Review and synthesis of research in Chemical Education from 2000-2010

A White Paper for

The National Academies National Research Council Board of Science Education
Committee on Status, Contributions, and Future Directions of Discipline Based Education Research

January 2011

REVISED DOCUMENT

Marcy Hamby Towns, Ph. D. and Adam Kraft
Department of Chemistry
Purdue University
West Lafayette, IN
mtowns@purdue.edu
Introduction

The initial charge of the scope of work was to summarize the empirical research on teaching and learning in chemistry. The DBER committee is likely aware that the research about teaching and learning in chemistry at the undergraduate level has not been thoroughly reviewed and characterized. Thus, a synthesis of the peer-reviewed literature has the potential advantage of identifying what the research unequivocally demonstrates, what it suggests, and where it is silent.

Methods

The focus of this review is on the peer-reviewed literature in chemistry education research. The NRC developed guidelines for the inclusion and exclusion of research that are outlined in Table 1.

Table 1. Guidelines for Including and Excluding Studies in the Review

<table>
<thead>
<tr>
<th>Include</th>
<th>Exclude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research or reviews/ syntheses that examine:</td>
<td>Research on:</td>
</tr>
<tr>
<td>• Instructional approaches</td>
<td>• Changing the practice of science education</td>
</tr>
<tr>
<td>• Student learning of science concepts</td>
<td>• Faculty professional development</td>
</tr>
<tr>
<td>• Student attitudes or beliefs toward science</td>
<td>• The development of chemistry education research as a profession</td>
</tr>
<tr>
<td>• The development and validation of measurement instruments</td>
<td>[NOTE: These topics will be addressed in subsequent meetings/papers.]</td>
</tr>
<tr>
<td>• Individual qualitative or quantitative research studies</td>
<td>• Editorials</td>
</tr>
<tr>
<td>• Comprehensive reviews of qualitative and quantitative research</td>
<td>• Papers that focus solely on defining a new theoretical framework</td>
</tr>
<tr>
<td></td>
<td>• Descriptions of an instrument without providing results of its</td>
</tr>
<tr>
<td></td>
<td>effectiveness</td>
</tr>
<tr>
<td></td>
<td>• Descriptions of innovative curricula or instructional techniques</td>
</tr>
<tr>
<td></td>
<td>without results of their effectiveness</td>
</tr>
<tr>
<td>• Research on undergraduate education, including at 2-year colleges</td>
<td>• Research on post-baccalaureate education</td>
</tr>
<tr>
<td>• Research on high school science learning where appropriate</td>
<td></td>
</tr>
</tbody>
</table>

It was determined at the outset that research outside of the US focusing on undergraduate education should be included. Research on high school populations of students where the impact or findings extended to the undergraduate population, such as David Treagust’s well-known misconceptions research, was also included.

The field of chemistry education research publishes predominantly in a small set of chemistry education and science education journals that are listed in Table 2. The 2009 two-year and five-year Impact factors for indexed journals are given where available [1].
Table 2: Chemistry education Journals and Science Education journals with impact factors.

<table>
<thead>
<tr>
<th>Journal title</th>
<th>2-year impact factor</th>
<th>5-year impact factor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chemistry Education Journals</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemistry Education Research and Practice</td>
<td>0.742</td>
<td>None*</td>
</tr>
<tr>
<td>Journal of Chemical Education</td>
<td>0.586</td>
<td>0.677</td>
</tr>
<tr>
<td>The Chemical Educator</td>
<td>Not indexed</td>
<td>Not Indexed</td>
</tr>
<tr>
<td>Australian Journal of Education in Chemistry (AUS)</td>
<td>Not Indexed</td>
<td>Not Indexed</td>
</tr>
<tr>
<td><strong>Science Education Journals</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Research in Science Teaching</td>
<td>1.910</td>
<td>2.805</td>
</tr>
<tr>
<td>Science Education</td>
<td>1.625</td>
<td>2.800</td>
</tr>
<tr>
<td>International Journal of Science Education</td>
<td>1.047</td>
<td>1.614</td>
</tr>
<tr>
<td>Journal of Science Education and Technology</td>
<td>Not indexed</td>
<td>Not indexed</td>
</tr>
</tbody>
</table>

Also included in the study were a few articles outside of this set of journals known to the researchers. For example, Mike Stieff’s work has appeared in the *Learning and Instruction* and is embedded in organic chemistry.

To analyze the studies we built a database composed of the following dimensions.

- Citation
- Research question
- Theoretical framework
- Study context
- Data collection and analysis methods
- Findings
- Limitations
- Year

Each article identified from a specific journal was read and compared to the inclusion criteria. Those that failed to meet the criteria were excluded from the study. Those that remained were incorporated into the database structure.

As the database was developed specific areas of research emphasis in the field began to emerge. The constant comparative method was used to code individual studies, thus the categories of research emphasis began to emerge and were entered into the database. The final categories of research were pedagogies, misconceptions, particulate nature of matter (PNOM), instrument development, student achievement, and miscellaneous.

**Results**

Initially research from the past 15-20 years was targeted. However, it became clear that a 10-11 year time frame provided literally hundreds of studies to read, analyze, and synthesize. The journal titles, years surveyed, and number of articles are listed in Table 3.

Table 3: Listing of journals and number of articles.
Each study was coded into categories based upon the focus of research. The design of the study was classified as qualitative, quantitative, or mixed. In the case where there was a treatment(s) and control group compared, that was also noted. Table 4 lists the number of studies by focus of research and research design used in the study.

Table 4: Number of studies in each category of research and the number of studies in each category by design.

<table>
<thead>
<tr>
<th>Focus of Research</th>
<th>Number of studies</th>
<th>Qualitative design studies</th>
<th>Quantitative (Treatment versus control)</th>
<th>Mixed (treatment versus control)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedagogy [2-162]</td>
<td>161</td>
<td>32</td>
<td>95 (34)</td>
<td>34 (6)</td>
</tr>
<tr>
<td>Misconceptions [163-232]</td>
<td>70</td>
<td>36</td>
<td>20 (3)</td>
<td>14 (0)</td>
</tr>
<tr>
<td>Particulate Nature of Matter (PNOM) [233-277]</td>
<td>45</td>
<td>20</td>
<td>16 (6)</td>
<td>9 (2)</td>
</tr>
<tr>
<td>Instrument development [278-290]</td>
<td>13</td>
<td>0</td>
<td>8 (0)</td>
<td>5 (0)</td>
</tr>
<tr>
<td>Student achievement [291-308]</td>
<td>18</td>
<td>1</td>
<td>17 (1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Miscellaneous [309-379]</td>
<td>72</td>
<td>24</td>
<td>40 (1)</td>
<td>8 (0)</td>
</tr>
</tbody>
</table>

Table 5 further describes each research category giving typical research questions, the percentage of high school (HS), general chemistry, organic chemistry, and upper division courses that were included as study contexts within the category. For example, in pedagogies 29% of the 157 studies (see Table 4) had a presence in high school classrooms. The typical duration of the studies is listed as an average in the table. However, it is very clear from the data that the most frequent length of time during which a study takes place is one semester. Across the entire data set “one semester” is the mode of typical duration. Typical data sources are listed for each category of research as well.
Table 5: Research questions, course level, duration, and data sources.

<table>
<thead>
<tr>
<th>Focus of Research</th>
<th>Typical Research questions</th>
<th>Typical courses and location (Majority in lecture or interviews)</th>
<th>Typical duration in semesters (Average given below. The mode is 1 semester)</th>
<th>Typical Data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedagogies</td>
<td>There are no “typical” research questions, however, frequently the questions center on the effectiveness or impact of a particular pedagogical implementation. Examples: To what extent is the PLGI reform an effective teaching practice in a high enrollment college chemistry course? To what extent is the reform an equitable teaching practice, with regard to equality of outputs, in the same setting?</td>
<td>29, 68, 16, 17&lt;br&gt;29 out of 157 studies focused on lab.</td>
<td>2</td>
<td>Test scores, end of course grade, attendance, concept test scores, field notes, surveys, attitude instruments, motivation instruments, laboratory reports, pre/post tests, demographics, spatial ability tests (FASP, PVROT [380], etc.), lab practicals, interviews, transcripts of classroom interactions, SATV, SATM</td>
</tr>
<tr>
<td>Misconceptions</td>
<td>There are no “typical” questions. The questions center on investigating conceptual understanding of students; Example: 1. Were propositional claims and/or causality included in students’ explanations? 2. What was the variation in students’ explanations according to the above [in (1)] two criteria? 3. What was the variation in students' notions of models? 4. What is the interpretation of combined</td>
<td>38, 70, 7.5, 17</td>
<td>1.8</td>
<td>Interviews and free response questions,</td>
</tr>
<tr>
<td>Instrument development</td>
<td>Particulate Nature of Matter (PNOM)</td>
<td>Research questions were guided by construct of interest.</td>
<td>Interviews and free response questions, 2-tiered diagnostic tests, student artifacts, pre/post tests, GALT</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Research questions focus on relating measures to student achievement in course or ACS exam. Example: Relation of performance on GALT to performance in first and second semester general chemistry</td>
<td>These questions focus on student understanding at the particulate level. Example: Does the use of computer animations at the particulate level accompanied by electronic potential maps (Elpots) improve conceptual understanding? Example: How do students at all levels complete Lewis structures? Example: What is the impact of dynamic and static visuals on college students understanding of concepts related to nucleophilic substitution and elimination reaction mechanisms in organic chemistry.</td>
<td>35, 65, 18, 17</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>18, 56, 50, 33</td>
<td>2.3</td>
<td>Interview and quantitative analyses, concept maps.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test scores; end of course grades, GALT, TOLT, GPA, Patterns of Adaptive Learning Survey (PALS), Motivated Strategies for Learning</td>
<td>22, 71, 80, 0</td>
<td>2.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summarized Findings

Pedagogies

This was the largest and most extensive body of research in the literature. These studies focused on investigating the effectiveness of an instructional strategy. Over 60% of the research designs were quantitative with 20% using a treatment versus control design. The research had its greatest presence in general chemistry classrooms where the students were college of science (or college of liberal arts and sciences) and college of engineering majors.

The findings from the research demonstrate that socially mediated forms of learning in a wide variety of frameworks such as PLTL, Peer Led Guided Inquiry, SCALE-UP and recitation like settings that employ small group learning produce positive outcomes. These findings include significantly higher test scores, higher final grades, better conceptual understanding, lower course withdraw rates, and positive impacts on attitudes.

One of the most robust studies was carried out by Tien, Roth, Kampmeier [87] on the impact of using the Peer Led Team Learning (PLTL) model in first semester organic chemistry. The study employed a mixed methods design using SAT, sex, ethnicity, total points, and course grade as variables. The treatment and control groups were students enrolled in the course from 1992-1994 that received the traditional lecture driven model, and the 1996-1998 group that received the PLTL model. SAT was used as a co-variate because the treatment and control groups were not statistically similar.

The findings demonstrated that students in the treatment group earned significantly more total points and earned higher course grades. In addition the qualitative outcomes of the study informed the quantitative outcomes. Findings from the qualitative portion indicated the connection to the improved student performance through having the students be a part of a community of learners, supporting and promoting reflection and explanation as the students worked problems and negotiating meaning in their groups, and helping the students to develop concept specific problem solving strategies and general thinking strategies.

Lewis and Lewis carried out another robust three-year study using a hybrid Peer Led Guided Inquiry (PLGI) [53] model where one lecture per week was replaced with a peer led problem solving session in first semester general chemistry. The study
investigated the effectiveness of PLGI and also asked if it was an equitable practice. The data included ACS final exam score, SAT math and verbal scores, exam scores, final grades, sex, and ethnicity.

To analyze the data Hierarchical Linear Modeling (HLM) was used. The results demonstrated that PLGI was associated with improved exam performance and it helped all students. There were no statistically significant differences in performance between sex or ethnic groups, and there were no interactions between PLGI and SAT scores.

The most compelling data that engages a predominantly under-represented minority student population comes from a study by Akinyele at Howard University [3]. This university is an HBCU with approximately 7,000 undergraduates, 83% of whom are African American (non-Hispanic). The course was a general-organic-biochemistry course, where the general chemistry was taught in the fall semester, and the OB content was covered in the spring. Students registered for chemistry did not know which sections would be taught in a PLTL format or a traditional format.

The results in this study were similar to those from the Tien et al. and Lewis and Lewis studies [53, 87]. The mean overall percent score in the course was significantly higher for the PLTL group every semester from Fall 2004 to spring 2007. The students did not have significantly different SAT scores in any semester, nor did they have significantly different attendance scores between the PLTL group and the traditional group. The only negatively outcome that Akinyele cites was the high withdraw rate during the first semester of implementation in Fall 2004 [3]. After that semester the withdraw rate dropped to nearly half that of the traditional course.

Beyond the positive outcomes for students there is another attribute of these studies that is unequivocal. In every case where a form of small group learning was implemented that required peer leaders (PL), the PLs tasked to lead the groups were recruited from a pool of high achieving students and they were trained on a weekly basis. These are critical attributes of successful implementations.

**Laboratory**

There are 29 studies that focused on the laboratory environment [21, 25, 28, 33, 42, 43, 45, 49, 62, 65, 73, 75-77, 90, 93, 95, 97, 98, 104, 127-129, 131, 135, 142, 143, 148, 354]. The findings from the research are suggestive of the following approaches.

1. Implementing a cooperative PBL format in organic chemistry had a very positive effect on student attitudes and perceptions. The researchers collected observations during laboratory as well as videotaping and further analyzing student interactions. Two types of interactions were discovered, those in the student’s main laboratory group where learning and overall synthesis of results occurred, and those within a secondary group where students cross checked their results and reinforcement of data analysis occurred. Based upon student surveys, they found this pedagogical
approach to be more like research rather than a “cookbook” recipe following laboratory [25]. A PBL format in general chemistry also produced a similar preference among students for the PBL approach as opposed to the traditional laboratory [148].

2. There are studies which suggest possible approaches to supporting inquiry or developing inquiry skills in the laboratory [42, 43]. Hofstein and coworkers compared the number and level of questions posed by students who were instructed to critically read a scientific article [43]. The students had either been in the treatment group and exposed to a laboratory curriculum that included inquiry experiments, or the control group that followed a traditional curriculum. Students in the inquiry group who had experience in developing questions out performed the control group by asking more and better questions for future laboratory investigations. In another study lead by Hofstein the goal was to develop, implement, and assess the learning outcomes of inquiry-based laboratory experiments in the context of high school chemistry in Israel [42]. The study demonstrated that by conducting the experiments students were able to practice inquiry skills such as asking questions, hypothesizing, and suggesting a question for further investigation using an experiment that they planned. The analysis of students’ laboratory reports clearly demonstrated that they improved their abilities regarding inquiry learning in the chemistry laboratory. Finally, in a study to investigate the cognitive and motivational effects of explicit pre-experimental activities in planning investigations [65]. The experimental group was able to craft a significantly higher number of questions that focused on causes than the students in the control group. The authors note that the pre-experimental activities were elements of an inquiry cycle that had been broken down into sub-processes. They speculate that automatizing such sub processes requires repeated exposure to the entire routine.

The body of research doesn't point towards one incontrovertible approach to laboratory. Further there is only one study that explores the goals faculty have for laboratory [16]. The results of this study demonstrated that the goals differ by course within the chemistry curriculum. The goals that span the curriculum are 1) to master laboratory skills and techniques, and 2) to develop critical thinking skills and experimental design skills. Across courses faculty gave voice to varied and subtly different meanings to mastering laboratory skills and techniques. Further, critical thinking skills discussed by general chemistry and organic chemistry faculty are related to the development of experimental designs discussed by faculty in the upper division courses.

**Misconceptions/Conceptual Understanding**

The research on misconceptions between 2000-2010 continues the tradition of describing student conceptual understanding of a variety of chemical phenomena including atomic structure, elements, compounds, and mixtures, solution chemistry intermolecular forces, equilibria, kinetics, and thermodynamics. The investigations are
most commonly set in the general chemistry classroom and fewer studies consider students in organic or upper division courses. The importance of the research findings in organic and upper division courses is to point out that misconceptions are highly resistant to change. All one has to do is to read Sozbili’s work in thermodynamics to be convinced that misconceptions students hold in high school remain in upper division courses [167, 189, 218, 379].

Henderleiter, Smart, Anderson, and Elain [176] carried out a qualitative study on how students understand, explain, and apply hydrogen bonding to physically characterize molecules. The 22 volunteer participants were completing their second semester of organic chemistry. The interview protocol given in the article (p. 1127) focuses on defining hydrogen bonding, describing when it occurs amongst molecules, the physical properties of molecules, and spectroscopy.

The findings strongly suggest that after nearly four semesters of chemistry students hold on to misconceptions. The faulty conceptual information and tendency to memorize disconnected bits of information make it very difficult for students to use concepts meaningfully to analyze problems placed before them.

For example, 18 out of 22 students could correctly describe the relationship between electronegativity of the atom bonded to H and the polarity of the bond. However, 13 students generated many possible atoms that could hydrogen bond (Cl, S, P, and C) in addition to F, N, and O. All students stated that hydrogen bonding could occur between molecules of one type (water), or two types such as water and propanol. However, five students believed hydrogen bonding could be induced. Another question probed the students’ understanding of intramolecular hydrogen bonding. Five students believed that intramolecular hydrogen bonding generated new covalent bonds.

When probed about physical properties 16 students could accurately describe trends in boiling points for a set of alcohols, but five could not. When given 1-butanol, 1-butanal, and propanoic acid 11 students correctly identified the compound with the highest boiling point and used appropriate reasoning. Seven other students relied on rote memorization with only one recalling the trend in boiling point correctly. Three students used resonance arguments and one used pKa.

What is troubling in these responses is that after four semesters of chemistry students cannot easily identify concepts that are relevant or irrelevant to the problems presented to them. Further, a reliance on rote memorization with little meaningful reasoning connected to the memorized bits is of little use in analyzing and interpreting data.

Beyond describing student misconceptions there is the important question of how they can be effectively addressed. What does the research say about pedagogical approaches to conceptual change? In eleven years there have been few studies to address this issue. However, there are theoretical models that point towards how conceptual change can be fostered in the classroom. The model of conceptual change developed by Posner et al [381] proposed four conditions that must occur prior to a
student modifying a misconception: the student must be dissatisfied with the current conceptual model, the new conceptual model must be comprehensible and plausible, and the new conceptual model must be better at explaining the observation than the previous conceptual model.

Zimrot and Askenazi [231] formulated an interactive pedagogical approach to conceptual change that focused on demonstrations in a first semester general chemistry course. They developed and implemented interactive lecture demonstrations where students chose a pre-determined outcome before the demonstration (the students had specific worksheets), discussed their ideas with a peer, voted for an outcome using clickers, then the instructor sought engaged feedback from the class for the outcomes. The instructor then performed the demonstration, noted the correct outcome from the predictions, then asked students choose an explanation from four choices. The students discussed the explanations and indicated their choice using a clicker. The instructor then led a discussion of each explanation noting the inconsistencies of those that were incorrect referencing the alternative conceptions associated with the choice and the explanatory power of the correct choice.

To determine the impact of interactive lecture demonstrations the researchers interviewed 12 students 3 times each during the semester. The interviews asked students to recall a demonstration, their initial prediction, and how they would explain the demonstration now. If the student’s explanation changed they were probed as to why.

A quantitative component was also carried out in a treatment versus control design across two fall semesters. In the first fall semester after week 8 all the interactive steps of the demonstrations were removed. The fall semester a year later retained the interactive steps for the entire semester. The two groups of students had similar GPA values in biology, physics, and math, which were not significantly different. At the end of both fall semesters a conceptual test was administered.

The results of the interviews demonstrated that the student’s ability to recall and explain the demonstrations is characterized by how they changed their prior models. Students who were characterized as low level conceptual change could not recall the demonstrations or the explanations. Essentially they did not undergo conceptual change because they failed to recognize any conflict. Medium level conceptual change students could recall the demonstration and the conflict in their predictions and answers, but they could not generate a scientifically correct explanation. High-level conceptual change students recalled the demonstration and can give a scientifically correct explanation.

The quantitative results were also revealing. The conceptual test was split into two parts to control for the first 8 weeks when all students received the interactive treatment (this score was used as a covariate). Thus the score on part two depended upon the second 8 weeks where the students were associated with one of two conditions. An ANCOVA showed a statistically significant difference in the part two conceptual score
(\(p < 0.001\)) with a medium positive effect size of \(d = 0.48\) favoring the more interactive condition. Thus, the research suggests that it is possible to carefully craft undergraduate learning environments that promote conceptual change.

What is needed is more research that documents with robust data how misconceptions can be addressed and change through effective pedagogies.

**Particulate Nature of Matter (PNOM)**

The particulate nature of matter is indigenous to chemistry. There is an expectation among chemistry faculty that students will develop fluency between the symbolic, macroscopic, and particulate domains [382, 383]

![Diagram showing symbolic, macroscopic, and particulate domains]

However, an abundance of research has pointed out student difficulties with the particulate domain. In the years 2000-2010 this was a vigorous area of research.

The Lewis structure pedagogy is a method to help students develop an understanding of bonding in molecules and eventually through VSEPR an understanding of molecular shape. Faculty in general chemistry would emphasize how molecular shape and electron distribution would influence physical properties, polarity, solubility, and miscibility. Faculty in organic chemistry and biochemistry would emphasize the structure-function relationships. Thus, an understanding of matter at the particulate level is critical to understanding the behavior and interactions of molecules.

Cooper, Grove, Underwood, and Klymkowsky [244] investigated how students at all undergraduate levels draw Lewis structures. The study used a mixed method design. Twenty-one students from freshman to senior level (6 graduate students as well) were interviewed and asked to draw Lewis structures for \(\text{NH}_2^+\), \(\text{NO}\), \(\text{CH}_4\text{S}\), \(\text{C}_2\text{H}_6\), and \(\text{C}_3\text{H}_7\text{NO}\). During the interview they were asked what essential features Lewis structures contained what information could be gleaned from them, and how they could be interpreted. The quantitative portion asked 166 undergraduates in general (\(n = 32\)) and organic chemistry (\(n=134\)) to submit Lewis structure drawings and to discuss the functionality of Lewis structures via an online data collection system.
The students demonstrated a lack of clarity about drawing Lewis structures and during the interviews varying degrees of frustration. When the number of atoms increases from six to seven there is a precipitous decline in the accuracy of the drawings—a decrease from 80% to 30% correct. This is a reflection of moving to a multicenter atom where the connectivity of the atoms is not evident to the student unless he or she recognizes the molecule.

The students focus their attention on drawing Lewis structures to the exclusion of practicality. The students in the study had great difficulty in predicting structure function relationships. Only 30-40% of the students indicated that Lewis structures could be used to determine the shape of the molecule. And perhaps worst of all only 56% of the general chemistry students and 31% of the organic chemistry students stated that chemical information could be obtained from Lewis structures.

Nicoll [252] interviewed 56 students across general, organic, inorganic, and physical chemistry asking them to draw the Lewis structure for CH$_2$O and to make a model of the molecule using playdoh and sticks. Her field notes from the study included a coding scheme, arrangement of the atoms, size of the playdoh balls, geometry, color, and placement of the sticks that represented bonds. Images of the student models and the coding scheme were included in the article. The correct Lewis structure is shown below.

![Lewis structure of CH$_2$O](image)

Figure 1: The Lewis structure of CH$_2$O.

The research demonstrated a very mixed understanding of this simple molecule. Although 70% of the students wrote and built a trigonal planar molecule (which is correct), 39% placed the oxygen atom in the center rather than the carbon atom. Further, 17% of the students made the carbon atoms and the oxygen atoms the same size.

Nicoll found that bonding among atoms is not well understood. Note that the 39% of the students who placed oxygen in the center might have believed that oxygen could form four bonds. There were also students who formed a more hydrogen peroxide structure with the atoms, thus indicating that carbon only formed two bonds.

What both the Cooper and Nicoll studies unequivocally demonstrate is that students do not improve their understanding of molecular models and associated concepts as they progress through the curriculum [244, 252]. In both studies there was a lack of understanding of how atoms are placed in a molecule. Essentially Cooper’s study demonstrated that Lewis structures for multi-centered molecules could be accurately drawn only after the student knows the correct structure. Further students lack an
understanding of the proper concepts to interpret the structures. Finally, there is very little understanding of the physical and chemical properties that may be discerned from these drawings.

However, there are studies that point to pedagogical techniques that target specific concepts at the molecular level. Sanger and Badger [257] conducted a study to discern if particulate animations accompanied by electrostatic potential maps used with the traditional methods improved student conceptual understanding. The study used a treatment versus control design. The 36 students in the control group received instruction on molecular polarity and miscibilities using static drawings, models, and demonstrations. The 36 students in the treatment group received molecular animations and electrostatic potential maps in addition to the traditional instructions. The research included the 1997 ACS special exam (second term) scores that demonstrated no significant differences between the treatment and control group.

The findings demonstrate that students who received the treatment condition were significantly better at identifying polarity of molecules, miscibilities of liquids, and intermolecular interactions between water and salt, and soap, grease, and water.

Abraham and co-workers have investigated student understanding of organic chemistry concepts under conditions of traditional static visuals and animations. Aldahmas and Abraham [237] investigated the difference between animations and static visuals on student understanding of nucleophilic substitution and elimination reaction mechanisms. 142 students volunteered for the study. 71 were randomly assigned to the animation condition, 71 were randomly assigned to the static condition, and the remaining 101 students in the course who didn’t volunteer formed a reference group. The treatment group viewed animations of nucleophilic reaction mechanisms that included 3D representations of molecules. The control group viewed static visuals on a computer screen that contained 2D representations of molecules and equations representing nucleophilic reaction mechanisms. The reference group attended lecture and were assigned readings that covered identical material. The treatment and control group completed two spatial ability tests, the PVROT [380] and the Find-a-shape-puzzle, the FASP test.

All three groups completed an exam unrelated to the reaction mechanism content. The exam scores were not significantly different indicating that the groups had an equal ability to learn organic chemistry concepts. On a content test related to nucleophilic reaction mechanisms there were significant differences in performance. The students viewing the animations (treatment group) scored 10% higher than the static (control group), and 22% higher than the reference group. On open-ended questions graded by the researchers and experienced teaching assistants (IRR = 99.7%) the animation group scored 16% higher than the static group, and 30% higher than the reference group all of which were statistically significant. The static group scored significantly higher on the content test and open-ended questions than the reference group.
To determine the effect of spatial ability on these results a two-way ANOVA using group designation as one independent variable (animation or static) and spatial ability as the independent variable. Students were classified as high, medium, and low based upon PVROT and FASP scores. The dependent variable was the score on the content test related to nucleophilic reaction mechanisms. There were statistically significant main effects and interaction effects. Perhaps not surprisingly the high spatial ability animation group attained a higher test score than any other group (Fisher PLST post-hoc analysis used, P = 0.0015).

The authors note in their conclusions that animations linked to a clear explanations likely account for the results of this research. They also note that there is a body of research on computer animations used in chemistry that should encourage faculty to use 3D animations in their instruction.

Abraham, Varghese, and Tang [233] carried out an investigation of the effectiveness of three kinds of molecular representations on student understanding of stereochemistry concepts in a first semester organic chemistry course. The conditions in the study were a computer visualization group, a hand-held ball and stick model group, and 2D perspective drawing group, and a reference group that did not volunteer for the study. The three treatment groups were randomly assigned and spent nearly equal amounts of time (60-70 minutes) exploring the treatment activities that covered the same content and used the same examples. The reference group did not engage in these activities and thus serve as a measure of the importance of “time on task.”

The students engaged in four types of assessments. The treatment groups completed the PVROT and were given a quiz immediately after the treatment. All students completed a regularly schedule content test before the treatment unrelated to stereochemistry. All students also completed a regularly scheduled exam written by the course instructor two weeks after the treatment that contained 13 multiple choice questions and 3 open-ended questions. Two of the multiple-choice questions and one open-ended question were chosen by researchers for analysis because their relationship to the stereochemistry treatment content.

Based upon a lack of statistical significance between the groups on the unrelated content test, the PVROT scores, and the unrelated test items on the test given after the treatment, the treatment groups were considered to be equivalent in their spatial ability and all groups were considered to be equivalent in their ability to learn organic chemistry content. No statistically significant differences were found on the quiz given immediately after treatment.

There were statistically significant differences found on the test items that pertained to stereochemistry (ANOVA, F = 13.418, P< 0.0001). The differences between groups were analyzed using a Newman-Keuls test that showed statistically significant differences in the following: the computer animation group scored 15% higher than the other treatment groups, and 37% higher than the reference group. The hand-held model group and the 2D perspective drawing group scored 22% higher than the
reference group. The authors note that they were not able to find statistical differences using a two-way ANOVA that indicated spatial ability or gender played a role in performance on the stereochemistry questions. This is in contrast to previous work by Bodner that showed spatial ability was a factor in student performance on organic chemistry questions requiring the mental manipulation of 2D representations of molecules [384]. The results of this study show that computer animations can be effectively used to help students learn stereochemistry concepts.

Synthesizing the results of work by Abraham and Sanger [233, 237, 257] the research results strongly suggest the value of using computer animations to help students learn particulate level concepts. In both studies the authors encourage practitioners to use the best practices possible in order to help students learn these particulate concepts. In Sanger’s work and Abraham’s 2009 study both researchers note the importance of clear explanations coupled to the animations. In Abraham’s 2010 study he notes the importance of linking 2D and 3D representations.

The question of transfer of knowledge from animations can be addressed by a set of studies from Kelly and Jones [249, 251]. These studies focused on students’ ability to connect macroscopic phenomena to microscopic representations, then to transfer that knowledge. In the first experiment 18 students watched a molecular level animation of NaCl dissolving into water. They were asked to draw at the macroscopic and microscopic level what took place. After viewing molecular level animations of NaCl dissolving all 18 students drew water interacting with sodium and chloride ions and no students had retained drawings with ion pairs.

The second study was conducted one week later and it focused on discovering what features of their drawings transferred from the first experiment to the second. The second interview began with the same participants watching a video of NaCl (aq) being mixed with AgNO₃ (aq). The students were then asked to draw what they saw and to explain it, and to draw what they would see at the molecular level and explain it.

In the second study, only two out of 18 students drew NaCl (aq) with the ions separated after watching the video demonstration of the solutions before and after mixing. During the interview 15 out of the 18 students were reminded of the animation they watched the week before, dissolving NaCl in water. After prompting, nine out of 15 students changed their NaCl (aq) drawings to represent separated ions. Some, but not all, of these students created drawings where the ions interacted with water.

Essentially the set of experiments suggests that it is difficult for students to transfer their mental models and conceptual understanding of NaCl dissolving and of NaCl (aq) without prompting. The experimental protocol was changed because the researcher noticed the first three students didn’t make a connection to the previous animation. Thus, the research suggests that students need explicit instruction that connects previous microscopic representations to the “new” content that the faculty may wish to have students depict and understand at the microscopic level.
Teichert, Tien, Anthony, and Rickey [264] conducted a investigation on the effect of context on student’s particulate level understandings. Particularly interesting was the use of a NaCl (aq) solution that allows for comparison to Kelly and Jones work.

In the Teichert et al. study general chemistry students (n=11 from a community college and n = 8 from a R1 institution) had completed a laboratory module that emphasized developing models of an aqueous solution at the molecular level. As part of the laboratory the students measured the conductivity of a NaCl (aq) solution and completed laboratory activities that encourage them to explain at the molecular level in words and pictures what was happening in the solution.

After the laboratory module, 19 students volunteered to be participant in a two-part interview study. Part one asked the student to make predictions about the conductivity of NaCl (aq) and AgNO₃ (aq) solutions and to draw each solution at the particulate level. Students were next asked to predict the product(s) of mixing these two solutions, predict the conductivity, and to make a particulate level drawing of the mixed solution. The researcher measured the conductivity of each solution and the students were allowed to make changes to their diagrams and explanations.

Part two of the interview introduced the students to one colligative property, BP elevation, via an explanatory paragraph. The students had not been exposed to this content in their general chemistry course. The salient feature of colligative properties emphasized to the students was that properties depend on the total number of solute particles in the solution, not their identity. Students were asked to explain their understanding of the paragraph then predict the relative BP elevation of pairs of aqueous solutions including 0.02 M and 0.06 M C₆H₁₂O₆ solutions, 0.02 M NaCl and 0.02 C₆H₁₂O₆ solutions, and 0.12 M CaCl₂ and 0.15 M KCl solutions.

In part I of the interview eighteen out of 19 students correctly described a NaCl (aq) solution as consisting of separated ions and predicted that the solution would conduct electricity. Seventeen students were able to create appropriate drawings with separated ions, but in this investigation the interaction with water was not discussed as it was in the Kelly and Jones studies [249, 251]. Fifteen out of 19 students drew correct representations of an AgNO₃ (aq) solution by separating the silver and nitrate ions. However, when asked to represent the mixed NaCl (aq) and AgNO₃ (aq) solutions, about half of the students (53%) drew AgCl as molecules in their pictures rather than a distinct ionic solid with an array of positive and negative ions.

However, when the context changed to the BP elevation questions, only ten out of 19 students correctly drew sodium and chloride ions separated the NaCl (aq) solution. Nine of the students reverted to drawing NaCl as ion pairs. These students were prompted to compare their drawing to their previously drawn NaCl (aq) solution from the conductivity portion of the interview. This comparison caused six students to change their drawings to separated ions in solution.
What is striking in this experiment is nearly half of the students failed to draw an appropriate molecular level view of NaCl (aq) in the second part of the interview even though 15 minutes earlier 95% of the students drew the solution as separated ions. Teichert et al suggest that inconsistency in the use of particulate models may be due the activation or lack of activation of cognitive resources in a given context. They go on to suggest that identifying where these context dependent models of particulate level behavior are found to occur will be an important area of research.

Given that the Teichert et al study [264] and the Kelly and Jones [249, 251] work share the features of having students represent NaCl (aq) solutions at the particulate level, and that they were asked to transfer that knowledge either 15 minutes or one week later it is interesting to synthesize their findings.

The transfer of knowledge, when it occurs and how it occurs has not been widely studied in the field of chemistry education research. These three studies strongly suggest that the transfer of particulate level models does not easily occur at least in the case of general chemistry students who could be considered novices or at least much less experienced in the field. The actions by the researchers during the interview protocols suggest that with prompting—either by suggesting that the students recall a previously viewed animation or by asking the students to compare their representations of NaCl (aq), that transfer between the two tasks was facilitated for some, but not all, students.

**Instrument Development**

In the past 11 years reliable and valid instruments that measure student understanding of concepts, cognitive expectations of learning chemistry, beliefs about chemistry and learning chemistry, attitudes towards chemistry, and course perceptions, have been developed for use in chemistry and chemistry education research. Beyond new instrument development tests of formal reasoning ability have been compared to determine if there is an advantage of one test over another. The instruments are divided into broad areas as follows: conceptual understanding; cognitive expectations, metacognition, and formal reasoning ability; beliefs, attitudes, and perceptions; and other instruments

It is critically important that those carrying out research and assessments use reliable and valid instruments. Instruments such those described below allow researchers to compare data and results across studies.

*Conceptual Understanding*

**Chemistry Concepts Inventory** [285]

Mulford and Robinson developed a 21-item instrument that measures a student’s conceptual understanding in chemistry at the general chemistry level and has a
Cronbach $\alpha$ value of 0.704 (pre) and 0.716 (post). The development of the instrument was strengthened by a mixed methods design. Content from the first semester of general chemistry for science and engineering majors was matched to misconceptions in peer-reviewed literature. A free response instrument was piloted in order to develop responses for a multiple-choice format. Based upon the pilot data 22 questions were developed for the final instrument. In general chemistry there was a pre/post test design to calculate Cronbach $\alpha$, and eight interviews verified the validity of the inventory.

Two-tiered Diagnostic: PNOM and bonding [286]

This is a 10 item diagnostic instrument to assess student understanding of PNOM and chemical bonding. The instrument is available from first author Othman and the instrument has a Cronbach $\alpha$ value of 0.66. The authors caution that this instrument has only been implemented and analyzed in one secondary school in Singapore, thus the findings are considered preliminary.

Two-tiered Diagnostic: Inorganic qualitative analysis [287]

This is a 10 item diagnostic instrument to assess student understanding of inorganic qualitative analysis. The instrument is available at the end of the article and has a Cronbach $\alpha$ value of 0.68. This instrument was also developed in a secondary school Singapore.

Two-tiered Diagnostic: Representational systems and chemical reactions diagnostic instrument [288]

This is a 15 item two-tiered diagnostic instrument to assess student understanding of chemical reactions using multiple levels of representation (symbolic, macroscopic, and microscopic). The instrument has a Cronbach $\alpha$ value of 0.65. It was developed in a single secondary school with students outside of the USA.

Cognitive expectations, metacognition, and formal reasoning ability

CHEMX [284]

Grove and Bretz developed a 47-item instrument with 7 clusters (effort, concepts, math link, reality link, outcome, lab, visualization) to describe students’ cognitive expectations for learning chemistry. The Cronbach $\alpha$ for the clusters ranges from 0.73 to 0.89 and for the entire instrument is 0.97. Analysis of student performance with respect to faculty indicates that during general chemistry the students move farther away from faculty expectations, then during the sophomore year move towards faculty and surpass the freshman level values in the junior year.

Metacognitive Activities Inventory [281]
Cooper and Sandi-Urena developed an instrument to measure a student’s metacognitive skillfulness in chemistry problem solving. The instrument was developed using a panel of experts model. Reliability and validity are addressed. The Cronbach \( \alpha \) values are 0.85 (pretest) and 0.91 (post-test) in the main study. The factor analysis was not clear as metacognitive factors are interdependent.

**GALT, TOLT, and TOLT + 2 [299]**

The GALT, TOLT, and TOLT+2 were compared in first semester general chemistry classes. No advantage was found to adding the two concrete items to the TOLT based upon reliability and discriminatory power. The GALT and TOLT were compared using first semester general chemistry students and preparatory chemistry students. Again, the reliability, discriminatory power, and potential item bias do not indicate that one instrument should be preferred over the other. However, it was noted that more items on the GALT instrument exhibited differential item functioning (DIF) than on the TOLT. Further, item 2 on the GALT demonstrates a potential bias against females in the general chemistry and preparatory chemistry populations. It also has one heteronormative item (#11) and demonstrates an absence of cultural sensitivity.

**Beliefs, attitudes, and perceptions**

**Attitudes Toward the Subject of Chemistry Inventory (ASCI) [279]**

Bauer developed an attitude towards chemistry instrument that is composed of a 7-point semantic differential scale anchored by opposing polar adjectives. Exploratory factor analysis demonstrated three factors: interest and utility, anxiety and intellectual accessibility with Cronbach’s \( \alpha \) values of 0.83, 0.77, and 0.78 respectively. The instrument proves a measure of the students’ emotional stance towards chemistry.

**C-LASS (Colorado Learning Attitudes about Science) [278]**

Barbera, Adams, Wieman, and Perkins modified an existing physics instrument to measure student beliefs about chemistry and the learning of chemistry. They conducted interviews to improve the validity of in the instrument that led to the removal of some items. The instrument has a Cronbach \( \alpha \) value of 0.89. They cite robustness values for nine categories (factors perhaps), but give no factor analysis data for each category or the Cronbach \( \alpha \) values.

**Chemistry Attitudes and Experiences Questionnaire (CAEQ) [282]**

Dalgety, Coll, and Jones developed an instrument to measure attitudes towards chemistry, chemistry self-efficacy, and learning experiences for first year chemistry students across a wide range of majors. The instrument has a Cronbach \( \alpha \) value of 0.76. Detailed factor analysis, KMO values, and discussion of validity are presented. The instrument is available from the authors in electronic form. The instrument was also described in Coll, Dalgety, and Salter [290].
Attitude Toward Chemistry Scale [280]

This is a modified form of Fraser’s test [385] of Science Related Attitudes based on four dimensions: enjoyment of chemistry lessons, enjoyment of chemistry laboratory, evaluative beliefs about learning chemistry in school, and behavioral tendencies for learning chemistry. The instrument is composed of 12 items to be rated on a seven-point Likert scale from strongly disagree to strongly agree. Factor analysis was carried out and four factors were found. The Cronbach α values for each scale were 0.86 (chemistry lessons), 0.84 (chemistry laboratory), 0.76 (beliefs about learning chemistry), and 0.76 (behavioral tendencies).

Other

Chemistry Course Perceptions (CCP) [283]

Reardon, Traverse, Feakes, Gibbs and Rohde developed an instrument via interviews to assess student’s chemistry course perceptions. They interviewed 10 students about their feelings towards chemistry and the course they were enrolled in. From the transcripts the survey was created. They found a Cronbach α value of 0.82 the instrument. Chemistry course perceptions are not a strong function of sex or ethnicity. The university at which the research was carried out has a goal to be come a Hispanic Serving Institution by 2012.

Rubric to characterize laboratory inquiry

Fay, Grove, Towns, and Bretz (2007, CERP), have developed a rubric chemistry faculty can use to examine the level of inquiry facilitated by the experiments used in their laboratory curriculum.

Student Achievement

There is a small group of studies that focused primarily on predicting student performance in specific courses. These studies were nearly all quantitative in design. They used combinations of GPA, grades in math and chemistry courses, test scores, final scores/grades, ACS exam scores, tests of logical thinking (TOLT and GALT), and measures of attitudes, motivation, or self-concept coupled with statistical analyses to predict performance or account for variance in grades in a course.

In terms of performance in general chemistry, the following studies present findings of note. Lewis, Shaw, Heitz, and Webster administered Bauer’s self concept inventory to 630 students [300]. The instrument has five subscales: chemistry, mathematics, and academic self-concept, academic enjoyment self-concept, and creativity self-concept. Using cluster analysis the students were collapsed into five clusters—high self-concept, high math and chem, average, low math, low self-concept. Students took the ACS first
term general chemistry exam at which point 411 students remained for analysis. The performance on the ACS exam was statistically significant between groups (ANOVA, \(F(4,344) = 7.678, p < 0.001\)), with the high self-concept and high math and chem groups significantly outperformed the remaining three groups according to a Tukey test. Thus, there is a clear relationship between higher self-concept and higher ACS exam scores.

The GALT and TOLT are frequently used measures of logical thinking and formal reasoning. Jiang, Xu, Garcia, and Lewis [299] analyzed the TOLT, TOLT + 2, and GALT as described in the instrument development section. They recommend using the TOLT as a more balanced and culturally sensitive instrument.

Bird [291] has found evidence that students who are formal thinkers as opposed to transitional or concrete thinkers as classified by their score on the GALT score significantly higher on the ACS general chemistry exam (F=17.99, p < 0.01). However it is unclear which of the means were different. The students in this study were more than 99% Hispanic, almost all Puerto Rican. The author also demonstrated that in both semesters of general chemistry final course grades differed significantly by operational level (formal, transitional, or concrete) \(X^2 = (8,N=466) = 52.89, p < 0.001\) and \(X^2 = (8,N=466) = 52.48, p < 0.001\).

The role of motivation and cognition in learning general chemistry content was explored by Zusho, Pintrich, and Coppola [305]. 458 students in two introductory chemistry courses completed surveys at weeks 5, 10, and 15. The first survey included demographic information, self-efficacy, and task value beliefs. The second and third assessed goal orientation, self-efficacy, task value, interest, anxiety, and use of cognitive and self-regulatory strategies. Most items were adapted from the Patterns of Adaptive Learning Survey and the Motivated Strategies for Learning Questionnaire. Final grades were also included for analysis.

The student's motivation as measured by self-efficacy (F(2,443) = 15.10, p < 0.001), task-value (F(2,443) = 91.40, p < 0.001), and performance goals (F(1,443) = 11.662, p < 0.001) all significantly decreased with time. Only mastery goals showed no significant differences. Measures of affect including interest and anxiety showed no significant differences with time.

Strategy use changed in unusual ways during the semester. Rehearsal strategies (F(1, 452) = 77.51, p < 0.001) and elaborative strategies (F(1, 451) = 180.77, p < 0.001) both decreased with time while organizational (F(1, 449) = 251.92, p < 0.001) and metacognitive strategies (F(1, 405) = 18.01, p < 0.001) increased.

To analyze the motivational and cognitive measures by performance the students were divided into three performance groups (high achieving scores above 81%, average achieving 70%-80%, and low achieving below 69%). They found that high-achieving students' self-efficacy scores increased significantly with time and low achieving
students’ self-efficacy scores decrease significantly \((F(4, 420) = 22.99, p < 0.001)\). Ratings of interest also varied with time by performance level in a similar fashion.

In a study to determine significant factors of performance in organic chemistry, Turner and Lindsay [303] found that performance in general chemistry was the most consistent and best predictor of performance in organic chemistry. This finding is based upon significant correlation values and stepwise multiple regression analyses.

Miscellaneous

These studies do not easily fit into the categories that emerged from the data. The types of research questions, contexts, methods, findings and limitations are broad—essentially they are not easily grouped by the nature of the research question, course, population, data collection techniques or analysis, etc. The inclusion of a study in the miscellaneous category does not indicate \textit{a priori} that the study was flawed in design, data collection, or analysis. It also does not indicate that the study was limited.

However, some of the studies focus on contexts that severely limit the impact of the findings. Unfortunately some studies are flawed and limited in ways that negatively impact the usefulness of the findings.

Examples of studies in the miscellaneous category by research question or research focus:

- How do students use their textbooks? [342]
- What are the effects of a remedial chemistry course (prep chem.) on the chemistry curriculum at a particular institution? [310]
- An analysis of the inclusion of people of color in 11 general chemistry texts. [323]
- What do undergraduate researchers do at ACS national meetings? Why do they attend? [325]
- Analysis of types of organic chemistry problems used on exams via three frameworks. [368]

What does the research state unequivocally, what does it suggest, and where is it silent?

At the conclusion of this review the charge was to communicate the findings in terms of what is known unequivocally based upon the data, what the data suggests, and where any data or evidence has yet to find a voice.

The data strongly supports the use of socially mediated learning (in whatever form the faculty wishes) as a method of improving outcomes. There is incontrovertible evidence that in programs which use peer leaders as facilitators that the peer leaders must be trained throughout the semester. Further the peer leaders must be drawn from an academically high achieving pool of students.
There is a wealth of data on student misconceptions. However, missing from the research base are investigations of pedagogical approaches that facilitate conceptual change and evidence that change has occurred. Further, it would be important to know how durable the change actually is—how long does it last?

The research unequivocally demonstrates that students do not develop strong particulate models of matter nor the concepts associated with them as they progress through the curriculum. Further there is indisputable evidence that transferring conceptual understanding of particulate behavior from one context to another is difficult for students. Transfer of knowledge has not been a vigorous area of research in chemistry education.

There is evidence in multiple studies that suggest the value of using particulate level animations to help students learn particulate level concepts. Researchers emphasize the importance of using appropriate pedagogical techniques when using such animations. The studies show mixed results pointing towards spatial ability (actually the rotational spatial ability factor as measured by PVROT) as a mediating factor in student performance.

A wide variety of reliable and valid instruments including ACS Exams [386] for the entire curriculum are available for faculty use. The issue is not a lack of instruments, rather its that faculty frequently do not use them but rather opt for individually generated instruments or surveys which perhaps are neither reliable nor valid. For research outcomes to be compared across studies it would be incredibly helpful to the field if those engaged in research used widely available reliable and valid instruments.

In terms of predicting student achievement there is not a large research base to draw from in CER. Be that as it may, the research does suggest that higher GALT scores are associated with higher scores on the ACS general chemistry exam. There also appears to be a positive relationship between self-concept and ACS general chemistry exam score. Focusing on motivation and cognition there is data to suggest that high achieving students’ self-efficacy scores increase with time while low achieving students’ self-efficacy scores decrease with time.

In addition to these comments about the research base there are two others which emerged from conducting the review. Studies that disaggregate the data by sex or ethnicity would be enormously helpful to the field. Currently it is often the case that the course in which the research takes place is identified and perhaps the number of males and females is given. When appropriate to the research design and questions researchers should be encouraged to disaggregate the data.

Based upon this eleven-year review it is possible to state that there are a rather small number of findings that CER has established unequivocally. How the research agenda can be shaped to drive forward the field to establish findings in a robust incontrovertible manner would be an outstanding next step.
In addition, funding sources that are renewable to allow for the development of research programs lasting longer than 3-5 years must be widely and consistently available in order for the field to develop robust findings. Long term or renewable funding would allow research programs to evolve and advance rather than to start and stop every 3 to 5 years as funding begins and ends.

Literature Cited

386. Holme, T.A. American Chemical Society Division of Chemical Education Examinations Insitute. 2011 [cited 2011 January 31].