An Analysis of Field Test Results for Assessment Items Aligned to the Middle School Topic of Atoms, Molecules, and States of Matter

Cari F. Herrmann Abell and George E. DeBoer
AAAS Project 2061

NARST Annual Conference
Baltimore, MD

March 30 to April 2, 2008

In this paper we report the results of a field test of assessment items aligned to the middle school chemistry topic of atoms, molecules, and states of matter administered to 3750 sixth, seventh, and eighth grade students of diverse backgrounds in the spring of 2007. Our focus here is on the misconceptions that students reveal and their grade-to-grade growth in understanding of the targeted science ideas. This work is part of a larger multi-year project funded by the National Science Foundation to develop an online collection of assessment items that are precisely aligned with national content standards in middle school science. Each item is developed using a procedure designed to evaluate an item's match to important science ideas and its overall effectiveness as an accurate measure of what students do and do not know about those ideas. During item development, pilot testing is used to obtain feedback from students about the items. Then scientists and science education experts review the items using criteria that ensure content alignment and construct validity. After revisions are made based on the reviews, the items are field tested on a large national sample to determine the psychometric properties of the items and clusters of items.

Introduction

As K-12 educators prepare for widespread testing in science mandated by the *No Child Left Behind* legislation (2002), concerns are growing about the quality of science assessments and their alignment to state and national content standards (American Federation of Teachers, 2006). In response to these concerns, AAAS Project 2061 is developing an online collection of assessment items in science that are precisely aligned with those standards. Each item is developed and analyzed using a procedure designed to evaluate an item's match to important science ideas and its overall effectiveness as an accurate measure of what students do and do not know about those ideas. (See DeBoer et al., 2007 for a more detailed discussion of the procedure.) This paper describes the field test results of a set of assessment items aligned to the middle school topic of atoms, molecules, states of matter, and changes of state. It includes a comparison of how well students in our sample did on each idea and a grade-to-grade, growth-of-understanding, analysis.

Methodology

The field testing reported on here included 3750 students in grades six through eight. The sample was generated through the help of the Building a Presence Program of the National Science Teachers Association. A subset of the teachers who volunteered for field testing was selected in the order they responded to the invitation. A total of 48 teachers from 19 states participated in the field testing. Approximately 26% of the students were in sixth grade, 37% in seventh grade, and 37% in eighth grade. Students from a wide range of urban, suburban, and rural school districts across the country responded to the items. Approximately 41% of the students were students of color, and 12% of the students indicated that English was not their primary language. Half of the students were female and half were male. The field test was administered in the spring of 2007.

Because we were testing more items than students could complete in a typical class period, four different test forms were created that contained subsets of the available questions. Also, because we were interested in describing both the students' understanding of the topic as a whole and their understanding of the ideas within each topic, two of the test forms included a random selection of all of the items and two of the test forms included items from selected clusters of the ideas being tested. In addition, for each form of the test, half of the students took the items in reverse order so that the last items on the test would not be disproportionately omitted if students ran out of time.

During field testing, students answered multiple choice questions related to seven key ideas. The ideas are based on Chapter 4, Section D of *Benchmarks for Science Literacy* (*BSL*) (AAAS, 1993) and Physical Science Content Standard B of *National Science Education Standards* (*NSES*) (NRC, 1996). The seven key ideas are:

Idea A: All matter is made of atoms.
Idea B: All atoms are extremely small.
Idea E: All atoms and molecules are in constant motion.
Idea F: There are differences in the spacing, motion, and interaction of atoms and molecules that make up solids, liquids, and gases.
Idea G: For any single state of matter, changes in temperature typically change the average distance between atoms or molecules. Most substances or mixtures of substances expand when heated and contract when cooled.
Idea H: Changes of state can be explained in terms of changes in the arrangement, motion, and interaction of atoms and molecules.
Idea I: For any single state of matter, the average speed of the atoms or molecules increases as the temperature of a substance increases and decreases as the temperature of a substance decreases.

Each key idea was further clarified in order to state precisely what students would be expected to know. These clarification statements act as item writing specifications that ensure a close alignment between the items and the learning goals. For example, the clarification statement for Idea A says:

> Students are expected to know that matter is anything that has mass and takes up space. They should know that matter includes all gases, liquids, and solids, which make up all living and non-living things. They should know that light, heat, and electricity are not matter. Students are also expected to know that all matter—solids, liquids, and gases—is made up of discrete particles (atoms), rather than being continuous, and that these atoms *are the matter* rather than *contained in matter*. In other words, the atoms are not floating or embedded in some other substance, such as air or a liquid. Additionally, students are expected to know that matter can exist even when it cannot be seen. For example, they are expected to know that gases or vapors are matter even though some of them cannot be seen. Students should also know that anything made up of atoms is matter. Students are expected to know that it is because atoms take up space and have mass that all matter takes up space and has mass. It is not expected that students will know ideas about the internal structure of atoms or the existence of subatomic particles, which are included in benchmarks 4D/H1 and 4D/H2.

During the development of the assessment items, student misconceptions were incorporated into the distractors. (See Sadler, 1998 for a discussion of distractor-based multiple choice testing.) The following is a list of the main misconceptions that were tested.

1. Biological materials, such as cells, are not matter (Novak et al., 1991; Stavy, 1991). Used as a distractor in one of the items.
2. Atoms or molecules of a solid are not moving when no motion of the substance is visible (Novak et al., 1991; Lee et al., 1993). Used as a distractor in seven of the items.
3. Atoms or molecules of solids are hard and atoms or molecules of gases and liquids are soft (Novak et al., 1991; Lee et al., 1993). Used as a distractor in five of the items.
4. The molecules of the gas state are the lightest and the molecules of the solid state are the heaviest (Griffiths et al., 1992). Used as a distractor in two of the items.
5. Solid substances do not expand or contract with changes in temperature (AAAS Project 2061 Pilot testing, 2006). Used as a distractor in three of the items.
6. Heat is made of “heat molecules” (Berkheimer et al., 1988; Lee et al., 1993). Used as a distractor in five of the items.
7. The number of atoms or molecules of a substances increases when the temperature increases (AAAS Project 2061 Pilot testing, 2006). Used as a distractor in three of the items.
8. The mass of the atoms or molecules of a substance decreases when the temperature increases (AAAS Project 2061 Pilot testing, 2006). Used as a distractor in four of the items.
9. The molecules of a substance break down into individual atoms during a change of state. For example, molecules of water become atoms of hydrogen and oxygen when water boils (Osborne et al., 1983; Renstrom et al., 1990). Used as a distractor in four of the items.
10. Molecules change weight/mass during a change of state (Griffiths et al., 1992; Chang, 1999). The weight of a substance can change when a liquid changes to a gas (Stavy, 1990). Used as a distractor in two of the items.

Findings

Because we wanted to make grade-to-grade comparisons in student performance and because our sampling procedures did not ensure that the sixth, seventh, and eighth grade students could be considered equivalent, an analysis of covariance was performed controlling for a number of factors including gender, whether the students identified themselves as non-Hispanic white vs. non-white, whether English was the student's primary language, and the student's performance on a linking item from Idea A (All matter is made of atoms) that appeared on all of the test forms. The estimated marginal means for the overall percent correct for all items combined are reported in Table 1. Analysis of covariance showed that differences in performance by grade are significant at the 0.01 level of significance (F=6.93). A Bonferroni post hoc test showed that all of the grade-to-grade comparisons were significant on the 0.01 level. The greatest difference was between sixth grade students (37.0% correct) and seventh grade students (45.3% correct). Rather than seeing a continued increase in percent correct from seventh to eighth grade, we found a statistically significant decrease from 45.3% correct to 40.4% correct.

Table 1
Overall Percent Correct by Grade

Grade	6th Grade	7th Grade	8th Grade
% correct (N)	37.0% (950)	45.3% (1368)	40.4% (1378)

We then repeated the analysis described above at the individual idea level. The results of this analysis, reported in Table 2, are based on all of the students who answered an item for a particular idea. A separate analysis of covariance was conducted for each idea, again controlling for gender, whether the student self identified as non-Hispanic white vs. non-white, whether English was the student's primary language, and the student's performance on a linking item from Idea A (All matter is made of atoms) that appeared on all of the test forms.

Overall, the results indicate that students had the most difficulty with the idea of thermal expansion and with the idea that atoms are always in motion. Students were most successful with questions testing the idea that all matter is made of atoms and that atoms are extremely small. No attempt was made to statistically rank student performance across the seven ideas.

Table 2
Field Test Results by Idea

Key Idea	6th grade % (N)	7th grade % (N)	8th grade % (N)	Total % (N)
Idea A: All matter made of atoms	52.5% (950)	54.9% (1368)	55.8% (1378)	54.5% (3696)
Idea B: Extremely small	46.2% (722)	53.4% (1004)	49.8% (1023)	50.0% (2749)
Idea E: Constantly in motion	30.7% (946)	34.6% (1360)	31.5% (1375)	32.3% (3681)
Idea F: States of Matter	39.0% (950)	54.4% (1367)	43.1% (1377)	45.7% (3694)
Idea G: Thermal Expansion	29.5% (950)	34.9% (1368)	35.5% (1377)	33.5% (3695)
Idea H: Changes of State	31.3% (722)	40.8% (1006)	36.1% (1025)	36.2% (2753)
Idea I: Speed Changes with Temperature	41.0% (949)	48.1% (1365)	43.3% (1375)	44.3% (3689)

The differences in the percentage correct are significant for Idea F (states of matter) ($F=6.97$, $p<0.01$), and Idea H (changes of state) ($F=4.27$, $p<0.05$). For Idea F, post hoc comparisons show that the seventh and eighth grade students performed significantly better than the sixth grade students. For Idea H, the seventh and eighth grade students performed significantly better than the sixth grade students, and the percent correct for the seventh grade students was significantly higher than the percent correct for the eighth grade students.

We also conducted a distractor-level analysis to determine which of the misconceptions listed above were the most common for these students at each grade level. The results below show the relative strength of a wide range of misconceptions and grade-to-grade comparisons on each of these misconceptions. As with the idea-level comparisons made above, a separate analysis of covariance was conducted for each misconception, controlling for gender, whether the student self identified as non-Hispanic white or non-white, whether English was the student's primary language, and student performance on a linking item from Idea A (All matter is made of atoms) that appeared on all of the test forms (see Table 3). For all of the misconceptions listed, the difference in performance among the grades is not significant. No attempt was made to statistically rank the strength of the misconceptions.

Table 3
Percentage of Students Holding Misconceptions

Misconception	6th grade % (N)	7th grade % (N)	8th grade % (N)	Total % (N)
1. Cells are not matter.	33.7% (950)	28.0% (1368)	27.9% (1378)	29.7% (3696)
2. Atoms of a solid are not moving.	35.3% (950)	26.6% (1367)	36.2% (1377)	32.6% (3694)
3. Atoms of solids are hard and atoms of gases and liquids are soft.	23.8% (950)	16.5% (1367)	17.2% (1377)	19.0% (3694)
4. Molecules of the gas state are the lightest and molecules of the solid state are the heaviest.	16.0% (938)	13.5% (1361)	18.1% (1372)	15.9% (3671)
5. Solid substances do not expand or contract with changes in temperature.	52.6% (946)	43.9% (1367)	47.8% (1370)	47.9% (3683)
6. Heat is made of "heat molecules."	28.4% (950)	26.7% (1367)	25.9% (1376)	26.9% (3693)
7. The # of atoms of a substances increases when the temperature increases.	21.5% (943)	24.5% (1363)	19.5% (1367)	21.8% (3673)
8. The mass of the atoms of a substance decreases when the temperature increases.	19.0% (949)	17.0% (1368)	18.0% (1375)	18.0% (3692)
9. The molecules of a substance break down into individual atoms during a change of state.	36.1% (722)	37.6% (1007)	35.4% (1076)	36.4% (2805)
10. Molecules change weight/mass during a change of state.	18.9% (716)	17.7% (1002)	23.2% (1021)	20.0% (2739)

Discussion

One of the most interesting and surprising findings of our field testing was the overall trend in performance that showed seventh grade students performing significantly better than sixth grade students and the performance of eighth grade students falling off compared to that of seventh grade students. The most obvious explanation for this finding would be that more of the students in our sample studied chemistry in the seventh grade than in the sixth or eighth grade. Although

we do not have conclusive evidence of this, information reported by the teachers suggests that more of our seventh grade students were taught these ideas than were sixth or eighth grade students. In addition, a number of the sixth grade teachers indicated that although their students were studying states of matter and changes of state, they were not studying them at the atomic/molecular level. Often these ideas are taught at the substance level in the sixth grade (e.g., solids have a fixed shape and volume, liquids take the shape of the container and have a fixed volume, and gases take the shape of the container and fill the volume of the container). This may explain why the sixth grade students performed poorly on the items about the states of matter and changes of state when compared to the performance of the seventh and eighth grade students.

It is also interesting to note that even though what students know seems to follow a pattern that shows improvement from sixth to seventh grade but not from seventh to eighth grade, the misconceptions did not follow the same pattern. There is much more similarity from grade to grade in the misconceptions that students held. Perhaps students are letting go of less tightly held misconceptions as they are learning chemistry ideas, but the misconceptions identified in this study persist throughout middle school.

Conclusions

Field testing on a large national sample provides valuable information about students' content knowledge and gaps in that knowledge. In this study, we found that students in middle school were most familiar with the ideas that all matter is made of atoms and that atoms are very small. They had the most difficulty with the idea of thermal expansion and with the idea that atoms are always in motion.

A grade-to-grade comparison study revealed that the sixth grade students had significantly more difficulty with ideas related to states of matter and changes of state than the seventh and eighth grade students. The comparatively low performance of sixth grade students on these two ideas in particular can probably be explained by the fact that most sixth grade students are just beginning their studies of atoms and molecules and often have not been instructed on the atomic/molecular level explanations of states of matter and changes of state.

The most puzzling finding was that, on a number of ideas, eighth grade students performed less well than seventh grade students and only slightly better than sixth grade students. Perhaps that was due to our sampling procedures, but we statistically controlled for relevant variables that should have equated the samples for the three grades. If not due to differences in the samples, it appears that most of the knowledge gains made by the seventh grade students were lost by eighth grade. It may be that the knowledge was too insecure to be retained without further instruction. This would suggest that without repeated instruction on those ideas throughout the middle grades, the ideas are too easily forgotten.

We also investigated the prevalence of a number of specific misconceptions related to the topic of atoms, molecules, and states of matter. Testing multiple misconceptions on the same population of students allows us to compare the relative strength of those misconceptions and to

ascertain if these results are consistent with results previously reported in the research literature. We found that all of the major misconceptions reported in the literature were confirmed through our field testing. The strongest of them appear to be the ideas that atoms of solids are not moving (32.6%) and that molecules break down into individual atoms during a change of state (36.4%). In addition, we identified three other misconceptions not previously reported. These are the ideas that solid substances do not expand or contract with changes in temperature (47.9%), the idea that the number of atoms of a substances increases when the temperature increases (21.8%), and the idea that the mass of the atoms of a substance decreases when the temperature increases (18.0%). Finally, we found that the strength of the misconceptions did not significantly change from grade to grade, suggesting that these misconceptions are resistant to the instruction students are currently receiving about atoms, molecules, and states of matter.

This work is important to science educators because assessment items that are closely aligned to the content standards can be used in science education research to more effectively measure what students know about specific ideas in science. The work is also important because it reveals in considerable detail what middle school students currently know about the ideas in national standards documents and how well we are achieving the goal of science literacy for all.

Acknowledgements

This work is funded by the National Science Foundation National (Grants # ESI 0227557 and ESI 0352473).

References

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

American Association for the Advancement of Science. (2006). Unpublished results of pilot testing.

American Federation of Teachers. (2006). Smart testing: Let’s get it right (Policy Brief No. 19). Washington, DC: Author.

Berkheimer, G. D., Anderson, C. W., Lee, O., & Blaskeslee, T. D. (1988). *Matter and molecules teacher's guide: Science book*. East Lansing, Michigan: Michigan State University.

Chang, J. Y. (1999). Teachers college students' conceptions about evaporation, condensation, and boiling. *Science Education*, *83*, 511–526.

DeBoer, G. E., Gogos, A., & Herrmann Abell, C. F. (2007, April). *Assessment linked to science learning goals: Probing student thinking during item development.* Paper presented at the

annual conference of the National Association for Research in Science Teaching, New Orleans, LA.

Griffiths, A. K., & Preston, K. R. (1992). Grade-12 students' misconceptions relating to fundamental characteristics of atoms and molecules. *Journal of Research in Science Teaching*, *29*(6), 611–628.

Lee, O., Eichinger, D. C., Anderson, C. W., Berkheimer, G. D., & Blaskeslee, T. D. (1993). Changing middle school students' conceptions of matter and molecules. *Journal of Research in Science Teaching*, *30*(3), 249–270.

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

No Child Left Behind Act of 2001, 20 U.S.C. § 6301 *et seq.* (2002).

Novak, J. D., & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal*, *28*(1), 117–153.

Osborne, R. J., & Cosgrove, M. M. (1983). Children's conceptions of the changes of state of water. *Journal of Research in Science Teaching*, *20*(9), 825–838.

Renstrom, L., Andersson, B., & Marton, F. (1990). Students' conceptions of matter. *Journal of Educational Psychology*, *82*(3), 555–569.

Sadler, P. M. (1998). Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments. *Journal of Research in Science Teaching*, *35*(3), 265–296.

Stavy, R. (1990). Children's conceptions of changes in the state of matter: From liquid (or solid) to gas. *Journal of Research in Science Teaching*, *27*, 247–266.

Stavy, R. (1991). Children's ideas about matter. *School Science and Mathematics*, *91*(5), 240–244.