Improving Middle School Students’ Understanding of Core Science Ideas Using Coherent Curriculum

Joseph Krajcik, Sung-Youn Choi, Michigan State University, 237 Erickson Hall, East Lansing, MI
Krajcik@msu.edu, choisung@msu.edu

Namsoo Shin, LeeAnn M. Sutherland, University of Michigan, 610 E. University Ave., Ann Arbor, MI
namsoo@umich.edu, lsutherl@umich.edu

Abstract: This study explores the influence of coherent curriculum on student understanding of core chemistry ideas across the middle grades. Participants in urban and suburban schools used materials designed based on what is known about promoting student learning of science, and that link ideas within and across grade levels in a coherent manner. Using a cross-sectional approach, student learning was assessed using an independent measure aligned with a learning progression on the transformations of matter. To develop the learning progression and corresponding assessment, we used construct-center design. Using IRT and analysis of variance, we demonstrate that students’ latent ability changes across grade level, illustrating that curriculum coherence can, indeed, support students in building understanding of core chemistry ideas. We also found that boys and girls did not perform differently, in contrast with other findings, particularly in urban U.S. middle schools, in which girls outperform boys.

Introduction and Rationale
Research shows that students: 1) demonstrate little understanding of the core ideas of chemistry, 2) lack interdisciplinary connections among key ideas, 3) fail to apply core ideas to understand chemical phenomena, and 4) do not use scientific reasoning and practices (Johnson 2000; Nakhleh, & Samarapungavan 1999). Because research illustrates that many students learn ideas superficially and that such limited understanding is not likely to prepare students for the world in which they live now and will live as adults, deeper understanding of the core ideas of science becomes instrumental for problem-solving and decision making in a competitive global economy, as well as for the personal well being of each citizen. Derived from research, our premise is that students need to develop integrated understanding of core ideas to actively contribute to and to use 21st century knowledge, practices, and skills to problem-solve and to make decisions. To help students build integrated understanding, one promising approach is developing and using coherent curriculum materials, designed such that ideas build upon one another over time and across contexts such that richer understanding and greater connectedness among those ideas result (Margel, Eylon & Scherz, 2008; Roseman, Linn & Koppal, 2008; Shwartz, Weizman, Fortus, Krajcik, & Reiser, 2008).

Curriculum coherence—“presenting a complete set of interrelated ideas and making connections among them explicit” (Roseman, Linn, & Koppal, 2008)—has been identified as a critical predictor of student achievement (Schmidt, Wang, & McKnight, 2005). Unfortunately, most traditional science textbooks and instruction in the U.S. do not support deep, integrated learning due to a lack of coherence (Kesidou & Roseman, 2002). Using learning-goals driven design (Krajcik, McNeill, & Reiser, 2008), we developed a middle school science curriculum for grades 6-8 with coherence as a central design tenet. Investigating and Questioning our World through Science and Technology (IQWST) (Krajcik, Reiser, Sutherland & Fortus, 2011) is a project-based curriculum comprised of biology, chemistry, physics, and Earth science units. The effects of using coherent curriculum on learning remain largely untested in the U.S. due to a dearth of coherent curricula and of appropriate measures to assess learning. We have researched use of the materials in classroom contexts to examine how the linking of ideas in a coherent manner supports student learning of core science content.

Research questions
Few studies report the result of using a coherent curriculum in the classroom. This situation arises because most curricula do not link content through and across years. The purpose of this study was to examine whether the three units in the chemistry strand, which link content, build understanding of targeted chemistry ideas, and to examine students’ performance by gender, school, and teacher. As such, the primary research question was: When students experience coherent curriculum materials, does understanding develop across time? We predicted that continuous growth in understanding would result as students experienced learning tasks and their ideas became more integrated using coherent materials across the middle school years. We also asked if students’ performance differ by gender, school of teacher.

Methodology
IQWST’s coherence was accomplished by using research literature to “identify clusters of science ideas that interrelate… and build[ing] an instructional sequence to foster more complex understandings over time”
Learning Progressions (LPs) provide a developmental perspective on helping students learn. Interactions sequence successively more complex as learners move from less understanding to more sophisticated understanding as they engage in new learning experiences (NRC, 2008; Smith, et al., 2006). Learning progressions (LPs) provide a developmental perspective on helping students learn.

Student Materials
This study focuses on the chemistry sequence across 6th, 7th, and 8th grades. Each unit takes approximately 10 weeks to complete. We designed each unit iteratively, following cycles of choosing standards to address, creating standards-based learning performances, developing materials, and collecting and analyzing student data, as well as incorporating feedback from teachers and students during and following classroom enactment (Shwartz, et. al, 2008). Learning core ideas and practices across content areas and time provides students with opportunities to develop, reinforce, and use their understanding on an ongoing basis throughout their three-year experience with IQWST. This enables students to develop understanding that they can call on and apply in newly encountered contexts. Scientific practices include designing investigations; data gathering, organization, and analysis; modeling; and constructing evidence-based explanations. Another important feature is that it is project based (Shwartz et. al, 2008) in that students perform investigations and seek information to find solutions to important and meaningful questions. Each unit uses a driving question (DQ) that anchors content in a context that is meaningful to students. DQs are rich, open-ended questions that use everyday language to connect science learning with authentic interests and curiosities students have about the world. By definition, a driving question must be meaningful, worthwhile, feasible, and “real world.” As such, the DQ is used to link the various learning tasks and investigations, central to coherence of the materials. In the process of finding solutions to the driving questions, students learn important science content. IQWST also provides students with opportunities to develop literacy in science through extensive reading and writing opportunities.

Instrument Development
Explaining what happens during transformation phenomena (chemical reaction, phase change, expansion and contraction) involves incorporating ideas from topic areas including: the structure of matter, conservation, forces and energy. We iteratively built an instrument based on the learning progression (LP) by considering the sets of ideas that students need in order to explain a range of transformations of matter. To develop the learning progression and corresponding assessment, we followed a construct-center design approach (CCD) (Shin, Stevens, & Krajcik, 2011). We carefully unpacked the ideas related to the structure and interaction of matter and examined the learning literature on students’ development of these ideas. Content experts reviewed this work to determine the validity of the learning progression and assessment items. Their feedback guided revision. As we designed items, we required that students incorporate ideas from multiple content areas, when possible, within a single item, regardless of the level of the LP. We developed assessment items to measure how students applied ideas within and across topics to explain phenomena involving transformations of matter. Items were validated through rounds of internal and external review, followed by pilot studies. To ensure that students interpreted the questions and representations as intended, each item was accompanied by a set of questions that probed student understanding of the item (DeBoer, et al., 2008). Each item was piloted with students from grades 6-14. Semi-structured interviews with a subset of students supplemented the pilot test data.

To measure the entire construct (LP) and to ensure that we had a reliable instrument, we generated a sufficient number of items to develop 8 different forms. Based on analysis of pilot data, we developed four test forms, each containing 14 items for 6th and 7th grade, and four test forms, each containing 15 items, for 8th grade. Each test form contained three pairs of linking items for a total of six linking items. The different forms were used to adjust the overall difficulty of the tests. We eliminated items with a discrimination value of less than 0.30 unless we could identify a problem with the clarity and/or wording of the question or one of the possible answers. Assessment items can be considered distal (Ruiz-Primo, et al., 2002) and can serve as an independent measure of curriculum effectiveness, as they were designed to measure students’ level on the LP and did not consider whether these ideas were represented in the particular curriculum or how the curriculum sequenced those ideas. Because the items on the various forms were developed independently of the curriculum,
we were interested in determining how students performed on items that aligned with IQWST learning goals and those that did not. We classified each item either aligned or not aligned with IQWST. We then ran Multidimensional IRT analysis by treating each item group differently. This analysis indicates that IQWST students performed similarly on IQWST-aligned items and non-IQWST aligned items. The distributions of items are located between -4.2 and 2.9 logits. These findings indicate a wide range in the difficulty level that covered easy-to-difficult items that are appropriate for measuring 6th to 8th grade students’ chemistry idea.

Participants and Data Collection

Data are reported from 6 schools that used IQWST materials during national field trials but completed the assessments after their participation in field trials had concluded. Not all the schools involved in this study took part in the national field trials. Two of the schools represent middle-to-upper-middle class socioeconomic status, and 4 of the schools were from neighborhoods that are lower-SES, with a majority of students eligible for free and reduced lunch. We administered the items in the fall, 2011. Seventeen teachers took part in the study; several taught more than one grade level. A total of 1519 students took part, with 481 students in 6th grade who had only begun to study IQWST; 519 students in 7th grade, who had studied IQWST for one year, and 519 students in 8th grade, who had studied IQWST for two years. The study design is cross-sectional.

Results

We estimated students’ latent ability parameter based on item response theory (IRT, McNamara, 1996) analysis. Students’ latent ability is considered a continuous latent variable using the partial credit model (PCM) (Masters, 1982). The raw score, which is the number of items correct, does not present a broad picture of test performance because it can be interpreted only in terms of a particular set of test questions. Latent student ability allows direct comparisons of student performance between specific sets of test items. Items were evaluated in terms of the goodness of fit (chi-square=9,487, df=120, p<.001). This shows that the items have different difficulty levels and that IRT is the appropriate method of analysis. First, we used fit index (using MNSQ, T-value >.40) and discrimination (r <.20) to identify poor items. The index provides information about whether items discriminate or not. In the IRT output, the mean square (MNSQ) of fit indices indicate the extent to which the item fits the item response model. The MNSQ, T-value, and discrimination are also good showing that the items are a valid measure. Second, we discussed items in face-to-face meetings to decide how to revise or to eliminate the items from future use of the instrument and further analysis.

Table 1. Mean and S.D. by grades

<table>
<thead>
<tr>
<th>Grade levels</th>
<th>N</th>
<th>Mean (S.D)</th>
<th>Mean diff.</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>6th</td>
<td>481</td>
<td>454 (84)</td>
<td>(7th-6th)</td>
<td>49* 58%</td>
</tr>
<tr>
<td>7th</td>
<td>519</td>
<td>503 (94)</td>
<td>(8th-7h)</td>
<td>37* 39%</td>
</tr>
<tr>
<td>8th</td>
<td>519</td>
<td>540 (102)</td>
<td>(8th-6th)</td>
<td>86* 102%</td>
</tr>
</tbody>
</table>

(* p<.001)

To facilitate interpretation, we determined students’ latent ability parameter to ability score to have an average score of 500 points and a standard deviation of 100(1), with about two-thirds of students scoring between 400 and 600 points. To answer the main research question regarding differences among the grade levels, we calculated whether there was a mean difference of students’ latent ability parameter to ability score to determine whether students improved across the years. We also calculated the effect size by dividing the mean differences by standard deviation of the lower grade (see Table 1). Significant differences were found between 6th, 7th, 8th grade (F=107, p<.001). From 6th to 7th grade, students’ ability is improved by 58% of 6th grade standard deviation. The greatest difference was between 6th and 8th graders. Between 8th and 7th graders, the least difference occurred. We speculated that this occurred because the instrument contained few assessment items that focused on the chemical transformations, the focus of the 7th grade curriculum that would influence students’ performance on the fall, 8th grade assessment. Overall, however, these results support our hypothesis that students improve across grade levels.

Figure 1 shows students’ ability score in a graphical representation, illustrating how students developed in their ability to respond to the chemistry items from 6th through 8th grade. The graph shows a shift in student ability score. Most students’ ability ranged from 300 to 750 with mean of 500. In 6th grade, most students were located from 300 to 600 with a mean of 454. In the 7th grade, the range was from 350 to 750, and 300 to 800 in 8th grade. The variance of 8th grade students (102 ability score) is wider than others. We speculated that this wider variance occurred because although the 7th grade curriculum supported students in understanding chemical transformations, the fall 8th grade assessment forms did not focus on chemical transformations.

To explore the if there is a relationship in learning due to learning, we calculated the mean difference in students’ ability score to determine whether difference existed between genders. Table 2 shows the number of
students, mean, standard deviation, mean difference and t value for each category (male or female). There was
no significant gender difference (t=1.39, p=0.166). This is an important finding as it shows that IQWST supports
student learning regardless of gender. Typically, particularly in urban schools, females outperform males
(Geier, et al., 2008).

To compare the effect of teacher, gender and school to students’ ability score by each grade, we
conducted ANOVA. The Table 3 shows the effect size of teacher, gender, and school by each grade. Overall, the
results show that differences in teacher (F=35.3, p<.001) and school (F=52.5, p<.001) exist and that these tend
to increase by grade. The differences in the teacher variable results in 16.2 standard deviations in 6th grade
(F=16.5, p<.001), 17.5 in 7th (F=14.8, p<.001), and 21.2 in 8th (F=18.4, p<.001). The differences associated with
school variable makes 16.4 standard deviation in 6th grade (F=19.8, p<.001), 16.6 in 7th (F=21.7, p<.001), and
21.5 in 8th (F=29.4, p<.001). These results indicate that teacher variables such as teaching strategies, teaching
experience, or student-teacher relationships; and school characteristics such as environmental supports, study
hours, or school culture may impact students’ ability scores.

Table 3. Effect size of variables on students’ ability score by grade

<table>
<thead>
<tr>
<th>Variable</th>
<th>6th grade</th>
<th>7th grade</th>
<th>8th grade</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher effect</td>
<td>16.2% *</td>
<td>17.5% *</td>
<td>21.2% *</td>
<td>26.5% *</td>
</tr>
<tr>
<td>Gender effect</td>
<td>0.4%</td>
<td>0.9%</td>
<td>-0.1%</td>
<td>0.1%</td>
</tr>
<tr>
<td>School effect</td>
<td>16.4% *</td>
<td>16.6% *</td>
<td>21.5% *</td>
<td>14.5% *</td>
</tr>
</tbody>
</table>
(* p<.001)

Discussion

Our work focuses on coherence around core ideas of science that are foundational across science disciplines (i.e.,
biology, Earth science, physics, and chemistry) and important for all citizens to understand in a science- and
technology-based society. These core ideas were the focus across IQWST units. Our hope in conducting this
work was to illustrate that when coherence is built into curriculum materials, students subsequently can build
integrated and deep understandings over time. Additionally, it was our hope that this study would provide
support for the idea that curriculum coherence may increase student achievement and support students in
building integrated understandings of core chemistry ideas over time. To test this, we administered assessment
items developed and aligned with a learning progression for the transformation of matter. In developing the
items, the sequence of ideas in IQWST was not considered. As such, these assessment items serve as an
independent measure of student understanding (i.e., they are not linked to the curriculum), and can be
considered distal from the curriculum and instruction specific to the IQWST classroom.

Four primary findings resulted from this analysis. First, we demonstrate that students’ latent ability
changes across grade level, supporting the hypothesis that coherent curriculum materials can support students in
developing integrated understanding, as evidenced by participants’ developing understanding of core chemistry
ideas. This finding illustrates the value of designing coherent curriculum that supports students, over time, in
building understanding of core ideas. It is important to realize, however, that the IQWST curriculum materials
also made use of other important features, such as driving questions, to connect unit activities. Second, boys and
girls did not perform differently. This differs from previous findings. Previous research has shown that the
science achievement of boys in urban schools often lags behind that of girls (Geier, et. al, 2008). Our findings illustrate that the design features of IQWST, including coherence and the use of driving questions, supports both boys and girls in understanding core ideas of chemistry. A third finding is that IQWST students performed similarly on items aligned to IQWST learning goals (e.g., particle model) and on items not aligned to IQWST learning goals (e.g., electrical forces at the molecular level). It may be that using IQWST enabled students to read and have a general sense of the appropriate response by eliminating non-appropriate responses on the assessments. With a focus on core ideas and practices, the curriculum materials align closely with the new Framework (NRC, 2011). As such, we would hope to see this positive effect. A fourth finding is that teachers and school variables influence student achievement. The results provide feedback for both the design of the curriculum materials and the design of learning progressions. The study has two key limitations. First, we used a cross-sectional design rather than tracking students longitudinally. We are, therefore, continuing to collect data from the 6th and 7th grade cohort in this study for future analysis. Second, teachers’ professional development experience varied. Because IQWST takes an approach to teaching and learning that is new for many, more thorough professional development and support might have led teachers to enact the materials with greater fidelity, which would influence outcomes. Despite these limitations, the results point to the importance of using coherent curriculum materials to support student learning.

Endnotes
(1) The ability scores were normalized so that the mean and S.D of scores was 500 and 100 using the formula:
\[ t_2 = 500 + \frac{100}{SD_1} (t_1 - m_1) \]
where \( t_1 \) is the latent student’s ability parameter, and \( m_1 \) and \( SD_1 \) are its mean and S.D.

References


