

7-2011

Assessing Secondary and College Students' Implicit Assumptions about the Particulate Nature of Matter: Development and Validation of the Structure and Motion of Matter Survey

Marilyne Stains

University of Nebraska-Lincoln, mstains2@unl.edu

Marta Escriu-Sune

University of Massachusetts Boston

Myrna Lisseth Molina Alvarez de Santizo

Universidad Galileo and American School, Guatemala City, Guatemala

Hannah Sevia

University of Massachusetts Boston

Follow this and additional works at: <http://digitalcommons.unl.edu/chemfacpub>

 Part of the [Chemistry Commons](#)

Stains, Marilyne; Escriu-Sune, Marta; Molina Alvarez de Santizo, Myrna Lisseth; and Sevia, Hannah, "Assessing Secondary and College Students' Implicit Assumptions about the Particulate Nature of Matter: Development and Validation of the Structure and Motion of Matter Survey" (2011). *Faculty Publications -- Chemistry Department*. 57.
<http://digitalcommons.unl.edu/chemfacpub/57>

This Article is brought to you for free and open access by the Published Research - Department of Chemistry at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications -- Chemistry Department by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Assessing Secondary and College Students' Implicit Assumptions about the Particulate Nature of Matter: Development and Validation of the Structure and Motion of Matter Survey

Marilyne Stains,[†] Marta Escriu-Sune,^{‡,§} Myrna Lisseth Molina Alvarez de Santizo,^{||} and Hannah Sevian^{*,†,||,§,○}

[†]Department of Chemistry, University of Nebraska—Lincoln, Lincoln, Nebraska 68588, United States

Departments of [‡]Biology, [†]Chemistry, and ^{||}Curriculum and Instruction and [§]Center of Science and Mathematics in Context, University of Massachusetts Boston, Boston, Massachusetts 02125, United States

^{||}Universidad Galileo and American School, Guatemala City, Guatemala

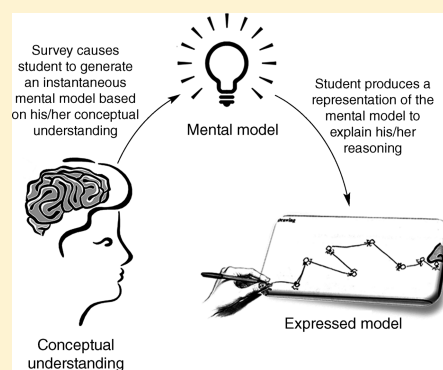
[○]Division of Undergraduate Education and Division of Research on Learning in Formal and Informal Settings, National Science Foundation, Arlington, Virginia 22230, United States

S Supporting Information

ABSTRACT: Development of learning progressions has been at the forefront of science education for several years. While understanding students' conceptual development toward "big ideas" in science is extremely valuable for researchers, science teachers can also benefit from assessment tools that diagnose their students' trajectories along the learning progressions. In this paper, we describe the development and validation of a teacher-friendly survey, the Structure and Motion of Matter (SAMM) survey, designed to measure students' trajectories along aspects of a research-based learning progression on the particulate nature of matter. Specifically, the survey assesses students' implicit assumptions about four concepts: the structure of solute and solvent substances in a gas solution, the origin of motion of gaseous solute particles, and their trajectories. The process to ensure the translation validity (face and content validity) of the survey is described. Criterion validity study results indicate that the SAMM survey is well grounded in theory, and the test–retest study results indicate that the survey is also reliable. Finally, the development of an Excel-based scoring scheme associated with the SAMM survey is also described. Inter-rater reliability studies indicate that the scoring scheme can be used reliably.

KEYWORDS: Elementary/Middle School Science, First-Year Undergraduate/General, High School/Introductory Chemistry, Chemical Education Research, Testing/Assessment, Gases, Kinetic-Molecular Theory, Learning Theories, Solutions/Solvents

FEATURE: Chemical Education Research



How students *should* progress toward understanding "big ideas" in science (i.e., concepts that are the foundations of the discipline and necessary for future learning) has long been of interest to the science literacy movement.^{1,2} More recently, researchers have placed an emphasis on characterizing how students *actually* develop conceptual understanding of these ideas.^{3,4} Such characterizations, termed learning progressions (LPs), were recently defined by a consensus among researchers to be research-based informed hypotheses of how actual understanding of a "big idea" develops over a long period of time, multiple grades, or educational levels.⁵

The particulate nature of matter (PNM) is one of these "big ideas".^{3,4} Recently, several science education researchers have worked on developing LPs that describe how learners should advance in their understanding of specific constructs of the PNM, such as the atomic–molecular theory⁴ and models of the structure,

behavior, and properties of matter.^{3,6} The research on LPs about the PNM builds on several decades of science education research on many different aspects of students' understanding of matter.^{7–14} Talanquer¹⁵ compiled, analyzed, and organized results of a vast array of this research into a hypothetical LP. His LP does not describe how students' understanding about matter should progress over a specific period of time (e.g., one grade); rather, it describes the trajectories students have been found to take in their paths toward understanding the "big idea", independent of grade level and curriculum experienced. Specifically, his LP describes the development of students' implicit assumptions about the structure of matter, its properties, dynamics, and interactions that constrain their thinking. Certain

Published: July 19, 2011

aspects of this LP, in particular the structure of matter and dynamics, are the topics targeted in the survey presented here. Specifically, our survey is designed to measure the progress in implicit assumptions about the Structure and Motion of Matter (SAMM) that constrain students' reasoning about phenomena emerging from particles' interactions. The LP presented in Talanquer's article¹⁵ was the starting point for the development of the SAMM survey and its associated scoring scheme.

Having access to well-researched and clearly defined LPs of core science concepts is invaluable knowledge for science teachers. It can assist them in planning activities that address critical understandings toward achieving conceptual understanding of the "big idea" that anchors an LP at the upper end and in interpreting formative assessments of students' progress toward that anchor. Thus, assessing where students stand and how they make progress along LPs is critically important. During the last several years, the science education community has developed research-based LPs for a wide variety of concepts in different science domains: see, for example, the special issue of the *Journal of Research in Science Teaching*.¹⁶ The development of LPs has gone hand-in-hand with the development of assessment tools that help characterize where students stand along LPs. However, few researchers have developed assessment tools that can be implemented easily by science teachers in their own classrooms. Most assessment systems are Web-based^{3,17} and rely on ordered multiple-choice items, which can become problematic in characterizing specific students' level of understanding when students answer inconsistently to different items designed to measure the same concept.^{18–20} In contrast to online assessment systems, the SAMM survey is a paper-based instrument with open-ended questions that instructors can easily and reliably score using a scoring scheme in the format of a programmed Excel spreadsheet (see the online Supporting Information). Open-ended items can overcome the weakness associated with the multiple-choice assessment format by providing rich qualitative information (e.g., explanations and drawings), which can be used to make sense of inconsistencies.

In this paper, we describe the development and validation of the SAMM survey, designed to assess students' mental models of diffusion and the implicit assumptions about the structure of solute and solvent substances in a gas solution, the origin of motion of gaseous solute particles, and their trajectories that constrained these mental models. The survey was designed to be used as both a formative and diagnostic assessment tool for science educators, as well as a research instrument. From a teaching practice perspective, our intent was to develop a survey that can be implemented easily by instructors in secondary and postsecondary classrooms, and can offer detailed information on students' assumptions about the structure and motion of matter. From a research perspective, we developed this tool because we are interested in exploring the ranges of students' mental models of diffusion and the implicit assumptions about the structure and motion of matter that constrain these mental models. The goal of this paper is to describe the development of the survey and its validation. Findings related to students' mental models about diffusion and the implicit assumptions that constrain these mental models will be presented in a later paper.

PHASES OF THE STUDY

Pilot Studies

The development and validation of the SAMM survey and associated scoring scheme involved several waves of pilot studies and a full implementation of the final version of the survey.

Table 1. Student Participants in the Pilot Studies

School	Level or Course	Pilot Study 1 ^a	Pilot Study 2 ^a	Total
Middle school	8th grade	65	86	151
High school	Regular	14 (B)	—	14
	AP	16 (B)	35 (C, P)	51
University	Freshman	62 (C)	137 (B, C, P)	199
Total		157	258	415

^aB is biology; C, chemistry; P, physics.

The two pilot studies conducted with students from 8th grade to the first-year level in college had four purposes, aiming to:

1. Control the length of the survey (no longer than 15 min)
2. Ensure that questions in the survey were understandable by a wide variety of students with different cultural and linguistic backgrounds and different preparation in science
3. Test and identify questions that provide rich information about students' understanding of diffusion and thinking about matter
4. Test and refine the scoring scheme

Full Implementation and Interviews

Implementation of the final version of the survey allowed us to conduct further validity studies on the final product. Interviews were also conducted at the pilot and full implementation stages with purposeful samples of students ($N = 25$). The goals of these interviews were to ensure that:

1. Questions in the survey were understandable by a wide variety of students with different cultural and linguistic backgrounds and different preparation in science
2. Our analysis of students' answers on the survey reflected what students intended to say
3. The scoring scheme represented an accurate description of all the different ways of thinking students demonstrated on the survey

PARTICIPANTS

All of the studies described above were conducted with volunteer participants from a large urban school district and a public research university. Both settings serve a highly diverse student population, with high proportion of minorities (i.e., Hispanic, African American, and low socio-economic status students) and immigrants from the Caribbean and Africa. Description of the students who participated in the pilot studies and the educational experts who were consulted are provided in Tables 1 and 2, respectively. Once the survey was finalized through the pilot studies and interviews, we collected a total of 647 surveys using the final version of the survey from a large pool of students with different levels of schooling. Of these, 471 surveys were complete and valid for analysis (see Table 3).

INSTRUMENT DESIGN

Conceptual understanding is best assessed through generative questions, which require students to spontaneously create mental models based on existing understanding and use these models to make sense of the situation rather than relying on recall and memorization.²¹ Traditionally in science education, mental models have been characterized through the analysis of expressed models. Mental models are private and personal cognitive representations, while expressed models are models derived

Table 2. Educational Expert Participants

		Scoring Scheme Development ^a	Content Validity ^a	Total
Science teachers	Upper-elementary school	3	—	3
	Middle school	10	—	10
	High school	16 (AP-C) 33 (B, C, P)	—	49
Content experts	Graduate students	—	26 (C)	26
	Faculty	—	3 (B, C, P)	3
Total		62	29	91

^a AP-C is AP chemistry; B, biology; C, chemistry; P, physics.

Table 3. Student Participants in the Full Implementation of the SAMM Survey

School	Level or Course	Biology	Chemistry	Physics	Total
Middle school	8th grade	N/A	N/A	N/A	79
High school	Regular	54	22	42	118
	Honors	0	30	58	88
	AP	21	26	23	70
University	First year	24	61	31	116
Total					471

from the mental models placed in the public domain through the use of different modes of representation. The survey design followed this assessment strategy as Figure 1 illustrates.

Science education researchers have argued about the nature of the relationship between expressed and mental models.²² In this study, expressed models are the explanations students write and draw to answer problems posed in the survey. We do not consider them to be a complete, in-depth representation of students' conceptual understanding of the topic. Rather, we think of them as the understanding at the front of students' minds and as a representation of the assumptions that guide their thinking at the time of assessment.

With the following constraints in mind, the instrument was designed to:

- Be open-ended
- Be teacher-friendly (short, with concepts targeted relevant to secondary and college science courses)
- Be understandable by students from 8th grade to college with various cultural and linguistic backgrounds
- Be valid and reliable
- Provide clear evidence of students' mental models of diffusion
- Help identify students' implicit assumptions about the following four concepts: (i) the structure of a gaseous solute (perfume); (ii) the structure of a gaseous solvent (air); (iii) the origins of motion of solute particles (why do molecules of perfume move?); and (iv) the trajectories of solute particles (how do molecules of perfume move?)
- Stimulate an instantaneous mental model representing the student's conceptual understanding, rather than cause rote recall of previously learned explanations that may not have been fully understood

The final version of the SAMM survey is a 15-min, open-ended instrument containing three multipart questions (see the online Supporting Information). In the multipart questions, students are asked to explain how perfume molecules traverse a room in

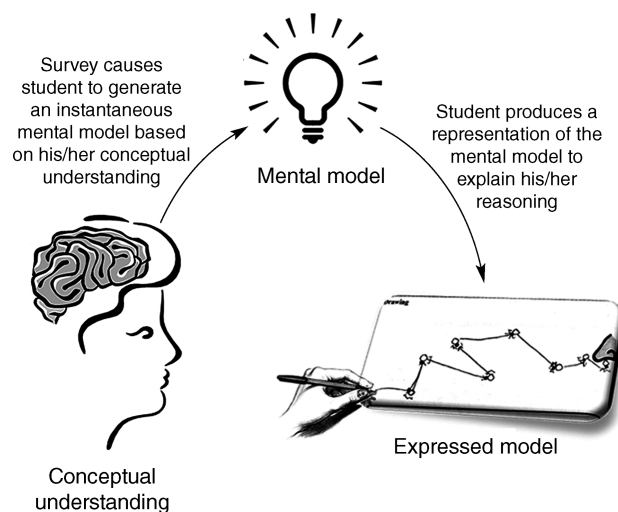


Figure 1. Process of capturing conceptual understanding through generative questions.

various situations through drawings and by answering follow-up questions to explain different aspects of the drawings. This particular situation of diffusion was chosen over many other options considered (e.g., diffusion of a drop of dye into a liquid) because it was not used as an example in any of the curriculum materials in grades 6 through the undergraduate level in the school and university system from which participants were drawn in this study. A detailed description of the relationships between questions in the survey and the four concepts targeted is provided in the online Supporting Information. A Flesch–Kincaid test of readability indicates that the survey is at a 4th grade reading level.

SCORING SCHEME DEVELOPMENT

The scoring scheme was also designed under specific constraints and intended to:

- Reflect a research-based learning progression
- Identify the implicit assumptions students hold about the four concepts described above
- Represent the majority of students' ways of thinking on the SAMM survey
- Be informative by providing enough details for instructors to be able to plan learning activities accordingly
- Be clear and easy to use by teachers
- Be reliable

The scoring scheme, which evaluates the implicit assumptions about the four concepts described above that constrain students'

Table 4. Implicit Assumptions' Categories and Succinct Descriptions for the Origin of Motion of Solute Particles Concept

Category	Implicit Assumptions
0	Not coherent: Student does not express any relevant information or the expression is so confusing as to be not interpretable
1	Motion of molecules is caused by external forces: Student assumes that external forces are the cause of particles' movement
2	Molecules' motion is conditioned by certain properties or features of the substance but external forces are the main agent of movement: Student recognizes implicitly that molecules move but still mentions an external force as the cause of movement
3	Molecules' motion is conditioned by certain properties or features of the substance: Student expresses the idea that molecules move by themselves but does not explicitly say that they are in constant motion
4	Molecules' motion is conditioned by certain properties or features of the substance AND molecules are always in motion
5	Molecules are always in motion: Student explicitly recognizes that molecules' motion is an intrinsic property of the molecules

Table 5. Implicit Assumptions' Categories and Succinct Descriptions for the Structure of Gaseous Solvent (Air) Concept

Category	Implicit Assumptions
0	Not coherent: Student does not express any relevant information or the expression is so confusing as to be not interpretable
1	Ignored/Absent: Student never mentions or draws air
2	Macroscopic: Student assumes that the air is continuous
3	Macro/Microscopic: Student assumes that air is continuous sometimes and particulate at other times
4	Microscopic: Student assumes that air is particulate

expressed models of diffusion, was initially developed based on a previously published LP about the PNM.¹⁵ In developing this LP, Talanquer reviewed a large body of science education research on students' understanding of the PNM with a focus on longitudinal and cross-sectional studies about students' development of understanding of matter, and studies in cognitive science and developmental psychology about cognitive constraints. Through the analysis of these studies, he identified implicit assumptions that constrain students' reasoning about matter. He summarized these findings in an LP describing the progression of these cognitive constraints along four dimensions related to the understanding of matter:¹⁵

1. The structure of matter: toward the idea that matter is particulate
2. The properties of matter: toward the idea that properties of a substance emerge from particles' interactions
3. The dynamics of matter: toward the idea that particles' motion is an intrinsic property of particles
4. The interactions of matter: toward the idea that particles interact through intrinsic forces

In developing our scoring scheme, we refined the categories describing cognitive constraints that Talanquer identified by analyzing surveys collected during the two pilot studies (see Table 1) and interviews with volunteer participants from the pilot and full implementation studies. The analysis of these data through grounded theory allowed us to identify and characterize categories describing cognitive constraints that were not present in Talanquer's LP. For example, many students relied on a mixture of two cognitive constraints from Talanquer's LP of the origin of motion (category 1 and 3 in Table 4). We thus added this transition as a separate category (category 2). We incorporated these new categories into the scoring scheme and tested it with a new population of students with different characteristics than the previous one (e.g. different grade level). This iterative process was followed until all implicit assumptions that different populations of students demonstrated were accounted for. The scoring scheme that emerged from this analysis is thus grounded in theory and heavily relies on empirical data.

A short description of the categories for two of the four concepts targeted in the SAMM survey is provided in Tables 4 and 5. A more detailed description of each category and indicators in students' answers to the SAMM survey that characterize them is provided in the online Supporting Information for all four concepts.

We originally developed a holistic version of the scoring scheme, in which scorers were asked to look for consistency in several different indicators throughout the survey (e.g., does the student draw circles in question 1a and write the word molecules/particles in questions 1b, 1c, and 1e). However, the pilot testing of this version with science teachers who had various experiences in teaching science (see Table 2) indicated that teachers, regardless of their teaching experience, prefer a scoring scheme that evaluates students' answers to individual questions on the survey. We thus developed a programmed Excel spreadsheet (see the online Supporting Information and the attachments at ref 23 for a Google Docs version) based on the logic of the holistic scoring scheme to address this issue and to minimize variation due to scorers' subjectivity. In the spreadsheet, students' answers are scored according to the presence or absence of indicators that are associated with each category, which is measured through a series of yes/no questions about each question in the SAMM survey (e.g., did the student draw dots, circles, or use letters to represent the perfume?). To minimize misinterpretation of some of the yes/no questions, we included examples of "yes" answers taken from surveys that were collected during the full implementation of the survey. The spreadsheet is programmed using a sequence of logical functions (e.g., IF, AND, OR) to derive a category for each concept based on the answers to these yes/no questions. The spreadsheet was tested to ensure that it provided the same results as the holistic scheme.

■ SCORING SCHEME RELIABILITY

Two of the authors (M.E.-S. and M.S.) scored independently 12% ($N = 55$) of the surveys collected from the full implementation of the survey (see Table 3). These surveys were selected randomly and represented a variety of school levels and subject

Table 6. Comparative Results for Inter-Rater Reliability on Four Concepts Targeted in the SAMM Survey

SAMM Concept Measured	Weighted <i>K</i> Coefficient Values (<i>N</i> = 55)	
	Before Sharing Scores	After Sharing Scores
Structure of gaseous solute	0.78	1.00
Structure of gaseous solvent	0.86	0.97
Origin of motion of solute particles	0.51	0.81
Trajectories of solute particles	0.60	0.85

matters. Following independent scoring session, scores were shared. For each disagreement, each scorer rescored the survey independently, double-checking for omissions. Weighted Cohen's *K* coefficients were calculated before and after viewing each other's scores in order to determine the inter-rater reliability.²⁴ On the basis of the accepted interpretation of *K* values,²⁵ results indicate a moderate to almost perfect agreement before scores were shared and almost perfect agreement after re-evaluation of inconsistencies (see Table 6). Most disagreements were due to one of the raters missing a detail that the student wrote or drew (e.g., faint dots in a drawing).

In an editorial published in this *Journal*,²⁶ Bunce explained the importance of testing the validity and reliability of a research instrument. In the following sections, we describe the different validity and reliability tests that were performed on the SAMM survey.

■ VALIDITY STUDIES FOR THE SAMM SURVEY

Translation Validity

Translation validity tests ensure that findings emerging from data analysis adequately represent the theoretical construct on which the instrument is based.²⁷ For the present survey, students' answers need to provide a comprehensive view of students' mental model of diffusion and implicit assumptions about the structure and motion of matter.

Face Validity

Face validity is concerned with the impression that the instrument can collect the needed information. We piloted the first version of the survey by asking a group of students representative of our targeted population (see Table 1, Pilot Study 1) to answer the survey to the best of their ability and to write comments if the questions or instructions were unclear. We also organized a focus group with participants who had a wide variety of expertise in chemistry, from a chemistry postdoctoral fellow to a person who had taken general chemistry 10 years ago. The goal of this focus group was to ensure that the questions were understandable regardless of the level of preparation in science. From the analysis of these data, we determined the effectiveness of the SAMM survey and improved it. Because of space limitations, details on how the survey was improved as a result of this pilot study, subsequent ones, and interviews are provided in the online Supporting Information.

We piloted the second version of the survey with another set of students representative of the targeted population (see Table 1, Pilot Study 2). Students from the pilot study who volunteered (*N* = 13) were presented with their own surveys and were asked

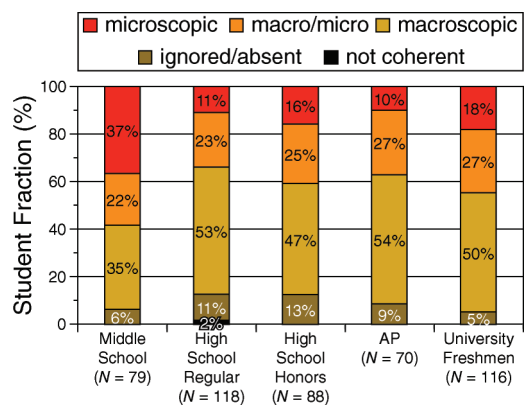


Figure 2. Distribution of students among the five categories of implicit assumptions about the structure of air.

to explain their answers during interviews taking place a week after survey completion. These interviews helped us ensure that students with various cultural and linguistic backgrounds understood the questions (wording and graphics) as we intended, and that we were interpreting their answers as they intended. Following analysis of these data, the SAMM survey was further refined into a final version (see the online Supporting Information for details).

Content Validity

An instrument with high content validity contains questions representative of the theoretical construct the instrument is meant to measure. Crucial to a high level of content validity is a clear and detailed definition of the theoretical construct. We designed the questions in the survey based on the Talanquer's LP about the structure of matter and its motion.¹⁵ In particular, the questions are specifically designed to assess students' implicit assumptions about the following four concepts: structure of a gaseous solute (perfume), structure of a gaseous solvent (air), origin of motion of solute particles, and trajectories of solute particles. Expert chemists' assumptions about these concepts are presented in the user manual for the scoring scheme (see the online Supporting Information).

To ensure content validity, we collected surveys from and interviewed a group of science faculty and chemistry graduate students (see Table 2). Analysis of these data confirmed that the survey represented our theoretical construct effectively. The expert chemists also provided feedback on the correct use of language. For example, the chemicals represented in question 3 of the survey are not perfumes, but the gases do have detectable odors, so the language was changed to "scented gas" for accuracy.

Criterion Validity

A survey grounded in theory should provide conclusions that are expected based on the theoretical framework.²⁷ In our case, we expected differences between groups of students who had exposure to different curricula. In the school district where this study took place, 8th graders spend one-third of the year learning about molecules and their movement using the FOSS kit, *Chemical Interactions*.²⁸ We thus expected that 8th graders' expressed models, which represent the understanding at the front of students' minds, would be based on this particulate thinking. Because science curricula taught at the high school and first-year university level focus more on quantitative descriptions of matter, we expected that the particulate way of thinking about matter would not be at the front of these students' minds.

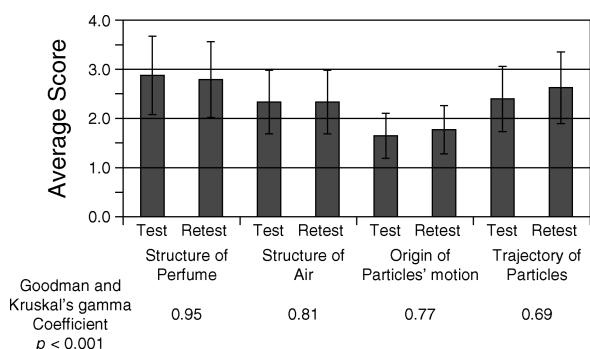


Figure 3. Test–retest results for each of the four concepts targeted in the SAMM survey ($N = 48$).

We thus hypothesized that more middle school students than high school and postsecondary science students would assume that both perfume and air are particulate.

This hypothesis was tested with data collected during the full implementation of the survey (see Table 3). Figure 2 illustrates the distribution of students by schooling level among the five different categories of implicit assumptions about the structure of air (see Table 5). As expected, a significantly larger proportion of middle school students assumed that air is microscopic compared to other groups of students, $\chi^2(16, 471) = 36.61$, $p = 0.002$, $V = 0.14$.

RELIABILITY STUDIES FOR THE SAMM SURVEY

Internal Consistency

Internal consistency has to do with ensuring that constructs are evaluated by more than one item on a survey. This is traditionally evaluated in multiple-choice surveys by calculating the correlation between items that intend to measure the same construct. As the SAMM survey relies on open-ended items, correlation analyses could not be conducted. However, both the survey and the scoring scheme were built in such a way that each of the four concepts are scored based on students' answers to five or six different questions in the survey.

Test–Retest

Survey reliability is traditionally determined through test–retest studies, which evaluate the reproducibility of results obtained from the survey.²⁹ The SAMM survey was administered twice, two weeks apart, with several classes of a second-year level psychology course offered at the public university ($N = 48$). We specifically chose psychology students because they likely do not learn new science related to the content of the survey during the two-week period (concurrent enrollment in a science course was probed in the survey). Thus, students should provide similar answers both times. Figure 3 presents the result of the correlation analysis along with the average score students obtained on each concept on the two occasions. Owing to the nature of the survey and its scoring, Goodman and Kruskal's γ coefficients were calculated instead of Spearman's ρ . The γ coefficient is similar to the Spearman's ρ except that the calculation takes into consideration possibilities for many ties that can arise with categories that have few levels.³⁰ Because our instrument and data set fitted these criteria, the calculation of Goodman and Kruskal's γ coefficients was deemed more appropriate. The correlation coefficients presented in Figure 3 show statistically significant and moderately strong to very strong correlation between the test and retest for all four concepts. These results indicate strong similarities in students'

answers despite the two-week gap and thus they show that the SAMM survey generates reproducible results.

All of the validity and reliability studies were focused on the structure of the instrument. The validity of the survey in term of its purpose (e.g., diagnostic tool vs formative assessment) is being evaluated in ongoing studies.

APPLICATIONS OF THE SAMM SURVEY

The SAMM survey was designed with both science teachers and science education researchers in mind, and should therefore be useful to both.

Science teachers should be able to use the SAMM survey as a formative assessment to explore their students' thinking and develop classroom activities accordingly. It is our hope that science teachers will find it useful to compare class results on this survey from the beginning of a unit or term to the end, and that this would provide feedback on the dimensions along which the class progressed. At the undergraduate level, chemistry faculty could use the survey as a diagnostic tool that would allow them to evaluate incoming students' assumptions of the structure and motion of matter, concepts that are essential for learning chemistry. Ideally, when preparing a lesson, instructors could filter instruction through a lens of knowing where students stand in their current thinking, to help ensure that the same lesson could serve to advance student thinking from different starting points.

Science education researchers should also be able to use the SAMM survey in multiple contexts. For example, Vosniadou found some differences between children in the U.S., Greece, India, and Samoa, in the mental models they hold about the concept of Earth.³¹ In an analogous fashion, we are currently using the survey to compare whether and how students' assumptions about the PNM differ in different contexts (Spanish-language versions of the survey and scoring scheme are available from the corresponding author). The instrument could also be used in longitudinal studies to characterize the evolution of students' thinking over time. Finally, the SAMM survey could be useful to test the effectiveness of a curriculum or a set of activities designed to teach students about the structure and motion of matter.

ASSOCIATED CONTENT

Supporting Information

Relationships between survey questions and concepts targeted; descriptions of the role of the pilot studies in the SAMM survey design; the SAMM survey; user manual for the Excel scoring scheme and the holistic scoring scheme; Excel scoring scheme. This material is available via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: hannah.sevian@umb.edu.

Notes

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

ACKNOWLEDGMENT

We would like to thank all the participants in this study, and Brian White in the Biology Department at the University of Massachusetts Boston, for many enlightening conversations during the instrument and scoring scheme development. This paper is based on work supported by the National Science Foundation (NSF), while one of the authors (H.S.) was working at the Foundation. This work was also supported, in part, by NSF award EHR-0412390.

REFERENCES

- (1) American Association for the Advancement of Science. *Benchmarks for Science Literacy*; Oxford University Press: New York, 1993.
- (2) American Association for the Advancement of Science. *Atlas of Science Literacy*; American Association for the Advancement of Science: Washington, DC, 2001.
- (3) Claesgens, J.; Scalise, K.; Wilson, M.; Stacy, A. *Sci. Educ.* **2009**, *93* (1), 56–85.
- (4) Smith, C. L.; Wisner, M.; Anderson, C. W.; Krajcik, J. *Meas.: Interdisc. Res. Perspect.* **2006**, *4* (1 and 2), 1–98.
- (5) Duschl, R. A.; Schweingruber, H. A.; Shouse, A. *Taking Science to School: Learning and Teaching Science in Grades K–8*; National Academy Press: Washington, DC, 2007.
- (6) Stevens, S. Y.; Delgado, C.; Krajcik, J. S. *J. Res. Sci. Teach.* **2010**, *47* (6), 687–715.
- (7) Adbo, K.; Taber, K. S. *Int. J. Sci. Educ.* **2008**, *31* (6), 757–786.
- (8) Eilam, B. *J. Res. Sci. Teach.* **2004**, *41* (10), 970–993.
- (9) Gabel, D.; Samuel, K. V.; Hunn, D. *J. Chem. Educ.* **1987**, *64* (8), 695–697.
- (10) Harrison, A. G.; Treagust, D. F. *Sci. Educ.* **1996**, *80* (5), 509–534.
- (11) Nakhleh, M. B.; Samarapungavan, A. *J. Res. Sci. Teach.* **1999**, *36* (7), 777–805.
- (12) Sanger, M. J. *J. Chem. Educ.* **2000**, *77* (6), 762–766.
- (13) Stains, M.; Talanquer, V. *J. Chem. Educ.* **2007**, *84* (5), 880–883.
- (14) Stains, M.; Talanquer, V. *Int. J. Sci. Educ.* **2007**, *29* (5), 643–661.
- (15) Talanquer, V. *Int. J. Sci. Educ.* **2009**, *31* (15), 2123–2136.
- (16) Duncan, R. G.; Hmelo-Silver, C. E. *J. Res. Sci. Teach.* **2009**, *46* (6), 606.
- (17) Kennedy, C. A.; Wilson, M. *Using Progress Variables To Interpret Student Achievement and Progress*, BEAR Technical Report: University of California, Berkeley: Berkeley, CA, 2006.
- (18) Briggs, D. C.; Alonzo, A.; Schwab, C.; Wilson, M. *Educ. Assess.* **2006**, *11* (1), 33–46.
- (19) Alonzo, A. C.; Steedle, J. T. *Sci. Educ.* **2009**, *93* (3), 389–421.
- (20) Steedle, J. T.; Shavelson, R. J. *J. Res. Sci. Teach.* **2009**, *46* (6), 699–715.
- (21) Vosniadou, S. Mental Models in Conceptual Development. In *Model-Based Reasoning: Science, Technology, Values*; Magnani, L., Nersessian, N., Eds.; Kluwer Academic Press: New York, 2002.
- (22) Gilbert, J.; Boulter, C. J. *Developing Models in Science Education*; Kluwer Academic Publishers: Dordrecht, Boston, 2000; p xi, 387.
- (23) Stains, M.; Sevian, H. The Structure And Motion of Matter (SAMM) Survey. <https://sites.google.com/site/sammsurvey/> (accessed Jul 2011).
- (24) Banerjee, M. *Can. J. Stat.* **1999**, *27* (1), 3–23.
- (25) Landis, J. R.; Koch, G. G. *Biometrics* **1977**, *33*, 159–174.
- (26) Bunce, D. M. *J. Chem. Educ.* **2008**, *85* (10), 1439.
- (27) Dalgety, J.; Coll, R. K.; Jones, A. *J. Res. Sci. Teach.* **2003**, *40* (7), 649–668.
- (28) Lawrence Hall of Science. *Chemical Interactions* (a module of the Full Option Science System for Middle School); University of California, Berkeley: Berkeley, CA, 2006.
- (29) Fink, A. *The Survey Kit*, 2nd ed.; Sage Publications: Thousand Oaks, CA, 2003; Vol. 8.
- (30) Ritchey, F. J. *The Statistical Imagination: Elementary Statistics for the Social Sciences*, 2nd ed.; McGraw Hill: Boston, MA, 2008.
- (31) Vosniadou, S. *Learning and Instruction* **1994**, *4*, 45–69.