assessments in both tracking and assisting student learning?
3. What can we learn from the assessment of student learning in integrated settings that can be applied to teaching practices in general?

Given the ambiguity in the field, our definition of curriculum integration needs to be made clear. In addition, our view of assessment and its relationship to integration also requires exploration.

THEORETICAL FRAMEWORK

Integration

While multidisciplinary, integration, interdisciplinary, and transdisciplinary represent different concepts, consistent across all of these concepts is the understanding that there is an *intermingling* of knowledge and skills across discipline and subject areas (Relan & Kimpston, 1993); however, the type of intermingling, the degree of intermingling, and the purpose of this intermingling varies. Often these concepts are placed on a continuum (Adler & Flihan, 1997; Applebee et al., 2007; Wineburg & Grossman, 2000). On the one end, *multidisciplinary* activities, also known as correlated activities (Adler & Flihan, 1997; Applebee et al., 2007), are concerned with how different disciplines (e.g., mathematics, science, geography, etc.) can complement one-another. The discipline content remains separate and is usually taught in parallel.

When planning involves examining how different disciplines complement each other, this is usually referred to as an *interdisciplinary* unit (Applebee et al., 2007 refer to this as *sharing*). Planning still begins with the different disciplines, with content being pulled around a common theme. Emphasis is placed on identifying the connections between the different disciplines and these connections are made explicit to students. In most cases, the emphasis is on specific curriculum content; teachers start planning with separate subject areas and they assess each subject area independently. The focus questions for teachers planning interdisciplinary units would be: what theme will link these curricula together, or what curricula can we link to this theme?

In contrast, a *transdisciplinary* or *restructured* unit typically starts with a question or project and asks: what do students need to know or be able to do to

answer this question or complete this project? The separation, identification, or linking to different disciplines is not the focus; the focus is the problem or project. *Transdisciplinary* units usually focus on real-world or real-life contexts to shape their questions or projects. In many transdisciplinary units, students generate the key questions under investigation (Beane 1997, 2005). This type of integration is democratic in nature, providing opportunities for students to question, explore, and actively participate in their immediate and global communities. In the fullest sense, this definition of integration represents a completely restructured curriculum.

The problem then arises when the word *integration* is used to refer to different points on the continuum. Beane's (1997, 2005) use of integration specifically refers to a complete restructuring of the curriculum. In contrast, Drake (1998) uses integration as a global term to describe all types of inter-, trans-, and meta-disciplinary mixing. Wineburg and Grossman (2000) took the opposite approach and avoid use of the term *integrated* anywhere in their edited book, choosing instead to have all contributing authors discuss the use of interdisciplinary curriculum in its various forms. In 1997, Adler and Flihan's literature review used integration as a term common to both shared and restructured curriculum; however, by 2007, Applebee et al. recognized that the word *integration* was being used in classrooms to describe all levels of curriculum work, from predisciplinary (referring to strict adherence to discipline boundaries) to restructured.

This progressive expansion of the term *integration* demonstrates both an increase in the use of the term, but also a broadening of its meaning to the point of meaninglessness. As a result, even though integration is referenced and encouraged in both national and international documents (e.g., American Association for the Advancement of Science, 1993; BC Ministry of Education, 1996; Curriculum Council of Western Australia, 1998; National Council of Teachers of Mathematics, 1995, 2000; National Research Council, 1996; Ontario Ministry of Education, 1998; PEI Ministry of Education, 2003-2004), it still remains difficult to determine any empirical effects on student learning given the variation to which it is implemented, understood, and defined. This point makes it critical for researchers to define their use of the word integration in relation to their research study.

To this end, we defined integration as a restructuring of curricula to enable the completion of a final project. In each case, teachers decided on the unit, determined the objectives, and planned the final project. In each case, teachers were responsible for working with and assessing specific curricula outlined by the Ontario Ministry of Education; these curricula are organized by subject areas. Consequently, the lines of each discipline remained a focus for the teachers involved; however, typical of most problem- or product-orientated integrated units (Drake, 1998), the assessments were complex and performance

based rather than traditional. Unfortunately, a review of research literature looking at the implementation of a range of integrated units demonstrates a lack of detail and information regarding assessment (Adler & Flihan, 1997). As a result, this paper aims to analyze the strengths and challenges associated with assessments of student learning in an integrated unit.

Assessment of student learning

The emphasis that shared integration places on evaluating meaningful performance tasks, regardless of discipline-specific knowledge, and the application of knowledge and skills to a problem or issue can result in many assessment challenges. Teachers often need to adapt curriculum learning outcomes to the problem or issue under investigation and develop their own rubrics and benchmarks to track students' application, rather than regurgitation, of knowledge (Drake, 1998). Earl (2003) refers to this as *understanding* or "knowledge in action" (p. 33). Recognizing the complexity involved when assessing a student's level of understanding in relation to a number of curriculum objectives, we searched for an assessment model that took a number of factors into consideration. Our search for an expansive assessment model from which we could effectively analyze student learning throughout an integrated unit led us to the assessment *for, of,* and *as* learning model.

When discussing assessment, Black and William (1998) assert that "a focus on standards and accountability that ignores the processes of teaching and learning in classrooms will not provide the direction that teachers need in their quest to improve" (p. 139). Recognizing that assessment is an important means for affecting instruction, not simply reporting student progress, is pivotal to advancing student success and understanding (Gronlund & Cameron, 2004, p. 11). In addition, a meta-review of research looking at classroom assessment revealed that "innovations that include strengthening the practice of formative assessment produce significant and often substantial learning gains" (Black & William, 1998). Elaborating on formative assessment in future studies, Black, Harrison, Lee, et al. (2004) observed substantial learning gains when teacher assessment practices included self-assessment opportunities, sharing criteria with students, and peer assessment. Consequently, it is not surprising that assessment that provides direction for teaching provides balanced opportunities for summative, formative, and student self-assessment (Burke, 2005).

Rethinking Classroom Assessment with Purpose in Mind, published by the Western and Northern Canadian Protocol (WNCP), provides a model for classroom assessment (WNCP, 2006). It includes a manual resource focused on "classroom assessment, not large-scale assessment... It is designed to provide a framework for thinking as teachers, administrators, and professional developers work together over time in developing and using assessment in their classrooms to differentiate and facilitate learning for all students" (WNCP, 2006, p. vii). *Rethinking Classroom Assessment* views achievement broadly; the authors advocate

for assessment practices that monitor daily classroom activities to improve student motivation (measured by work habits, persistence, and enthusiasm), support student academic performance, and provide increased opportunities to differentiate instructional and assessment practices so as to reach individual learner needs. The framework is structured around three different purposes in relation to assessment: assessment *as* learning, assessment *for* learning, and assessment *of* learning.

Assessment *as* learning focuses on teaching students how to assess themselves and others. Included under assessment *as* learning activities would be self-assessment, peer assessment, and sharing criteria with students. The focus for assessment *as* learning is enabling students to be critical evaluators of their own work; it requires an explicit understanding of the criteria for each assignment, enabling students to be able to identify whether they have met the shared criteria or not. Assessment *for* learning is most closely associated with formative assessment practices. The purpose of assessment *for* learning is an on-going assessment of student work and understanding to enable the teacher to modify and alter daily lesson plans and student activities. Assessment *of* learning, also referred to as summative assessment, represents a final assessment used to evaluate and rank a student. This assessment *of* learning is usually found in student report cards and is communicated to parents, other teachers, administrators, and the community at large.

To maximize assessment practices, *Rethinking Classroom Assessment* provides the following three guidelines: (a) utilize assessment practices that relate to all three purposes, (b) ensure congruence between the types of assessment tools and records and the assessment purpose, and (c) utilize tools and records that are valid and reliable. Utilizing the model provided by assessment *as, for,* and *of* learning, we examined student learning from two separate case studies involving teachers and students working through integrated units.

RESEARCH DESIGN

This dialectical (Greene, Caracelli, & Graham, 1989), mixed-methods study utilized teacher and student interviews, student pre- and post-tests, classroom observations, audio recordings of student group work, student motivation surveys, and the analysis of text materials to provide a variety of qualitative (thick descriptions) and quantitative (e.g., frequency counts, survey results) data. Given the domination of separate qualitative or quantitative studies in research literature, we felt it would be valuable to unpack what we mean by a *dialectical, mixed-methods* study; contrary to popular belief, mixed-methods is not simply a matter of gathering both qualitative and quantitative data.

Traditionally, qualitative and quantitative methods were considered separate paradigms: qualitative methods aimed to provide thick descriptions of events while quantitative data aimed to establish cause and effect (Creswell & Plano

Clark, 2007; Denzin & Lincoln, 2000; Tashakkori & Teddlie, 1998). Given these separate purposes, many researchers, to this day, believe that qualitative and quantitative data should (or can) not be mixed. In contrast, *paradigmatic* mixed-method researchers argue that, while there are philosophical differences associated with different methods, these "assumptions are logically independent and therefore can be mixed and matched" (Greene & Caracelli, 1997, p. 8). Consequently, differences in paradigmatic assumptions become irrelevant – matching appropriate methods to specific research questions and contextual variables (e.g., time constraints, accessibility, etc.) are all that matter. Within the past 20 years, an alternative mixed-methods position has emerged amongst numerous researchers.

The *dialectical* approach to mixed methods argues that there are paradigmatic differences between qualitative and quantitative research methods and researchers need to be aware of these differences so as to effectively tack back-and-forth between different methods. By deliberately using these different paradigms, we are able to envision new discoveries and understandings of our data (Greene & Caracelli, 1997). Given that our research questions attempted to both describe pedagogy (in relation to assessment) and link it to student learning, we required both thick descriptions as well as attempts to establish causal linkages. Consequently, we utilized a mixed-methods approach to gather both; however, we remained aware of the paradigmatic assumptions linked to our methods. Throughout our study, ensuring that we did not make causal attributions based solely on qualitative data was one of our greatest challenges. Our solution to this problem also helped address a second problem common to mixed-methods research.

A lack of data integration at the analysis stage is common to mixed-methods research (Maxwell & Loomis, 2003). As a result, rather than integrating data throughout analyses, researchers tend to analyze their qualitative and quantitative data separately, integrating them only at the conclusion stage of research. Aware of this, we consciously planned numerous opportunities to integrate (or mix) our analyses (e.g., move between the thematic analyses present in our observations of classroom activities and the quantitative analysis of student learning in the final interviews). By making these conscious points of integration, we were also able to monitor the data sources inherent to any causal attributions that emerged.

Given our unpacking of *dialectical mixed-methods*, we now focus on our specific data sources. Using stratified, purposive sampling (Teddlie & Yu, 2007), we identified two sites in one of Canada's largest public school systems and received voluntary and informed consent to observe their classrooms while teachers and students completed integrated units.

DATA SOURCES

At the first site we observed a pod of two teachers and fifty grade 8 students engaged in a toy-building unit. During the unit, students experimented with simple machines, surveyed other students in the school, and then designed, built, and marketed a moving toy. While we gathered observations on the class as a whole, we focused our observations and interviews on twelve students in three homogenous ability groups (total observation time = 26 hrs). At the second site we observed one teacher and twenty-six grade 6 students engaged in a thematic hockey unit. During the unit, students designed a hockey season schedule, created travel brochures, and built an ice rink unit (including lighting and sound systems) to scale. For this site we focused our observations and interviews on ten students in two mixed ability groups (total observation time = 33 hrs).

For both sites, we simultaneously collected qualitative and quantitative data, with some earlier data being used to structure the data tools used later in the study. This sequence resulted in three main stages of data collection: pre-unit, during unit, and post-unit. All three stages were completed at both sites. The pre-unit data collection included interviews with the site teacher(s) and each of the selected students. We placed special emphasis on having the teacher(s) identify the key learning academic goals for the unit. The student interviews had students describe what they thought integration was and included application problems related to the key learning outcomes identified in the teacher interview.

During the unit we completed observation records that both coded student behaviour and provided a descriptive written record of teacher instructions, teacher behaviours, and general observations of student behaviours. Given the high levels of group work inherent to integrated units, we audio recorded all small group discussions and problem solving. In addition, we conducted mini-interviews with students and teachers regarding perceptions of activities, teacher assessments and reasoning, and student *think-alouds* which aimed to expose student thinking, knowledge acquisition, and transfer. The post-unit data included a final teacher interview, final student interviews, and a whole class motivation survey. During the final interview we asked teachers to comment on specific aspects of the unit as well as identifying, for each of the key objectives identified in the initial interview, the *evidence* they had for assessing whether an individual student had or had not achieved that objective. The final student interviews asked students how meaningful they had found each activity throughout the unit as well as a series of knowledge and skill testing questions designed to assess each student's ability to apply the key learning objectives identified at the start of the unit.

RESULTS

Both sites utilized separate marking guides or rubrics for each subject area; while they considered this an integrated unit, they assessed each discipline independently to match the curricula set out by the Ontario Ministry of Education. These case studies also revealed a number of strengths and challenges with regards to assessment *as, for*, and *of* learning. At the first site, students were continually made aware of the criteria for which they were being evaluated; both teachers took a minimum of one period to explain and give examples of how students could meet the rubric designed for the final project (one for math and one for science). This opportunity for assessment *as* learning remained a focus for students who continually referred to their rubrics while building and marketing their toys. For example, the rubric described how students were to make explicit connections between their survey results (summarized by measures of central tendencies) and their schematics; all marketing display boards provided this comparison and, in the final interview, students were very successful at determining measures of central tendency when problem solving. In addition, both teachers encouraged students to refer to the rubrics while building their toys and provided opportunities for groups to give each other feedback with regards to the criteria. These self- and peer-assessment opportunities were reflected in student discussions which illustrated a common language for discussing their projects and a continued focus on the objectives listed in the rubrics.

For the math objectives, both teachers then provided numerous opportunities for groups to share their observations and data with the rest of the class, encouraging students to ask questions of each other and challenge each other's conclusions. These opportunities to share and ask questions led to numerous whole class discussions which focused on student-to-student conversations and thinking while providing opportunities for teachers (and researchers) to identify the problems students were having and provide alternative examples and clarifications. As a result, these assessments *as* learning opportunities in math led to assessment *for* learning opportunities. The teachers used these formative assessments to modify their lessons and support student learning. These discussions were referenced in student final interviews as being very helpful and students were very successful in applying those skills (e.g., calculating measures of central tendency and interpreting a graph) which involved this type of peer interaction and assessment.

In contrast to the strong emphasis on assessment *as* learning for both math and science, the first site did not engage in a large number of assessments *for* learning with regards to the science objectives. Students found the task of designing, constructing, and marketing their own movable toy completely engaging; even though they had watched videos, had group discussions, lectures, and completed experiments with regards to simple machines, applying that knowledge proved challenging for many students. As they worked through

their schematic development and toy construction, the task differentiated those students who understood and could apply what they had learned and those who could not. For example, one group included a pulley in their schematic; however, they did not correctly apply what they had learned regarding pulleys, attempting to use a pulley to change the direction of force rather than minimize the amount of force required. This error on the schematic was not addressed by the teachers and, when these students tried to build their toy, they were unsuccessful.

Similarly, another group of students tried to build a toy car with rotating wheels; when these students experienced difficulties, they dropped their original plan and ran out of time to complete their final project. In the end, they had to "Frankenstein" a car by taking the motor out of an existing toy car and building a new carriage. As a result, for the science objectives, assessments *for* learning were not identified or used to alter instruction. Instead, the differentiation between students who were and were not able to build a toy using a simple machine was used in an assessment *of* learning. Just as important, when asked to apply their knowledge of levers (as a type of simple machine) at the end of the unit, the majority of students were unsuccessful. Consequently, objectives that included an assessment *for* learning (i.e., math) positively supported student transfer of knowledge and objectives that did not include an assessment *for* learning (i.e., science) demonstrated low levels of knowledge transfer.

At the second site, the use of assessment *as* learning was not as prevalent. While the teacher did discuss the attributes of successful completion (e.g., all aspects of the ice rink need to be to scale, 40% of the games need to be in the Eastern division, etc.), these attributes were not linked to specific curriculum objectives (e.g., converting percents to decimals, differentiating between a parallel and series circuit, etc.). As a result, we observed students focusing on the final products and having it "look" like what was expected rather than focusing on how they were accomplishing each task. For example, students experimented with many different ways of calculating 40% of 80 games and, when they stumbled across an answer that made sense, they used that method; however, they were unsure of how to relate percents to decimals or which procedure to use for future problems. In the final interviews, when students were asked to calculate percentages, the majority of students were unsuccessful.

Similar to the first site, several tasks at the second site provided opportunities for assessment *for* learning. For example, one of the major tasks involved in the unit was the building of a scale model ice rink with a working light and buzzer system. This required students to apply their math skills associated with ratios and science skills related to electrical circuits, both of which had been taught prior to the integrated unit. While groups avidly worked on the task of building their ice rinks over a two-to-three week period, some groups

experienced a great deal of difficulty building the rink to scale and getting the lights to work. As noted at the first site, the integrated tasks clearly differentiated between those students who understood proportions and those who did not, those who knew how to build a parallel circuit and those who did not. In contrast to the first site, this teacher did identify those students who were unsuccessful (teachers at the first site did not identify those groups who were unsuccessful); however, in this case, the teacher *chose* not to intervene. When asked why she did not intervene, the teacher shared that she was concerned about providing assistance as it would affect the validity of her final assessment. When asked how she would evaluate a group that did receive assistance from the researcher, enabling students to correct their errors in scaling, the teacher replied that: "It's like a whole process, like they did get it, but it took them a while, so they may be a little bit lower because they had to get some assistance" (teacher, site #2). Consequently, this teacher chose not to utilize the information she had learned through her assessment *for* learning as she felt it would jeopardize the validity of her assessment *of* learning; the ice rink was her one major assessment tool for those particular objectives. When we tested these students on their ability to calculate measures to scale, the majority were unsuccessful. Even though time had been provided for students to attempt to "work" through the problem, their understanding was not adjusted and they were unable to apply those concepts after the completion of the unit.

CONCLUSION

Both sites illustrated the power of assessment *as* learning. At the first site, assessment *as* learning opportunities related to increased student ability to transfer knowledge. In contrast, at the second site, reference to the objectives remained global and vague, focused instead on the look of the final product. As a result, throughout the unit, students remained focused on the look of the final product as opposed to the specific objectives for the unit. This then led to an overall decrease in success on the curriculum objectives. These examples illustrate the value not only in sharing criteria with students and maximizing assessment *as* learning opportunities, but also the importance of focusing these assessment opportunities on specific curriculum objectives as opposed to global, vague project descriptors.

With regards to assessment *for* learning, both sites demonstrated the rich opportunities that integrated tasks provide in the creation of assessments *for* learning; in each case, the application of knowledge that was required for the completion of each integrated task clearly illuminated gaps in student understanding. This illumination of gaps in student understanding was especially true in the second case study when students' inability to determine percentages shocked the teacher as this was a skill that students had already "learned" and been "evaluated" on prior to the integrated unit; she was surprised that they were unable to transfer knowledge that she assumed they possessed, since she

had already evaluated them (traditionally, using written tests) as successful. These two case studies, if anything, clearly demonstrate the importance of providing real-life problems or projects whereby students are asked to apply knowledge and skills a teacher may think they already have; in many ways, these findings strongly support the use of alternative performance assessments to increase the validity of student evaluations.

However, it is important to note that both sites also demonstrated the challenges involved with assessment *for* learning. As illustrated by the first case study, teachers need to be monitoring for student successes and challenges and use the information they acquire from their assessments *for* learning to adjust their planning and clarify activities. At the first case study, time was a huge factor (e.g., report cards, coordinating with the other grade 8 teachers in the school, coordinating with incoming practicum students, etc.); these two teachers did not have the time to closely monitor each group's progress through the toy construction. As a result, even though student learning (or lack thereof) was clearly evident during the science-heavy schematic and construction day activities, this had not been identified by teachers and, consequently, instruction was not altered and student learning decreased. In relation to math, teachers at the first site did use the assessment information they gathered *for* learning to modify and adapt their lessons, increasing student success. In contrast, while the teacher at the second site did closely monitor groups through their construction phase, she chose not to use the information she gathered formatively out of fear that it would make her assessment *of* learning invalid. Putting the findings from both sites together, it becomes clear how important it is to (a) schedule time for ongoing assessment *for* learning – teachers need to recognize how vital it is to closely monitor group work and (b) use the information that you learn about student progress to provide specific feedback and alter activities so as to fill in the learning gaps which were identified by students. If, as at the second site, a teacher is worried about the validity of his or her assessment if they intervene, this can be managed by increasing the number and/or variety of assessments being used. Assessment *for* learning, when used for the purpose of adapting instruction, appears to increase students' ability to transfer knowledge beyond the immediate unit of instruction; if student understanding is a teacher's primary concern, learning gaps need to be dealt with earlier (formatively) rather than after the fact (summatively).

LOOKING BEYOND INTEGRATION

We now return to the final question we asked at the beginning of this paper: what can we learn from the assessment of student learning in integrated settings that can be applied to teaching practices in general? There are three things we suggest teachers can take away from these two case studies (in relation to assessment):

1. Even if a teacher is not using integrated units, it is important to include an alterative performance/project assessment opportunity that requires students to apply what they have taught/learned; these case study results strongly support the problem-solving model.
2. We found that the more an assessment moves from traditional tasks (e.g., exams, essays, reflections) towards an alternative project, product, or performance assessments, the more time is needed for assessment *for* learning opportunities. In this sense, group work and independent projects lead to an increase in the amount of supervision and guidance students need from their teacher(s).
3. Assessment *as* learning, when specific to identifiable objectives rather than overall global aims/understands, appears to increase students' ability to transfer knowledge (or *understand*). These case studies reveal how critical it is to share assignment criteria with students and teach students to be effective assessors of their own and others' work.

These results strengthen arguments for both the use of integrated tasks to assess student understanding and the importance of having assessment *for* learning as a priority; we advise teachers to plan these assessments *for* learning into the overall unit schedule as well as continually use what is learned to alter activities and redirect student learning.

LOOKING BEYOND THIS STUDY

Reflecting on the use of a dialectic mixed-methods approach for this study, a discussion regarding data analyses is pertinent. During the analysis stage, our planned opportunities to integrate the data were important and often led to a re-analysis of data. Two specific types of integrated analyses were especially valuable. First, data transformation led to our first *recycling* of data. After an initial thematic analysis of the observation records and teacher interviews (qualitative data), we were able to break down each unit into specific activities linked to specific learning objectives. Using this new thematic organization, we recoded our quantitative data enabling a quantitative description of the unit. For example, we were able to identify percentages of time spent on each learning objective, percentages of time per activity, as well as levels of on- and off-task behaviour by learning objective. This data transformation then led to an opportunity to consolidate our data.

After examining how much time and how on-task students were for each learning objective, we compared these results with students' ability to demonstrate each learning objective in their final interviews. This revealed an important inconsistency. Students' ability to successfully transfer knowledge did not correlate with the amount of time spent on each objective or student levels of on-task behaviour; something else was affecting student learning. This then

led to a re-visitation of the data. When consolidating our teacher and student interviews with our observation record descriptions of classroom activities, assessment emerged as a key factor. This then led to a second recoding of our qualitative data (another type of data transformation) using the theoretical model of assessment *as, of,* and *for* learning, the results of which were reported earlier in this article.

Consequently, a dialectic mixed-methods approach, given the size of our data sets and the expansiveness of the questions we were asking – trying to link pedagogy and student learning – was the best research design. However, it required numerous re-visitations and recycling of our data set. In addition as researchers, we needed to be adept in managing both qualitative and quantitative data, recognizing how to use and re-use data appropriately. Nevertheless, this increased time spent managing and analyzing data resulted in new insights and new areas of investigation. Further research examining the links between assessment *as, of,* and *for* learning and student success is warranted.

ACKNOWLEDGEMENT

The authors would like to thank the Social Sciences and Human Research Council for their generous funding of this project.

REFERENCES

Adler, M., & Flihan, S. (1997). *The interdisciplinary continuum: Reconciling theory, research, and practice* (Report Series 2.36). Washington, DC: Office of Educational Research and Improvement. (ERIC Document Reproduction Service No. ED414602)

American Association for the Advancement of Science. (1993). *Benchmarks for scientific literacy*. New York: Oxford University Press.

Applebee, A. N., Adler, M., & Flihan, S. (2007). Interdisciplinary curricula in middle and high school classrooms: Case studies of approaches to curriculum and instruction. *American Educational Research Journal, 44,* 1002-1039.

Beane, J. A. (1997). *Curriculum integration: Designing the core of democratic education.* New York: Teachers College Press.

Beane, J. A., (2005). *A reason to teach: Creating classrooms of dignity and hope.* Portsmouth, NH: Heinemann.

Berlin, D. F., & Lee, H. (2005). Integrating science and mathematics education: Historical analysis. *School Science and Mathematics, 105*(1), 15-24.

Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86*(1), 9-22.

Black, P., & William, D. (1998). Inside the black box. *Phi Delta Kappan, 80*(2), 139-148.

British Columbia Ministry of Education (1996). *Science 8-10: Integrated resource package*. Victoria: Author.

Burke, K. (2005). *How to assess authentic learning: 4th Edition.* Thousand Oaks, CA: Corwin Press.

Creswell, J. W., & Plano Clark, V. L. (2007). Choosing a mixed methods design. In *Designing and conducting mixed methods research* (pp. 58-88). Thousand Oaks: Sage Publication.

Curriculum Council of Western Australia (1998). *Curriculum framework*. Perth, WA: Author.

Drake, S. M. (1998). *Creating integrated curriculum: Proven ways to increase to student learning.* Thousand Oaks, CA: Corwin Press.

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning.* Thousand Oaks, CA: Corwin Press.

Greene, J. C., & Caracelli, V. J. (1997). Defining and describing the paradigm issue in mixed-method evaluation. *New Directions for Evaluation, 74,* 5-17.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255-274.

Gronlund, N. E., & Cameron, I. J. (2004). *Assessment of student achievement (Canadian Edition).* Toronto, ON: Pearson Canada.

Hargreaves, A., Earl, L., & Ryan, J. (1996). *Schooling for change: Reinventing education for early adolescents.* London: Falmer Press.

Hargreaves, A., & Moore, S. (2000). Curriculum integration and classroom relevance: A study of teachers' practice. *Journal of Curriculum and Supervision, 15*(2), 89-112.

Maxwell, J. A., & Loomis, D. M. (2003). Mixed methods design: An alternative approach. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioural research* (pp. 241-271). Thousand Oaks, CA: Sage.

National Council of Teachers of Mathematics. (1995). *Connecting mathematics across the curriculum.* Reston, VA: NCTM.

National Council for Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: NCTM.

National Research Council. (1996). *National science education standards.* Washington, DC: National Academy Press.

Ontario Ministry of Education. (1998). *The Ontario Curriculum, Grades 1-8: Science and Technology.* Toronto: Author.

Prince Edward Island Department of Education. (2003-04). *Intermediate program of study and list of authorized materials.* Charlottetown: Author.

Relan, A., & Kimpston, R. (1993). Curriculum integration: A critical analysis of practical and conceptual issues. In R. Fogarty (Ed.), *Integrating the curricula: A collection.* Palatine, IL: IRI/Skylight Publishing (pp. 31 – 48).

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches.* Thousand Oaks, CA: Sage.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research, 1*(1), 77-100.

Western and Northern Canadian Protocol for Collaboration in Education. (2006). *Rethinking classroom assessment with purpose in mind.* Winnipeg, MB: Manitoba Education, Citizenship, and Youth. Available at http://www.wncp.ca/assessment/rethink.pdf

Wineburg, S., & Grossman, P. (Eds.). (2000). Introduction: When theory meets practice in the world of school. In *Interdisciplinary curriculum: Challenges to implementation.* New York, NY: Teachers College Press, (pp. 1 – 16).

SHERYL MACMATH is a doctoral student at the Ontario Institute for Studies in Education at the University of Toronto. She is also an instructor for the Teacher Education Program at the University of the Fraser Valley. Her interests include curriculum integration, classroom assessment, and citizenship development within our elementary schools.

JOHN WALLACE is a professor at the Ontario Institute for Studies in Education at the University of Toronto. His interests include curriculum integration, qualitative research methodologies, and science education.

XIAOHONG CHI is a doctoral student at the Ontario Institute for Studies in Education at the University of Toronto. She is currently researching student cultural perceptions as they are developed through their reading activities.

SHERYL MACMATH est doctorante au Ontario Institute for Studies in Education de l'Université de Toronto. Elle est aussi formatrice au programme de formation des enseignants à l'Université Fraser Valley. Ses recherches portent sur l'intégration des programmes, l'évaluation en classe et le développement de la citoyenneté au sein des écoles primaires.

JOHN WALLACE est professeur au Ontario Institute for Studies in Education de l'Université de Toronto. Ses recherches portent sur l'intégration des programmes, la méthodologie des recherches qualitatives et les sciences de l'éducation.

XIAOHONG CHI est doctorante au Ontario Institute for Studies in Education de l'Université de Toronto. Ses projets de recherche portent actuellement sur les manières dont les perceptions culturelles des étudiants sont façonnées par leurs activités de lecture.