d. Research Narrative

Contribution of Project to Solving an Education Problem
This project targets Goal One at the level of high school students in science. The purpose of this project is to advance the state of the art in science educational software by developing powerful new artificial intelligence (AI) tools for computerized assessment of student work in science, building on and integrating with Quantum's successful interactive intelligent tutoring technology, beginning in the area of chemistry. This work addresses a longstanding, clearly articulated need for improved software tools in chemistry education.

Chemistry comprises the majority of the content standard for physical science in the National Science Education Standards [1], and yet is one of the most neglected areas in terms of quality educational software. Most chemistry tutorial programs are categorized as "computer-assisted instruction" (CAI), which has been in use for a long time [2,3]. The crucial weakness of chemistry CAI is its reliance on a fixed database of problems and possible right and wrong answers as the embodiment of its teaching knowledge. The student can only work with problems that are already programmed into the software. CAI has been criticized for its scripted "linear" instructional approach and shallow level of interactivity as too rigid and inflexible to support meaningful learning or add significant complementary value to other educational resources [4-7]. Another major problem is the lack of diagnostic power for student mistakes and inability to provide instruction on an individualized basis. Typically, the student only selects the final answer in multiple-choice form, and is unable to ask questions or get feedback and analysis on his or her own attempts at solving the problem.

To deliver the next generation of enhanced capabilities users currently demand, a fresh innovative approach is needed. Quantum has incorporated new concepts from the field of artificial intelligence (AI) [8-12] to develop meaningful individualized tutoring technology for chemistry. The fundamental shift from CAI is to replace foreknowledge of specific problems with a direct representation of chemical and pedagogical principles, and then *simulate reasoning* using these principles for the purpose of tutoring students on problems *given* to the program.

Through prior research supported by the Small Business Innovation Research (SBIR) program, Quantum has successfully developed, tested and brought to the classroom intelligent tutoring systems (ITSs) for chemistry education that serve as supplementary materials to any existing curriculum. The full details of this background work can be found in Quantum's publications and SBIR technical reports [13-19]. To accomplish this goal, this work required application of AI techniques from other disciplines [8-12] to chemistry education for the first time, and also the origination of completely new AI education methodologies, for which Quantum has already obtained one patent [20] and currently has another patent pending. Quantum's work is the first to address successfully the longstanding and clearly articulated needs for interactive tutoring in chemistry. By using AI, the Quantum software removed the severe limitation of a fixed database of problems, allowing students to get interactive assistance on problems *they* can enter into the tutor, and get feedback on their *own* work, which they can also enter in detail. In controlled student testing, the intelligent tutors have been documented to help student performance and understanding significantly [17,19], and the technology has been received enthusiastically as meeting important outstanding needs. Ellen M. Standafer, Vice President of Science Product

Development at Holt, Rinehart and Winston, which distributes the Quantum Tutors to schools, states "In my 25 years of chemistry education instructional materials development, I have never seen a software product that compares instructionally and pedagogically with the Quantum programs."

The current proposal involves the extension of these new AI capabilities to the area of student *assessment* in chemistry and integrating these AI assessments with interactive tutoring, using Quantum's successful prior tutoring technology as the foundation. Particularly with the advent of NCLB, the importance to teachers and students of sophisticated high-quality tools for assessment is on a par with interactive tutoring, and the goal of this work is to build improved assessment capabilities to achieve the same power and sophistication as recently achieved for interactive tutoring. This is an exciting new area with many possibilities.

The majority of current software assessment tools in chemistry are based on multiple-choice (MC) tests or similar techniques. While this is sufficient for some simple purposes such as easily graded tests, it is inadequate as a foundation for developing the more sophisticated individualized assessment capabilities needed by teachers. Human teachers can perform much more sophisticated assessments of learning than MC tests provide, and can also give students more beneficial criticism, when they require students to show *all* their work in detail on homework and tests; the problem is that teachers rarely have the time to perform such an in-depth analysis for each and every student. At the same time, the increasing trend is towards requiring schools to compile and analyze more and more detailed reports of student achievement, increasing the demand on the already-overburdened teacher's time. The potential benefits of robust, dependable AI-based software assessment tools are considerable for helping teachers increase their effectiveness and achieve the greatest return with their time and resources.

How does the proposed research fit within the context of current AI work in education? Again, no such capability exists in chemistry software – there has been no prior investigation of these issues with respect to chemistry content and pedagogy, and so the benefits brought to that field are entirely new. As a prime example, in the attached letter included in Appendix A, Dr. Steven Ritter, Senior Cognitive Scientist at Carnegie Learning, expresses his support for the proposed work and notes that it is complementary to their own efforts. In the general picture, the current work has its origins in the ACT theory of Anderson, Corbett, Koedinger, *et al.* [8-12]. [redacted] (b)(4) An important characteristic of the proposed work is that it is complementary to other successful AI-based assessment tools. One important example is the knowledge space theory [21] used in Assessment and Learning in Knowledge Spaces (ALEKS). The innovations proposed here, (b)(4) [redacted] (b)(4) [redacted] have broad potential to enhance assessment systems for content areas beyond chemistry.

Research Plan

Feasibility research on the assessment technology concept and development plan has successfully attracted Small Business Innovation Research Funding from ED and NSF [22,23] to develop initial assessment programs for two individual topics in chemistry and to conduct preliminary field testing (results summarized below). Based on the successful Phase I SBIR research, both projects were recently awarded Phase II funding. This SBIR work has established the feasibility of the project and also provided initial evidence of the technology's effectiveness in the classroom (summarized below).

However, for true practical impact it is necessary that the assessment technology cover the *entire* first-year chemistry curriculum; the scope of this entire development project is greater than can be accomplished within SBIR. Therefore, in this project we propose to build on this successful prior work to *complete* the development of a comprehensive integrated tutoring and assessment system for an *entire* first-year chemistry course, and accumulate further evidence that the system is effective in improving student learning and achievement. The successful prior SBIR work has substantially mitigated the risk factor typically involved in new development research and provided a proven research plan that can be used as a roadmap. In this project, Quantum has established partnerships to work with a team of teachers and schools including at-risk students across a range of rural and urban populations, racial, ethnic and minority groups, and low-income backgrounds.

The intervention being developed here is the ability to place a comprehensive set of AI tutoring systems with integrated AI assessment capabilities in the hands of each and every student in a chemistry class, for use *on demand* during study at home and at school. Though constructing interventions that make a meaningful difference in improving science achievement for disadvantaged and at-risk students is very difficult, it is clear, however, that one requirement for success is that students have the opportunity for sustained, quality practice with science skills and concepts on their own. Put another way, even the most innovative classroom instructional program will have its impact considerably curtailed unless it is supported by students engaging in meaningful practice on a regular basis, and at-home study plays a pivotal role here.

Research has established again and again that one-on-one instruction by a qualified human tutor is a highly effective method for improving achievement. Although hiring a personal tutor for every student is impossible, the Quantum intelligent tutoring technology used as the basis for this work provides breakthroughs over conventional tutorial software in several key respects, that enables it to be effective in one-on-one instructional situations whereas conventional tutorials are typically not significantly effective: students can use their *own* problems, get help on their *own* mistakes, and ask the tutor questions about the concepts and processes being studied and how these relate to their problems. Appendix B gives example computer screens and a sample transcript from a session with an intelligent tutoring system for balancing chemical equations.

From a practical standpoint, it is very important to note that though the Quantum Tutors can be (and are) used during classroom time, this is *not* a requirement and, in particular teachers are *not* required to change the way they teach in order to incorporate them. Rather, the AI technology is primarily intended to act in a *supplementary* role to *reinforce* classroom instruction, by giving students the individual assistance and assessment they need at home to get the most out of that day's lesson. Students use the AI programs simply by logging into an Internet site; the technology can be accessed from any computer with an Internet connection. Since the intelligent tutoring systems and proposed assessment technology work with problems that teachers and students create, it can effectively supplement *any* existing curriculum, as well as be used productively and synergistically with other curriculum materials. By making the AI tutoring and assessment available to students on demand, *including at home*, the quality and effectiveness of at-home study time will be increased, reinforcing and augmenting classroom instruction.

Preliminary empirical evidence that the intervention will be effective is available for the proposed assessment technology in Quantum's publications and SBIR technical reports [18,22,23] as well as for the underlying Quantum intelligent tutoring technology, which has been brought to schools through partnership with a major publisher (Holt, Rinehart and Winston) and shown to be effective in initial small-scale trials using experimental design with randomization [17,19]. Data on effectiveness are summarized in Appendix A for the tutoring and assessment technologies, respectively, from a study on each; further effectiveness results are also available in the aforementioned references. Additionally, interviews with students using the existing tutors show they find them engaging, interesting and fun to interact with.

The feasibility of the proposed assessment concept was demonstrated in SBIR research by developing *automated grading programs* for two different chemistry topics that can grade homework and tests in which *teachers* assign their own problems and *students* show their own work in detail. Both SBIR projects were subsequently awarded Phase II funding. This work provides a solid foundation for continued development of advanced assessment functionalities for helping meet the goals of the No Child Left Behind Act (NCLB). Important benefits of the AI tutoring systems that readily extend to assessment include:

- Ability for the assessment problems to be specified to the system by the user (*e.g.* the teacher), rather than limiting selection to a bank of pre-stored items.
- Deriving grades by analyzing the student's own work in its entirety, which he or she inputs in detail rather than picking multiple-choice answers, thus automatically grading answers for which the teacher instructs the student to "show your work."
- Ability to use a robust assessment rubric based on new conceptually oriented AI methods embodied in the tutors, including ability to analyze unanticipated student errors.
- Ability to embed assessment capabilities in existing online learning systems, and provide results in a format that can be readily interfaced measurement and reporting capabilities of online learning management systems.

No such capability previously existed in chemistry education software. The obvious advantage of MC tests is the convenience they provide because of limited teacher resources. The preliminary assessment work demonstrated the feasibility of providing the very same convenience in grading the student's *own* work in detail, the way teachers would if they actually had the time to devote to it. This itself represents a significant advancement over the status quo, and also serves as a foundation for development of more expansive assessment capabilities. As discussed below, extended development includes functionalities such as the ability to mark comments on student tests as feedback (just as the teacher marks comments in red), and in-depth assessment of multiple-problem histories for diagnosing student strengths and weaknesses.

Table 1 shows the current development status of the integrated AI tutoring and assessment technology project and indicates the portion of the development to be carried out under this proposal. When completed, the result will be a comprehensive system for the fundamental topics and concepts in first-year chemistry.

Table 1. Status of Integrated Intelligent Tutoring and Assessment Technology Development

Module	Intelligent Tutor	Automated Grader	Student and Teacher Assessment Reports	Statistical Analyses
Measurement	+	x	x	x
The Elements	+	x	x	x
Chemical Formulas	+	*	x	x
Mathematics of Chemical Formulas	+	x	x	x
Balancing Equations	+	*	*	*
Chemical Bonding	+	x	x	x
Oxidation Numbers	+	*	*	*
Chemical Reactions	+	*	x	x
Stoichiometry	+	*	x	x

+ = Technology and product development completed and released to schools
* = Development completed or to be completed under SBIR projects
x = Development to be carried out in currently proposed project

As seen in Table 1, the assessment functionalities to be developed for each chemistry topic fall into three major categories: *automated grading*, which constitutes the initial assessment framework for the topic, *student and teacher assessment reports*, which provide detailed feedback on individual student learning and achievement, and *statistical analyses* of assessment information on student groups (*e.g.* class, school, district).

The specific research objectives that will be accomplished are listed below. These will be carried out for each outstanding topic area listed in Table 1. The technical objectives and project scope take into account consultation with students, teachers, administrators and industry partners and are prioritized based on their needs.

Groundwork for assessment development

1) Formulate generalized algorithm for assessing student work on domain problems.
2) Design and implement extensions to intelligent tutoring system framework needed to accommodate assessment functionalities.

Automated grading of authentic student work

3) Implement algorithm for grading student work in existing intelligent tutoring system technology.

4) Design and implement web-based interface for recording student work for assessment purposes.

Validation and review of initial assessment program

5) Collect database of authentic student work samples for testing and validation of prototype assessment system.
6) Refine prototype assessment system based on results of initial testing.
7) Conduct initial external review of prototype assessment system by selected teachers and industry partners.

Advanced learning and achievement assessment features for student and teacher

8) Implement "red pencil" comments on student work as formative assessment feedback.
9) Develop reporting capabilities for teachers on student process and concept mastery.
10) Develop statistical analyses of student learning and achievement based on AI assessments.

External evaluation of the project

11) Formative field testing of selected assessment programs.
12) Evaluation of impact of assessment technology on student learning and achievement.

The detailed plan for each objective is discussed below.

1) *Formulate generalized algorithm for assessing student work on domain problems.*
Before computer implementation can begin, it is necessary to develop a detailed algorithm by which the student's work will be assessed. The main issues of complexity here are:

- A robust algorithm is needed in which the specific problem is given to the program dynamically, *i.e.* is not stored ahead of time. This is nontrivial due to the typical variety of problems involved, but is a necessary requirement to enable teachers to use their own problems for homework assignments and tests. Therefore, we refer to the algorithm to be developed as a *generalized* rubric algorithm.
- Since the student's own work rather than a MC answer is being graded, there must be a consistent and coherent way of handling [redacted] student errors. This is one of the most difficult aspects from the AI perspective since conventional AI techniques [redacted] are much better suited to diagnosing common [redacted] errors.

[redacted] The method should be capable of accounting for all salient features of the problem without explicit knowledge the problem ahead of time. The development of the grading algorithm will be led by Quantum's curriculum specialist, Mr. Dale A. Holder. Mr. Holder is an accomplished outstanding teacher with over 35 years of teaching experience; it is his career teaching expertise that has formed the basis of Quantum's successful intelligent tutoring system development.

(b)(4)

[redacted]

[redacted]

Laying the groundwork for promoting metacognitive examination of mistakes through assessment feedback to the student is very important. This in particular will greatly increase the benefits students can get from software assessment of their own work. More than just a grade, an opportunity will be provided to learn directly from the assessment itself. Therefore, the assessment rubric algorithms developed will include this conceptually-based approach to judging work based on consistency concepts as well as more conventional procedural criteria involving recognition of common errors.

2) *Design and implement extensions to intelligent tutoring system framework needed to accommodate assessment functionalities.*

Before the generalized rubric developed in Objective 1 can be implemented, certain modifications must be made to the underlying intelligent tutoring system to accommodate considerations that occur for grading student work but do not arise in the interactive tutoring case. Here we outline the extensions that must be made to the (content-independent) AI framework of the ITS in preparation for assessment function implementation. The design objective is to generalize the AI framework, which is implemented in C++, to allow the same expert knowledge used by the ITS to be leveraged just as powerfully for assessment purposes. This not only facilitates creation of important new functionality, but also, just as importantly, enables reuse of the same knowledge base, which contains a considerable amount of valuable pedagogical expertise on the topic and required a great deal of time and effort to develop.

The first general design modification is to create a more flexible representation of the knowledge base that is not explicitly tied to the framework of the ITS itself, but rather can simply be invoked by the ITS (as well as other functionalities) as needed. Put another way, a greater decoupling of knowledge base from its (potentially multiple) uses is needed, in the sense of the object-oriented paradigm of loose coupling and tight cohesion. This primarily means a greater separation of the system's [redacted] rules from the specific context of interactive tutoring, and greater decoupling of the output of the analysis of student work by these rules from knowledge of its intended use by the system (*i.e.* whether the objective is assessment or tutoring). Basically, a new layer of separation of the knowledge base from its multiple varied possible uses will be created, exactly in analogy with the design approach of separation of the core functionality of an application from its user interface(s).

The second main technical obstacle is the current "interventionist" design of the ITS. For pedagogical reasons, the ITS was designed to deal with student errors as soon as they occur, and does not allow the student to follow a "blind alley" in the solution path for more than one step before intervening with tutoring to help the student resolve the error. In the context of assessment, this design is not sufficient since no interactive feedback will be provided. The system must be augmented to allow student errors to "propagate," and be able to identify and factor out their effects on subsequent work in order to avoid double-penalizing the student.

The major categories of student step classifications used in the ITS are below, with a summary how each is modified or impacted by the addition of assessment:

- *Correct step*: The conventional definition is a step that appears in the "conflict set" (the set of all possible legitimate next steps in the solution path generated by the system's production rules). In an interventionist ITS only correct steps are generated by the system since only correct partial solutions are used as a starting point for step generation. In assessment, the system must also be given incorrect partial solution paths to analyze, and thus some "correct" steps may actually contain wrong answers as a result of building on incorrect prior work. It will be necessary to distinguish whether an erroneous step is simply the result of propagating a prior mistake or is a new, independent error.
- *Incorrect step*: The conventional definition is a step that is *not* in the conflict set. Given the extension to the correct step analysis described above, this definition may be sufficient.
- *Premature step*: This is defined as a step that does not directly follow from the other work done by the student so far at the granularity of student work representation. This step type must be accepted in the solution path for the purpose of propagating its effect and analyzing whether it does eventually become a correct step after later student work.
- *Uninterpretable step*: This is a step for which no meaning can be successfully inferred by the AI engine (a string of random characters is a trivial example). In the interactive ITS such steps may simply be disregarded and the student prompted to re-enter work, with clarification and scaffolding on how to enter the next step properly, but this is not adequate in the assessment context. Additional AI rules for inferring or partially inferring the intent of the student's ill-formed work will likely be necessary in order to make sure all deserved credit is awarded, so that, for example, a step which may be correct except for a minor secondary error (such as failure to capitalize a chemical symbol) is not invalidated altogether.

3) *Implement algorithm for grading student work in existing intelligent tutoring system technology.*

The next step is to implement the generalized rubric from Objective 1 for the purpose of assigning a grade to a student problem attempt. Using the AI framework extensions developed under Objective 2, this will be a straightforward task. The rubric matrix output will be transformed by a grading algorithm to be developed by Mr. Holder into a score, from 0% to 100% credit for the given problem. For initial purposes, this represents a logical and practical separation of implementing *diagnosis* of the student's work from implementing *feedback* to the student. However, the grading program will be based on the same thorough underlying rubric analysis information that will also be used as the foundation for the more elaborate and advanced assessment capabilities to follow. At the same time, the convenience value to teachers from this functionality alone is of considerable value.

A test set consisting of several mock "student inputs" will be developed for the prototype program, involving a variety of problems designed specifically to exercise all possible facets of the assessment algorithm. The assessment information will be output in similar fashion as the current ITS engines output XML for tutoring.

4) *Design and implement web-based interface for recording student work for assessment purposes.*

In order to develop a robust, in-depth automated assessment system, the student must be able to enter *all* problem-solving work directly into the system, not just the final answer to a problem. This requires an input mechanism that is considerably more sophisticated than those in current programs, which typically use only multiple-choice problems. In Objective 5 below, we will collect a database of student work samples for the purpose of ongoing testing and validation of the assessment engine. This will be done through the Internet, which is also how the technology will be disseminated to schools. Developing a web-based interface for collecting student work for assessment is therefore an immediately needed item.

The components of the student interface that must be developed are:

- A set of concise but effective student instructions for the interface, complete with examples illustrating its use.
- A set of web forms for students to enter their work on a problem.
- A mechanism for serializing and storing student work in a representation that is easy to recall and process.

The ITSs which will serve as the foundation for this work already possess web-based interfaces for entering student work. This is a good starting point, but is not entirely sufficient, as there are some key technical and pedagogical differences between the contexts of interactive tutoring and assessment with implications for the user interface design.

For example, one important requirement is that the order of steps in the student's work must be clearly discernible, as this is a factor in the assessment. In the tutoring context, this is easily handled since the student enters only a single step at a time and the ITS gives feedback immediately after each step. For assessment, however, interactive intervention of the system is obviously not desirable. Similarly, the interface form presented to the student must not be overly restrictive, to capture as rich and faithful a representation of the student's reasoning process as possible, and avoiding the possibility of scaffolding the student as he or she works through the problem. An illustration of an extreme at the other end of this spectrum is the use of multiple-choice questions, since this permits student to guess at the answer without doing any work at all. On the other hand, simply providing the student with a large blank input area is not satisfactory either. For one, the progression of the student's work is not always apparent from handwritten work on a "blank page." The best interface is typically expected to include separate blank input areas and instructing the student to enter each step in their solution in order, one line at a time.

Another main concern is when the student enters work that is uninterpretable or unintelligible to the system. This can readily happen through imprecise usage of chemical notation or merely through typing errors, and so the system cannot simply disregard such entries. With a human teacher/grader, obviously real-time interaction is not typically involved, but the human teacher does routinely interpret work that is "ill-formed" to determine if some credit is warranted in spite of this type of error. In the computer tutoring case, the ITS seeks clarification from the student interactively and also deliberately incorporates scaffolding in the process, neither of which is desirable here.

An important point of investigation is whether it will be necessary in some cases to implement some mechanism for minimal intervention by the system to seek clarification from the student of entries the system cannot interpret. We must ensure that the system is fully capable of interpreting the student's work to a sufficiently accurate degree to ensure that all the credit deserved is awarded, but without the possibility of scaffolding or otherwise altering or interfering with the student's work as a result of the intervention. In chemistry, problem-solving work fortunately can often be fully represented in chemical and mathematical symbolic notation that is readily interpreted by computer, so the issue of ill-formed work is not nearly as problematic as it would be with entirely free-form natural language, for example. The extensions to the ITS to be developed in Objective 2 will give the system some ability to "factor out" certain of these types of errors which should not materially affect the assessment, but burdening the system by requiring significant natural language interpretation, for example, is not practical.

XML will be used for representation and storage of the data collected by the interface, which permits flexible custom tagging of student input. The original ITS engines already work with XML, and XML representation of assessment information is an important requirement that has already been expressed to us for the system (see Objective 10).

5) *Collect database of authentic student work samples for testing and validation of prototype assessment system.*

For the purpose of collecting student work samples for initial testing and validation of the assessment system, we have partnered with several high schools (see Objectives 7, 11 and 12). Quantum also has an established ongoing relationship with Duquesne University that includes several successful past collaborations. A large database of authentic student work will be collected anonymously from students in first-year chemistry.

Initial data collection merely involves participating students working a set of problems designed by Quantum using the web-based interface developed in Objective 4. The scope of the problem set will be comparable to a single typical homework assignment, and the interface will be conveniently accessed through the Internet at Quantum's web site. Students will interact with the system anonymously, and the system will simply record student work in XML format in a database for later analysis – no feedback or tutoring will be involved in this activity. Data collection will be conducted at the normal time the particular topic is covered in the first-year chemistry curriculum. This will minimize the effort required on the part of participating students and instructors, minimizing any potential disruption.

6) *Refine prototype assessment system based on results of initial testing.*

Testing of the prototype using the student work database collected in Objective 5 will reveal ways the assessment system can be improved and perhaps also uncover unanticipated issues that will need further research focus. This will be quite valuable for enhancing the quality of the system in advance of piloting the actual assessment system in schools in Objectives 11 and 12.

7) *Conduct external review of prototype assessment system by selected teachers and industry partners.*

The involvement of schools and industry partners from inception in review of the technology ensures it is clearly designed to meet user needs and requirements. With all initial assessment systems for each individual topic, review and comment by teachers at selected partner schools will be conducted beginning in Year 1 and continuing for the duration of the project. Further evaluation activities for Years 2 and 3 are discussed in Objectives 11 and 12. Field testing in Year 1 of the project will consist primarily of formative evaluation via one-on-one interaction with selected teachers at a small number of schools. The first year will be focused primarily on development and comparatively few modules will be ready for testing during that time. For those modules that will be available early on, input and feedback will be gathered from teachers to inform improvements that can be made before actual trial with students. For these early review and evaluation activities, Quantum has partnered with schools in two rural systems in the Appalachian region, Russell Co., KY and Greene Co., PA (see letters of commitment in Appendix A). These schools will also continue with evaluation activities in Years 2 and 3; as a greater portion of the assessment system is developed, a more extensive formative evaluation will be conducted with a larger number of teachers and additional schools as described in Objectives 11 and 12.

Quantum has also established relationships with several major online educational providers, such as Holt, Rinehart and Winston, Prentice Hall and Pearson Learning Group to gather their input on the current project. These organizations have unanimously expressed enthusiasm for the proposed technology to meet needs for improved assessment functionalities, and already several practical issues have been identified, such as representation of assessment information in XML, embedding of assessment within existing online learning activities, and compatibility/interoperability with online learning management systems such as Blackboard and WebCT.

8) *Implement "red pencil" comments on student work as formative assessment feedback.*
The next category of objectives concerns development of more advanced assessment capabilities, building on the successful prior work. Better informing students and teachers so they can make better decisions to improve learning and achievement is a theme throughout this project. The next step beyond AI grading is to implement the ability for the system to generate individualized comments on student tests during the grading process as feedback. This is analogous to teachers writing comments in red on student papers.

The initial grading program represents a logical and practical separation of implementing *diagnosis* of the student's work from implementing *feedback* to the student and teacher, but is based on thorough underlying AI analysis that also lays the groundwork for building much more elaborate and advanced capabilities. Why? The key is that the AI rule-based representation has the link between problem work and its conceptual justification *built in* from the start, because the system works with the concepts directly rather than by rote. That is to say, the program is actually doing assessment of the student's own work by *reasoning* (in a simulated way) using the underlying chemical principles. This not only makes the proposition of explaining the student's shortcomings more plausible, it permits the practical implementation to share the core analysis code with the grading program and intelligent tutoring system.

This important property is demonstrated by a simple example. Consider the following solution for assigning oxidation numbers to strontium sulfite, $SrSO_3$, obtained by one student in initial validation activities, who assigned the oxidation numbers as follows:

Step 1: Sr = +1
Step 2: O = -6
Step 3: S = +5

The correct values are Sr = +2, O = -2 and S = +4. However, even though none of the student's answers is literally correct, there are several attributes of this solution that demonstrate partial understanding, and these are elucidated during the AI analysis. Though these details are not currently exposed when assigning a numerical grade, this generated information already contains the essential "raw material" for formulating an excellent assessment of the student's solution. This internal information must be transformed and refined into statements suitable for the student and/or teacher, just as is done with the AI's simulated reasoning in the intelligent tutoring systems. This is not trivial, but with the design and internal detail produced by the AI system as a starting point, it is now plausible to create formative feedback as "red pencil" comments. The example feedback below, which conveys the essence of the system's reasoning process, illustrates the capabilities planned.

- "Your oxidation number of Sr = +1 is not correct. Strontium is a Group 2 metal and so it always has an oxidation number of +2 in compounds. Perhaps you thought Sr is in Group 1 instead? Be sure you are clear on strontium's position on the periodic table; that is very important for this problem."
- "It looks like you wrote the correct total oxidation number of the three oxygen atoms, which is -6 as you said. However, chemists always write oxidation numbers for the individual atoms, not the total, so in the future you should write your answer as O = -2."
- "Your value of S = +5 does properly conserve charge given the oxidation numbers you assigned the other atoms. However, the problem is that the incorrect value you gave Sr before caused you to get the wrong answer for S here. The correct value is S = +4. See if you can get that answer using the right value for Sr."

Students will benefit tremendously from such detailed, individualized formative assessment feedback. The feedback is specific, leads students directly to the areas where there is a gap in their understanding, and in this way helps them not only to build content knowledge, but also to develop confidence in their ability to approach a similar problem. These types of comments, consisting of the proper mix of constructive criticism and positive reinforcement, are carefully constructed to maintain the best possible student attitude and are well received by students. Research shows that formative assessment used frequently as feedback to individual students is one of the most effective strategies available to teachers in meeting high standards of student learning [24]. In fact, this type of strategic and focused feedback has been shown to increase student efficacy since students are armed with the knowledge of the specific steps they can take to succeed in their next attempt at a similar problem. This AI tool will be extremely practical in this regard, enabling students to continually and consistently receive this level of quality feedback.

The online assessment information provided will furthermore be integrated directly back to the corresponding intelligent tutoring system for that topic. From the assessment feedback for any individual step in the student's solution attempt, for further assistance at the click of a button the student will be able to invoke the corresponding ITS to obtain interactive tutoring on the very same errors made on the test. This is an important and powerful example of integration of the AI assessment and tutoring capabilities, and has already been successfully implemented for an initial module in the SBIR research.

(b)(4)

[redacted]

This gives students good examples of how to examine their own work critically in the specific context of relevant concepts. This encourages students to develop and apply critical thinking skills by directly and clearly demonstrating their benefits, while at the same time providing them with the grist for doing so by pointing out applicable principles in the context of their own mistakes. This increased emphasis on student metacognition promotes critical thinking by providing a way for the student to examine his or her own work the *next* time, discovering similar errors without having to be told. Promoting metacognitive examination of mistakes through assessment feedback to the student is a very important factor. This will greatly increase the benefits students can obtain from software assessment of their work.

9) *Develop reporting capabilities for teachers on student process and concept mastery.*
The previous objective will make available a considerable amount of detailed diagnostic information for each single problem attempt. The next logical step is to build on this to develop analysis capabilities for extracting and condensing the most important information from a *group* of problem attempts. Multiple-problem analysis will be used to reveal patterns and trends in student errors, identify and characterize conceptual and procedural strengths and weaknesses, and provide summary reports on the current level of student achievement.

Reports will be generated for both the student and the teacher. Information can be broken down in several ways, such as by individual concepts, skills, and level of mastery. [redacted] (b)(4) An illustration of the analysis to be developed appears in Figure 2 below, in a mock assessment summary report for a struggling student who has taken a ten-problem quiz on oxidation numbers.

Figure 2. Mock Student Assessment Summary Report for Oxidation Numbers Quiz

Overall score:
- Your overall score for this quiz is 57.3%. That would indicate that some improvement is needed for success with later concepts that build on oxidation numbers.

Basic purpose:
- Your understanding of the basic objective in assigning oxidation numbers appears to need improvement. When elements were attempted, you assigned the correct oxidation number only 63% of the time.

Taking an appropriate first step:
- Selecting the element to assign first is one of your strong points. You selected an appropriate element to assign first 80% of the time.

Order of steps:
- It is important to remember that not all elements have their own specific assignment rules. Usually at least one element in a problem can only be assigned by proper application of the Sum Rule at the end. Your steps were in proper order only 50% of the time.

Common elements:
- You did a good job assigning oxidation numbers to the common elements that have their own oxidation number rules, like hydrogen and oxygen. On this problem set you were successful with these elements 87% of the time.

Polyatomic ions:

- Dealing with polyatomic ions seems to be a significant problem for you. It is important to remember that the sum of the oxidation numbers on any polyatomic ion does not equal zero. Instead, it must equal the charge on the ion. When polyatomic ions were involved, you were successful only 33% of the time.

Charge conservation:

- Conservation of charge is the most fundamental concept in oxidation numbers. Since violation of this concept is a very serious error, the Sum Rule is included as a reminder. You have successfully conserved charge only 73% of the time in this problem set.

Suggestions for review:

- From your performance on this problem set, I would suggest you review the following points:
    - Assigning oxidation numbers in polyatomic ions
    - Conservation of charge (Sum Rule)
    - The distinction between valence and oxidation number

Not only will students benefit from availability of this practical formative assessment information [24], it will also be a boon to classroom teachers, who can obtain individualized assessment reporting and diagnostic information for any student on demand, as if from a "virtual teaching assistant." Just as our initial assessment research showed that automated grading can be carried out with substantially increased reliability and accuracy, AI assessment reporting will give teachers information that is more thorough and detailed than can ever be compiled by hand. This directly serves the goal of providing high-quality diagnostic information for better informed decision-making and planning, such as in meeting goals of NCLB. It should be particularly noted that another substantial advantage of the AI diagnostic assessment is that it embodies the expertise of a proven outstanding career teacher (see Section f, Mr. Dale A. Holder).

10) *Develop statistical analyses of student learning and achievement based on AI assessments.*

The detailed assessment information developed in prior Objectives naturally lends itself to investigation of systemic patterns across classrooms, grade levels, schools and districts. To aid schools in meeting the goals of NCLB, publishers are already developing statistical analysis tools for conventional assessments such as multiple-choice tests. Quantum is bringing this statistical analysis to bear on the AI assessment information as well. To begin, these existing statistical tools will simply be applied to the richer, more detailed AI assessment data directly without modification. This application will immediately reveal trends in a given student population on process and concept strengths and weaknesses, by directly calculating, for example, the percentage of students diagnosed by AI analysis as needing more work on any selected individual concept. If desirable, based on end-user needs, the potential also exists for determining statistical correlations among various strength and weakness trends. In practice, the external evaluations will help indicate any modifications and refinements needed that do not arise in conventional statistical assessments.

The technology will make it practical to conduct automated AI assessment on a wide scale in helping teachers and administrators meet the goals of NCLB. Teachers whose instruction is supported by the AI tutoring and assessment technologies can harness the systematic feedback to collect, organize, analyze, disaggregate and report on student achievement data. Student achievement data is complex, but with the AI technologies there are tremendous opportunities for teachers and groups of teachers to identify strengths and weaknesses in curriculum and instruction when the feedback data that the AI can provide is properly analyzed and synthesized.

11) *Formative field testing of selected assessment programs.*

As discussed earlier in Objective 7, field testing in Year 1 of the project will consist primarily of formative evaluation via one-on-one interaction with selected teachers at a small number of schools (see letters of commitment from Greene Co., PA and Russell Co., KY schools in Appendix A). The first year will be focused primarily on development and comparatively few modules will be ready for testing during that time. For those modules that will be available early on, input and feedback will be gathered from teachers to inform improvements that can be made before actual trial with students. In Years 2 and 3, as a greater portion of the assessment system will have been developed, a more extensive formative evaluation can be conducted with a larger number of teachers and schools as described below.

As part of ensuring the technology is as effective as possible, *external* evaluation will be an important focus. The first such evaluation, during Year 2, will be field-testing of then-available AI programs for automated grading and the accompanying intelligent tutoring systems. This will be mainly a formative evaluation in advance of larger-scale testing in the final year of the development project (see Objective 12). This testing will yield valuable practical information on improvements and possible additional features during the development process.

In addition to gathering useful feedback in the context of larger, "real-world" testing on the value added by eliminating inconsistencies in assessment due to human error, deriving grades by analyzing the student's own work in its entirety, and the program's ability to analyze unanticipated student errors, we will also gather feedback on whether the convenience factor of automated grading is realized for teachers in a practical setting, and to determine features that increase student use of the tutorial system. This evaluation activity will serve to strengthen the product's practical value. The Year 2 formative field-testing will be conducted with approximately 10 teachers and their students.

For this objective, Quantum has obtained the services of experienced psychometricians at Wexford, Inc. (see Section f) to design and conduct external project evaluation. Wexford is the evaluator on Quantum's current SBIR projects and Quantum has an established working relationship with them. The evaluation plan outline developed by Wexford is as follows:

A. Questions to be Investigated

1. Does the technology perform to specifications?
2. Does it do so in a variety of circumstances?
    a. Connectivity: variety of connection speeds, cross-platform
    b. Access: from home, school and community points, including firewall issues
    c. Support: variety of support (from teachers or technologists)
    d. Technological Backgrounds: of teachers and learners
    e. Learner Profile: learning preferences, learning prerequisites

B. Recruitment of Participants in Field-Testing

Recruit 10 to 12 chemistry teachers each of whom teach four chemistry sections. During Year 2, the teacher will field-test the AI grading program with one section (class), and will field test the student tutorial program with another section. Teachers will be selected in California, Kansas, Kentucky, Maryland, New Mexico and Oklahoma to include a diverse group based on:

- Teacher background: years of teaching, major-minor-out of area
- Geographical: multi-state, rural-urban
- Students: multi-ethnic, year of student, socio-economic background, gender

Additional participating schools will be arranged by Wexford, as they have done in collaborating on Quantum's SBIR projects (see letter of commitment, Appendix A). Quantum's alliance with Wexford provides access to schools across rural and urban populations with students across a range of racial, ethnic and minority groups and low-income backgrounds.

C. External Evaluation Process and Timelines: Year 3

Evaluators, with information and support from the project director will complete the following tasks.

Task	**Timeline**
Identify field-testing teachers and their students (at 10 to 12 sites).	Quarter 1
After introduction of grading and tutorial programs, monitor uses by all participants.	Quarters 2 - 3
Review performance of technology program to determine if it meets specifications, and gather information on teacher attitudes and acceptance factors, convenience level, integration with and ease of use in classroom context, perceived value.	Quarters 2 - 3
Conduct focus groups of teachers at each site to determine perceived strengths, benefits, and needs of participants related to capabilities of program. What is teacher confidence in grading program, would they continue to use it? Has the tutorial program been helpful to students? What would increase its usability? Provide feedback to project director.	Quarter 3
Conduct surveys of teachers and students.	Quarter 3
Summarize and analyze data; complete report.	Quarter 4

D. Foundation of Field-Testing Design

The foundation of the field-testing design is the Design-Based Research paradigm. Shavelson, *et al.* [25] discuss the value of design studies. They posit that "The strength of design studies lies in testing theories in the crucible of practice; in working collegially with practitioners co-constructing knowledge, in everyday classrooms, school and community problems that influence teaching and learning and adapting instruction to these conditions; in recognizing the limits of theory and in capturing the specifics of practice and the potential advantages from iteratively adapting and sharpening theory in its context." The field-testing for this project is characterized by five of the hallmarks of design studies, in that it is:

1. Iterative: involves tightly linked design-analysis-redesign cycles.
2. Process-focused: traces an individual's (or group's) learning by understanding successive patterns in the reasoning and thinking displayed and the impact of instructional artifacts on that reasoning and learning.
3. Interventionist: testing theory and instructional artifacts by designing and modifying real-world settings.
4. Often multi-leveled in linking classroom practices to events or structures outside the classroom.
5. Utility-oriented: intent on improving the effectiveness of instructional tools to support learning; theory-driven in the sense of testing and advancing theory through the design-analysis-redesign of instructional activities and artifacts.

12) *Evaluation of impact of assessment technology on student learning and achievement.*
A major goal of the project is to give students and teachers the high-quality, real-time assessments they need so they can make better-informed decisions to improve learning and achievement. The AI technology can analyze, compile and report high-quality assessment information with depth and breadth that teachers simply would never have the time to create on their own. The second major evaluation effort, in Year 3, is field-testing of the new features and new content of the program, and a study to measure the impact on student achievement.

A. Questions to be Investigated
1. Does the technology perform to specifications?
2. Does it do so in a variety of circumstances?
3. Does the use of the expanded AI program improve student achievement?

B. Participants in the Study
The 10 to 12 teachers and their students (1 class per teacher) who participated in Year 1 continue in field-testing of extended AI assessment program and AI tutor. (These 10 to 12 sections of students are Group 1). With a second class the teacher continues his/her normal instruction, not using any functions of the AI program (Group 2). With a third class, s/he uses the expanded version of the AI program, but not the tutorial (Group 3). With the fourth class, s/he uses only the student AI tutoring program (Group 4).

C. External Evaluation Process and Timelines: Year 3

Task	Timeline	Responsibility
Project director introduces teachers to expanded versions of the AI program and tutorial as is appropriate for what functions they will be using for the study.	Before school starts	Project Director
Pre-test students in all 40 chemistry sections.	1st week of school	Teachers
Score pre-test.	September	Evaluators
Proceed as in Year 2 design.	Quarters 1-4	Evaluators
Conduct post-test of students.	Quarter 4	Teachers
Summarize data for Groups 1 – 4; Compare results of Groups 1 – 4.	May	Evaluators

D. Foundation of Design
   1. Design-Based Research Paradigm described in Objective 11 (responds to Questions 1 and 2).
   2. Quasi-Experimental Design (responds to Question 3).
      Using a pre-post design, with a Treatment 1 group, Treatment 2 group, Treatment 3 group, and Comparison group. Because it is not possible to have a randomly selected group in this test, statistical analysis will be used to determine the similarities and differences between the Groups based on the characteristics in Objective 11, Item B.

For information on human subjects in research, please see Section a, ED 424, 4i-m.

Personnel
As requested, brief information on key project team personnel is provided here with full details in Section f. Quantum has assembled a research team that collectively possesses considerable experience and strong expertise in chemistry content, pedagogy, psychometrics and evaluation. The project director holds a Ph.D. in chemistry from Carnegie Mellon University working with a Nobel Laureate, and has years of experience in developing cutting-edge artificial intelligence for education. The chief curriculum and pedagogy expert is an outstanding teacher with over 35 years of experience, who is an NCLB "Highly Qualified Teacher" and has been recognized as the most outstanding secondary science teacher in his home state of Kentucky. The external evaluation organization, Wexford, Inc. is expert in scientifically based evaluation methods and is recognized nationally for its leading work in collaboration, strategic planning, evaluation, and equity. Also, as previously discussed, Quantum has established partnerships to work with a team of teachers and schools including at-risk students across a range of rural and urban populations, racial, ethnic and minority groups, and low-income backgrounds. Furthermore, once completed, national dissemination of the resulting technology will occur rapidly through Quantum's established distribution network with long-term partners such as Holt, Rinehart and Winston.

Resources
The major project resource requirement consists of standard software development resources, available to Quantum in-house, and software development personnel, which will consist of existing staff developers and developers to be hired. Resources required at school field settings are minimal and commonly available; the main requirement is Internet access in school computer labs and for participating students at home. The external evaluation resources necessary are provided by the external evaluator and are included in cost of subcontract to Wexford, Inc. for evaluation.

e. Reference List

[1] National Research Council, National Science Education Standards (National Academic Press, Washington, DC, 1996).
[2] S. Lower, G. Gerhold, S. G. Smith, K. J. Johnson and J. W. Moore, "Computer-assisted instruction in chemistry," *Journal of Chemical Education* **56**, 219 (1979).
[3] M. A. Bell, R. P. Gladwin and T. A. Drury, "Implementing CAL in chemistry: Practical issues," *Journal of Chemical Education* **75**, 781 (1998).
[4] S. Pavlinic, A. H. Wright and P. D. Buckley, "Students using chemistry courseware – Insights from a qualitative study," *Journal of Chemical Education* **77**, 231 (2000).
[5] G. Kearsley, "Educational technology: A critique," *Educational Technology* **38**, 47 (1998).
[6] W. E. Brown, *Science Teachers Workshops for Computer Training*, NIH-SEPA (NCRR) Award #RR07642 (1997).
[7] J. Beck, M. Stern and E. Haugsjaa, "Applications of AI in education," *ACM Crossroads* **3.1** (1996), http://www.acm.org/crossroads/xrds3-1/aied.html.
[8] The Advanced Computer Tutoring Project, Carnegie Mellon University, Pittsburgh, PA, http://act.psy.cmu.edu/ACT.
[9] J. R. Anderson, *Rules of the Mind* (Lawrence Erlbaum Associates, Mahwah, NJ, 1995).
[10] K. R. Koedinger and J. R. Anderson, "Effective use of intelligent software in high school math classrooms," Proceedings of the 1993 Conference on Artificial Intelligence in Education. Charlottesville, VA: AACE (1993).
[11] K. R. Koedinger, J. R. Anderson, W. H. Hadley and M. A. Mark, "Intelligent tutoring goes to school in the big city," Proceedings of the 7th World Conference on Artificial Intelligence in Education, Washington, DC (1995).
[12] J. R. Anderson, A. T. Corbett, K. Koedinger and R. Pelletier, "Cognitive tutors: Lessons learned," *Journal of the Learning Sciences* **4**, 167 (1995).
[13] B. G. Johnson, *Advanced Software for Interactive Chemistry Tutoring*, NSF SBIR Phase I Award #DMI-0060379 (2000) and Phase II Award #DMI-0132003 (2001).
[14] B. G. Johnson, *Advanced Software for Interactive Learning of Applied Mathematics*, ED SBIR Phase I Award #EDOERI-00-000491 (2000) and Phase II Award #EDOERI-01-000422 (2001).
[15] D. A. Holder, B. G. Johnson and P. J. Karol, "A consistent set of oxidation number rules for intelligent computer tutoring," *Journal of Chemical Education* **79**, 465 (2002).
[16] B. G. Johnson and D. A. Holder, "A cognitive modeling tutor supporting student inquiry for balancing chemical equations," *The Chemical Educator* **7**, 297 (2002).
[17] M. B. Walsh, C. M. Moss, B. G. Johnson, D. A. Holder and J. D. Madura, "Quantitative impact of a cognitive modeling intelligent tutoring system on student performance in balancing chemical equations," *The Chemical Educator* **7**, 379 (2002).
[18] B. G. Johnson and D. A. Holder, "An artificial intelligence-based program for automated grading of student work in balancing chemical equations," *The Chemical Educator* **8**, 271 (2003).
[19] B. G. Johnson, D. A. Holder, M. C. Wheeler and K. L. Willis, "Quantitative impact of a cognitive modeling intelligent tutoring system on student performance in balancing chemical equations II: Impact at the high school level," *The Chemical Educator* (under review).

[20] B. G. Johnson and D. A. Holder, "Intelligent tutoring methodology using consistency rules to improve meaningful response," patent number US 6,540,520.

[21] J.-C. Falmagne, M. Koppen, M. Villano, J.-P. Doignon and L. Johannesen, "Introduction to knowledge spaces: How to build, test and search them," *Psychological Review* **97**, 201 (1990).

[22] B. G. Johnson and D. A. Holder, *Artificial Intelligence Software for Student Assessment in Chemistry Education*, ED SBIR Phase I Award #R305S020088 (2002) and Phase II Award #R305S030012 (2003).

[23] B. G. Johnson, *Artificial Intelligence Software for Student Assessment in Chemistry Education*, NSF SBIR Phase I Award #DMI-0232223 (2003) and Phase II Award (2004).

[24] P. Black and D. Wiliam, "Inside the black box: Raising standards through classroom assessment," *Phi Delta Kappan* **80**, 139 (1998).

[25] R. J. Shavelson, D. C. Phillips, L. Towne and M. J. Feuer, "On the science of education design studies," *Educational Researcher* **32**, 25 (2003).

f. Curriculum Vitae of Key Personnel

Principal Investigator
Benny G. Johnson*, Ph.D., President and CEO, Quantum, Simulations, Inc.*, holds Bachelor of Science degrees in Chemistry and Mathematics from the University of Kentucky, and received his Ph.D. in Theoretical Chemistry from Carnegie Mellon University in 1993, working with Professor John A. Pople, a 1998 Nobel Laureate in Chemistry. During his graduate work, Dr. Johnson focused heavily on the research and development of methods for quantum mechanical study of large molecules. He is world-recognized as a leading expert in the field of theoretical chemistry and as an accomplished software developer with fifteen years of experience writing scientific and educational software. Dr. Johnson is the author of over forty scholarly publications in academic journals and books, and has delivered invited lectures at many national and international conferences.

Over the past seven years, Dr. Johnson has developed the results of his scientific research and successfully released a full-featured quantum chemistry software product to a worldwide commercial market of industrial, academic and government researchers. As principal investigator on various research and development projects, Dr. Johnson has received over $2.5 million in SBIR and other awards through various agencies including the U.S. Department of Education, the National Science Foundation, the Department of Energy, the Air Force Office of Scientific Research and the Ben Franklin Technology Center of Western Pennsylvania. In addition to this work, Dr. Johnson has contributed substantially to several significant commercial and government chemistry research software projects, including Pacific Northwest National Laboratory's NWChem project, which is part of the Department of Energy's effort to solve environmental problems related to the Hanford nuclear site in Washington state.

Dr. Johnson has been involved in education at various levels for several years and has a comprehensive understanding of the teaching techniques, curriculum and content required to educate students of chemistry and mathematics. Dr. Johnson experienced outstanding success as a private tutor of chemistry and mathematics while at the University of Kentucky and, as a graduate student, he received an Outstanding Graduate Teaching Award from the Carnegie Mellon University Department of Chemistry. Based on his work in advancing the state of the art in intelligent tutoring systems for science and mathematics, he was recently invited to serve on the U. S. Department of Education's committee on the SCORM (Sharable Content Object Reference Model) initiative. SCORM is an emerging standard for classification and delivery of online educational content. Dr. Johnson's role is to provide expertise to the agency on intelligent tutoring systems as the SCORM standards are developed in this area. Dr. Johnson's proven ability in and commitment to technology, coupled with his education and research experience, provide the essential leadership and expertise required to complete the objectives of all phases of product development.

Dr. Johnson's selected publications relevant to the proposed work are listed below.

- D. A. Holder, B. G. Johnson and P. J. Karol, "A Consistent Set of Oxidation Number Rules for Intelligent Computer Tutoring," *Journal of Chemical Education* **79**, 465 (2002).

- B. G. Johnson and D. A. Holder, “A Cognitive Modeling Tutor Supporting Student Inquiry for Balancing Chemical Equations,” *The Chemical Educator* **7**, 297 (2002).
- M. B. Walsh, C. M. Moss, B. G. Johnson, D. A. Holder and J. D. Madura, “Quantitative Impact of a Cognitive Modeling Intelligent Tutoring System on Student Performance in Balancing Chemical Equations,” *The Chemical Educator* **7**, 379 (2002).
- B. G. Johnson and D. A. Holder, “An Artificial Intelligence-Based Program for Automated Grading of Student Work in Balancing Chemical Equations,” *The Chemical Educator* **8**, 271 (2003).
- B. G. Johnson and P. L. Corio, “Computer Construction of Reaction Mechanisms,” *Journal of Physical Chemistry* **97**, 12100 (1993).
- B. G. Johnson, “Development, Implementation and Applications of Efficient Methodologies for Density Functional Calculations,” in *Modern Density Functional Theory: A Tool for Chemistry*, J. M. Seminario and P. Politzer, Eds. (Elsevier Science B.V., Amsterdam, 1995), vol. 2, p. 169.
- B. G. Johnson, P. M. W. Gill and J. A. Pople, “The Performance of a Family of Density Functional Methods,” *Journal of Chemical Physics* **98**, 5612 (1993). *The most cited paper in Chemistry for November-December 1994, as reported in Current Contents.*

Dale A. Holder*, M.S., Chief of Curriculum Design, Quantum Simulations, Inc.,* holds Bachelor of Science degrees in Mathematics and Chemistry from Eastern Kentucky University, a Master of Arts degree in Science Education from Western Kentucky University, and a Rank 1 certification as an Educational Specialist from the University of Georgia. Mr. Holder has experienced outstanding success during his 35-year tenure as a high school teacher of chemistry and mathematics in the Kentucky public education system. He has also taught at Western Kentucky University in the advance-placement program allowing high school students to receive college credit. Due to his excellent record and outstanding teaching techniques, Mr. Holder was invited to provide instruction for freshman chemistry at the University of Kentucky’s Somerset Community College, where he taught on a part-time basis. In recognition of his teaching achievements and commitment to excellence, Mr. Holder has been honored at the county, state and university levels with twelve Outstanding Teacher Awards. The Kentucky Academy of Science named him Kentucky’s Outstanding Secondary Science Teacher for 1999. Mr. Holder meets NCLB requirements as a “Highly Qualified Teacher.”

For many years, Mr. Holder has taken advantage of university programs offering high school students the opportunity to enter statewide competitions in chemistry. Mr. Holder ensured that his students were able to further their knowledge and increase their confidence by extensively preparing and coaching them. His students consistently won first and second places against students from much larger schools with more modern facilities and resources, and higher levels of funding. Furthermore, in an economically depressed area where there has historically been a low matriculation rate of students to any higher education program, six of Mr. Holder’s students have pursued and received a doctoral degree in chemistry, including Dr. Johnson.

Despite only a limited budget for instructional materials and teaching aids, Mr. Holder maintained exemplary dedication to providing interesting and thought-provoking high school chemistry programs for every student. As an example, he has developed several innovative classroom teaching aids which he has successfully patented, and which are commercially distributed and marketed by Science Kit & Boreal Laboratory. Proven to enhance the ability of students to grasp and visualize key chemistry topics, these innovative aids include:

- "Magnetic Molecular Models Kit" - Using magnetic disks, the students see the recombination of elements as the teacher arranges them on the blackboard
- "Crystallization on the Overhead Projector" - Incorporating an overhead projector, this kit uses petri dishes to grow crystals which can be projected directly onto the classroom screen
- "Ionic and Molecular Interaction Demonstration Model" - Incorporating an overhead projector, this kit utilizes marbles, a variety of specially shaped magnets and inert plastic chips
- "Alien Formations: A New Silicate Garden" - A fascinating new twist to a traditional colloidal formations laboratory designed to spark interest in students of all ages, motivating them to look further into the wonders of chemistry
- "Old-Fashioned Appalachian Soap-Making Kit" - Illustrates the slow rate of many organic reactions through a hands-on activity with interesting historical and cultural connections

Mr. Holder's selected publications relevant to the proposed work are listed below.

- D. A. Holder, B. G. Johnson and P. J. Karol, "A Consistent Set of Oxidation Number Rules for Intelligent Computer Tutoring," *Journal of Chemical Education* **79**, 465 (2002).
- B. G. Johnson and D. A. Holder, "A Cognitive Modeling Tutor Supporting Student Inquiry for Balancing Chemical Equations," *The Chemical Educator* **7**, 297 (2002).
- M. B. Walsh, C. M. Moss, B. G. Johnson, D. A. Holder and J. D. Madura, "Quantitative Impact of a Cognitive Modeling Intelligent Tutoring System on Student Performance in Balancing Chemical Equations," *The Chemical Educator* **7**, 379 (2002).
- B. G. Johnson and D. A. Holder, "An Artificial Intelligence-Based Program for Automated Grading of Student Work in Balancing Chemical Equations," *The Chemical Educator* **8**, 271 (2003).
- N. W. Hunter and D. A. Holder, "A Teacher-Built Device to Demonstrate Molecular Motion," *Journal of Chemical Education* **69**, 63 (1992).

Selma L. Sax, *Senior Evaluator, Wexford, Inc.,* has more than ten years of experience in Evaluation, Strategic Planning, Policy Initiatives and Public Service. She holds teaching credentials in Massachusetts and California and a Pupil Personnel Credential in California. Ms. Sax was a school board member for eight years, serving as president of the Los Angeles Unified School District board twice. She has worked with political and educational leaders to assist them in planning and implementing educational technology in K-12 public education to improve teaching and learning. Governor Wilson of California appointed her for two terms to the Education Council for Technology in Learning where she served for two years as the Chairperson of the Council. Ms. Sax has also served as one of President George W. Bush's advisors on policy initiatives in educational technology to improve teaching and learning. Currently she is the Director of the Center for Quality Teachers at the Wexford Institute, and Co-Director of PT3*L3 "Learning, Linking and Leading," Wexford's 2001 Catalyst Grant funded through the Preparing Tomorrow's Teachers to Use Technology grant program at the United States Department of Education. She is the lead evaluator for three other PT3 projects and for the High Plains Regional Technology Consortium all of whom contract with Wexford, Inc as the evaluation agency. Ms. Sax also serves as the consulting evaluator on the MathSTAR project, a Star Schools program in its third year, and continues to do the same for TEAMSImpact, an eleven-year Star Schools grantee. Ms. Sax and Ms. Cassidy of Wexford, Inc. twice chaired the Evaluation Institute for the U.S. Department of Education's PT3 Grantees Meetings (2001/2002) to help train evaluators for PT3 grant programs.

Sheila Cassidy, *Executive Director, Wexford, Inc.,* has 25 years experience as an evaluator. She is committed to equity and excellence in education for all students, with a focus on the traditionally underserved, particularly children of color, language minority students and students from low income families. Ms. Cassidy has been a teacher in Philadelphia, Pennsylvania and Compton, California, as well as an administrator of bilingual programs in Lawndale, California. She completed her doctoral coursework at the University of Southern California in Curriculum Change and Evaluation. She has conducted research and evaluation for local, state, regional and national agencies and programs. Over the last decade, Ms. Cassidy has assisted agencies in program development (planning, design, funding, implementation and evaluation). She has authored instructional materials, evaluation and research reports and articles. She co-planned and facilitated the 1997 Distance Education Forum, the 1998 Distributed Education Forum and the 1998 Educators Summit on Evaluation of Technology Integration and Distributed Education, sponsored by the Office of Educational Research and Improvement in the U.S. Department of Education. She is the co-author of "*Developing a Vision for Distance Education in the 21st Century*."

Wexford, Inc. assists agencies and individuals to collaborate to solve educational issues, to create equitable and excellent educational opportunities for all learners. Wexford is recognized nationally for its work in collaboration, strategic planning, evaluation, and equity.

g. Budget Narrative

	Year 1	Year 2	Year 3
Personnel			
Benny G. Johnson, Project Director (b)(6) [redacted]	$26,250	$26,250	$26,250
Dale A. Holder, Curriculum Design (b)(6) [redacted]	14,000	14,000	14,000
Rebecca Renshaw, Administrative Support (b)(6) [redacted]	9,000	4,500	9,000
Gregory Voss, Financial Management (b)(6) [redacted]	4,000	4,000	4,000
(2) Software Developers (b)(6) [redacted]	124,000	130,200	136,710
(1) Q/A Analyst (b)(6) [redacted]	59,000	61,950	65,048
Total Personnel	**$236,250**	**$240,900**	**$255,008**
Fringe Benefits (10% of salaries for health/payroll taxes)	**23,625**	**24,090**	**25,501**
Travel			
One trip each year to Department of Education	850	1,000	850
Visits to Wexford Inc. and multiple school sites Year 1: meetings with Wexford, Inc. to establish protocol and to meet with teachers/school administrators for orientation and set-up Year 3: wrap-up with teachers/school administrators Note: the remainder of travel to school sites is included in the Wexford, Inc. costs	5,150	0	3,000
Total Travel	**6,000**	**1,000**	**$3,850**
Supplies (for school sites: includes CDs, printed instructions, interview and survey materials and Internet costs)	**3,000**	**0**	**0**
Other			
Subcontract to Wexford, Inc. for protocol development, evaluation and testing	**65,000**	**70,000**	**40,000**

Total Direct Costs	**$333,875**	**$335,990**	**$324,359**
Indirect Costs (61.79% of MTDC, less subcontracts)	**166,125**	**164,010**	**175,641**
Total Direct and Indirect	**$500,000**	**$500,000**	**$500,000**

h. Appendix A

Preliminary Evidence of Effectiveness of Intelligent Tutoring Technology
The following summarizes the results from an empirical study of the impact of the Quantum intelligent tutoring system for the chemistry topic of balancing chemical equations [19]. Further empirical results in another study on the effectiveness of the intelligent tutoring technology are available in Ref. [17].

Fifty-one students in Chemistry I in the Spring 2003 semester at Carmichaels Area High School located in rural Greene County, PA participated in the study. The study involved two instructional conditions: a treatment group that received regular classroom instruction and used the AI tutor as part of their at-home study activities, and a control group that received the same classroom instruction but did not use the tutor. Separate classes of the course were used as treatment (N = 30) and control (N = 21) groups. The same instructor taught both sections. A pre-/post-test design was carried out with each group. Results are summarized in Table A1. Mean scores as well as the percentage of students arriving at a *correct* solution are given for each group. All scores are normalized to a maximum of 100 points.

Table A1. Student Performance on Balancing Equations Pre- and Post-Tests

	Mean Score		Percentage Correct	
Equation	Treatment	Control	Treatment	Control
Pre-test:				
(1) $MnO_2 + CO \rightarrow Mn_2O_3 + CO_2$	77.9	79.8	66.7	66.7
(2) $C_2H_8N_2 + N_2O_4 \rightarrow N_2 + CO_2 + H_2O$	59.2	54.7	30.0	28.6
(3) $C_3H_5(NO_3)_3 \rightarrow CO_2 + H_2O + N_2 + O_2$	35.3	33.3	6.7	14.3
(4) $C_8H_{14} + O_2 \rightarrow CO_2 + H_2O$	43.2	46.5	6.7	19.0
Total	53.9	53.6		
Post-test:				
(5) $N_2 + O_2 \rightarrow N_2O_5$	83.2	76.7	80.0	66.7
(6) $Na + NaNO_3 \rightarrow Na_2O + N_2$	63.4	50.3	53.3	28.6
(7) $C_6H_{10} + O_2 \rightarrow CO_2 + H_2O$	76.2	57.6	66.7	33.3
Total	74.3	61.5		

There is a strong trend of improvement of treatment relative to control in mean scores. The groups have essentially identical mean total scores on the pre-test and less than five points difference on any given equation. Both groups showed gains between the pre- and post-tests (as expected), but the treatment group's mean score is 12.8 points higher than the control group on the post-test, *i.e.* more than a letter grade. Treatment outperforms control on all three post-test equations, with the difference between treatment and control greater for the more difficult equations, nearing twenty points on Equation (7).

Examining the percentages of *correct* solutions also shows a marked difference between treatment and control. The groups were again equivalent on the pre-test, with the control group actually performing slightly better than the treatment group Equations (3) and (4), but with both groups very low. On the post-test there are marked differences, however. Treatment again outperforms control on all three post-test equations, with the gap widening as the equations become more difficult to balance. For Equations (6) and (7), the treatment group students reached a correct solution twice as frequently as the control group.

Statistical significance of the improvement was assessed using a χ^2 (chi-square) analysis comparing the number of correctly solved problems in the treatment group to the control group. On the pre-test there was no significant difference in the performance of the two groups. However, on the post-test there was a statistically significant difference between the frequency of correct solutions in the treatment group expected on the basis of the control group and the actual frequency of correct solutions observed, $\chi^2(1, N = 30) = 11.9, p < 0.001$.

Note: Based on their successful experience with the Quantum Tutors, Carmichaels Area High School is continuing their participation in evaluation activities for the current project (see letter of commitment below).

Preliminary Evidence of Effectiveness of Assessment Technology (AI Automated Grading)
The following summarizes the results from Quantum's Phase I SBIR assessment technology project with NSF (subsequently awarded Phase II) [23]. Further empirical results in another study on the effectiveness of the AI assessment technology are available in Ref. [18].

This project showed feasibility by creating an automated AI grading program for student work on the chemistry topic of oxidation numbers and comparing the results of the AI program to the grading of actual teachers on a dataset of authentic student work. A large group of students took a test consisting of five problems of varying difficulty, in which they were instructed to show as much work as possible to obtain partial credit. Fifty student papers (for a total of 250 problem attempts) were randomly selected and identical copies were given to four experienced high school chemistry teachers to grade.

Table A2 shows the mean scores assigned by the four teachers and the AI grader for the five-problem validation test. All scores are normalized to a maximum of 100 points.

Table A2. Mean Oxidation Number Test Scores Assigned

	Teacher 1	Teacher 2	Teacher 3	Teacher 4	AI Grader
Entire Test	57.2	50.8	62.9	53.2	59.4
Problem 1	50.3	44.7	53.8	43.0	49.6
Problem 2	76.8	72.0	85.0	69.2	82.8
Problem 3	64.0	53.3	64.3	58.7	64.4
Problem 4	35.2	32.0	47.7	43.3	45.5
Problem 5	59.5	52.0	63.5	51.8	54.4

While mean scores alone do not ascertain the quality of the grades, the mean AI scores are entirely in line with those of the four teachers; no significant differences are observed. There are no systematic trends of under-awarding or over-awarding of credit with respect to the grading of the four teachers in the study. The AI program tends to be nearer the upper end of scores awarded, but lies firmly within the group of graders.

The more important question is whether the AI grader is *consistent* with the teachers. Inter-rater agreement is one of the most significant issues on the quality of the AI grading program. Pearson's correlation coefficient (r) was calculated for all pairs of graders. These results are presented in Table A3.

Table A3. Inter-Rater Correlation Coefficients for Oxidation Number Test Scores

	Teacher 1	Teacher 2	Teacher 3	Teacher 4	AI Grader
Teacher 1	--	0.853	0.864	0.473	0.864
Teacher 2		--	0.836	0.432	0.850
Teacher 3			--	0.549	0.916
Teacher 4				--	0.469
AI Grader					--

The correlation coefficients clearly establish that the AI grader performs just as well as the teachers on this real-world grading task. There is good agreement among Teachers 1 – 3 and the AI grader, with all correlation coefficients greater than 0.8. Not only does the AI grader correlate strongly with the grading of Teachers 1 – 3, these three teachers also agree with each other. Teacher 4 is an outlier, with whom neither Teachers 1 – 3 nor the AI grader agree closely; here the discrepancy was traced primarily to an error made by this teacher in working a problem when preparing a grading key. Though this was not anticipated, it vividly illustrates the potential impact of human error on the grading process; a major advantage of the AI grader is that such errors and inconsistencies are totally eliminated.

The bottom line is that the variation between the AI grader and the teachers is no greater than the differences among the teachers themselves, and by this practical measure the AI grader does the job just as well as the teachers. It is interesting to note that if the labels were omitted from Tables A2 and A3, it would be impossible pick out the AI grader from the group. This work clearly established the feasibility of a new AI-based approach to development of assessment technology for student work in chemistry. On this initial application, the performance of the AI grader in terms of accuracy, robustness and reliability was shown to be as good as or better than human graders, and clearly met the level required for providing substantive assistance.

Letter of Commitment from Wexford, Inc. (external evaluator)

Wexford

Wexford is a non-profit educational agency
28633 So. Western Ave., Suite 202
Rancho Palos Verdes, CA 90275
Voice: 310-548-0077
Fax: 310-548-0099

January 5, 2004

Dr. Benny G. Johnson
President
Quantum Simulations, Inc.
5275 Sardis Road
Murrysville, PA 15668

Dear Dr. Johnson,

This letter serves to confirm the commitment and availability of Wexford, Inc. to participate as the evaluators in your company's Mathematics and Science Education Research project proposed to the U. S. Department of Education. Ms. Cassidy and I are willing to commit our efforts to the successful completion of the project during the proposed period of performance of June 1, 2004 through May 31, 2007. If there is any deviation in these dates, we are able to commit to a revised period of performance through 2007. The effort required is determined to be a total of $175,000 over the three years of the project. As described in your research plan, Wexford will also arrange participation of additional schools in the evaluation during project years 2 and 3.

Sincerely,

Wexford, Inc.

Selma Sax
Sheila Cassidy

Letter of Commitment from Carmichaels Area High School (Greene Co., PA)

Carmichaels Jr. Sr. High School

Carmichaels Area School District

Principal - Mrs. Lyn N. Shlosky
Email- lshlosky@carmarea.org
Assistant Principal - Mr. John V. Menhart
Email- jmenhar@carmarea.org

300 West Greene Street
Carmichaels, PA 15320
724-966-5045
FAX 724-966-5556

19 December 2003

Dr. Benny G. Johnson
President/CEO
Quantum Simulations, Inc.
5275 Sardis Road
Murrysville, PA 15668

Dear Dr. Johnson:

This letter serves to express my enthusiastic support for and confirm my participation in your research project on artificial intelligence software for assessment in chemistry proposed to the U.S. Department of Education. I understand the nature of the commitment of time, space and resources to the research project that will be required if your application is funded, and am ready, willing and able to participate in testing and evaluation of the software during this three-year project as described in your research plan. I welcome this unique opportunity to be a part of this important work and look forward to my contributions to its successful outcome.

Sincerely,

Kevin Willis
Carmichaels Area High School

Letter of Commitment from Russell Co., KY schools

WINDOWS OF LEARNING
RUSSELL
COUNTY
SCHOOLS

Russell County Board of Education

P.O. Box 440
404 South Main Street
Jamestown, KY 42629

(270) 343-3191
Fax (270) 343-3072

Scott Pierce, Superintendent

December 18, 2003

Dr. Benny G. Johnson
President/CEO
Quantum Simulations, Inc.
5275 Sardis Road
Murrysville, PA 15668

Dear Dr. Johnson:

This letter serves to express our enthusiastic support for and confirm our participation in your research project on artificial intelligence software for assessment in chemistry proposed to the U.S. Department of Education. We understand the nature of the commitment of time, space and resources to the research project that will be required if your application is funded, and our teachers are ready, willing and able to participate in testing and evaluation of the software during this three-year project as described in your research plan. We welcome this unique opportunity to be a part of this important work and look forward to our contributions to its successful outcome.

Sincerely,

Scott Pierce
Superintendent

SP:js

"Better Education Makes Better Citizens"
An Equal Opportunity Employer M/F/D

Letter of Support from Carnegie Learning

LEARNING BY DOING®

CARNEGIElearning™

January 6, 2004
Benny Johnson
Quantum Simulations
5275 Sardis Road
Murrysville, PA 15668

Dear Dr. Johnson,

Carnegie Learning has been following the development of the Quantum Simluations technology and approach for some time, and we are very impressed with the progress you have made. We see your recent work on developing new approaches to assessment in chemistry to be complimentary to our own work in mathematics and hope that work of this sort will help to bring effective educational technology to a broader market.

As the leading publisher of research-based intelligent tutoring systems for mathematics, we support your proposal to develop software tools which enhance a teacher's ability to assess student understanding and improve instruction.

Should this work prove successful, we look forward to discussing with you and your colleagues the opportunity to expand the reach of this new kind of learning environment.

Sincerely,

Steven Ritter
Senior Cognitive Scientist

1200 penn avenue
suite 150
pittsburgh pennsylvania 15222-4210
phone 412/690/6284
facsimile 412/690/2444

i. Appendix B

Example of Quantum Intelligent Tutoring Technology
Figures B1 and B2 below are sample screens from the Quantum Intelligent Tutor for Balancing Chemical Equations. These illustrate a student entering her own problem (a difficult equation to balance), then entering her own attempt at solving the problem and getting assistance step by step in detail. Students access the tutor simply by logging into an Internet web site, and use the tutor entirely within the web browser (with nothing to download or install on the local machine). As per proposal preparation instructions, the screens below are shown in black and white; the actual web pages are in color.

Figure B1. Student Entering a Problem into Equation Balancing Tutor

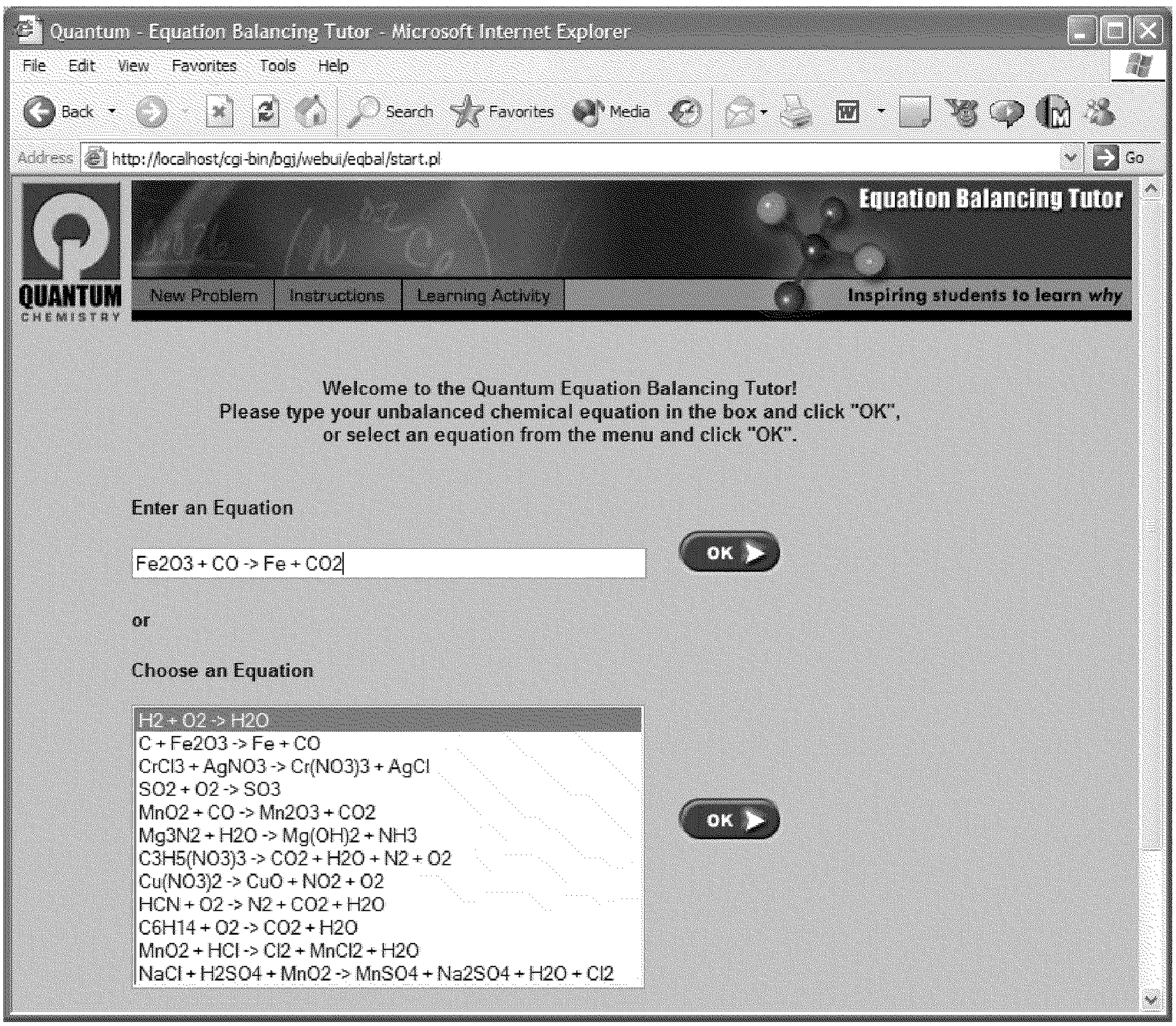

Figure B2. Student Getting Feedback on an Incorrect Step in Working the Problem

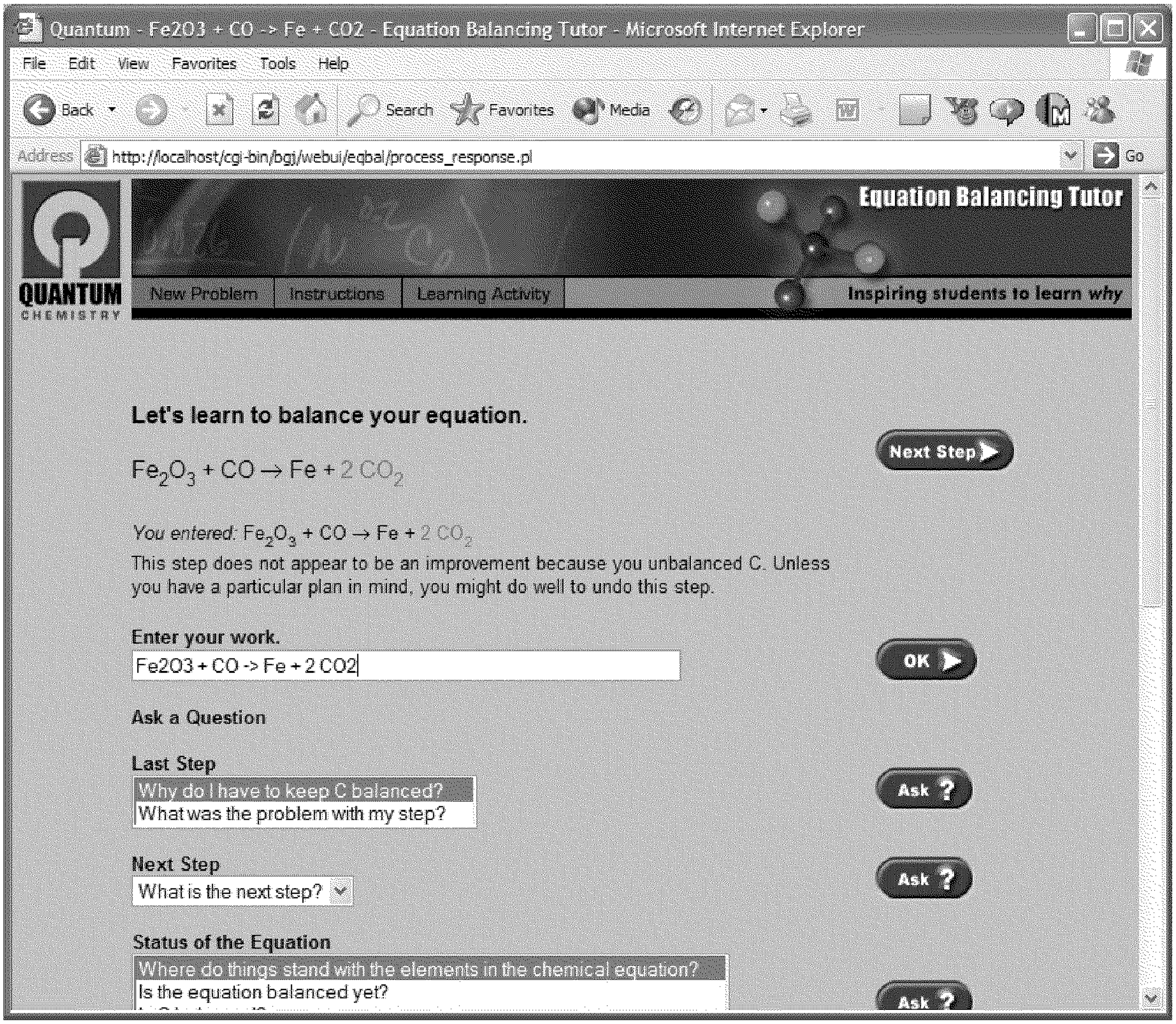

Figure B3 is the transcript from this example session. The dialogue between student and tutor illustrates three of the main advantages of the Quantum AI technology: students can get tutoring on their own problems rather than having to select from a fixed list of problems, can enter their own work and get feedback on their own mistakes, and can ask the tutor questions about the process and concepts involved at every step. In the transcript below, changes made to the equation by the student and the tutor are highlighted, and the underlined terms are hyperlinks to glossary entries in the original tutor web page; the student can get definitions and more information by clicking on them.

Figure B3. Transcript of Student's Session with Equation Balancing Tutor

Printer-Friendly Transcript – Equation Balancing Tutor

Tutor: Welcome to the Quantum Equation Balancing Tutor! Please type your unbalanced chemical equation in the box and click "OK", or select an equation from the menu and click "OK".

Student: $Fe_2O_3 + CO \rightarrow Fe + CO_2$

Tutor: Let's learn to balance your equation. You may enter a step or ask me a question.
$Fe_2O_3 + CO \rightarrow Fe + CO_2$

Student: $Fe_2O_3 + CO \rightarrow Fe + 2\ CO_2$
Tutor: This step does not appear to be an improvement because you unbalanced C. Unless you have a particular plan in mind, it might be best to undo this step.

Student: Why do I have to keep C balanced?
Tutor: Since carbon is the only element appearing in only two formulas (that is not a free element) it is the most important element to keep balanced. Then, the important thing is to balance the next element while still keeping elements you've balanced before balanced.

Student: $Fe_2O_3 + CO \rightarrow Fe + CO_2$
Tutor: Good. It looks like you are now back to where you were before. Use the suggestions in the balancing procedure as you make your next move.

Student: What is the next thing I should do?
Tutor: Of the currently unbalanced elements, O and Fe, the balancing procedure suggests that you should balance O. Be sure that when you balance O, you keep C balanced.

Student: Show me the next step.
Tutor: $1/3\ Fe_2O_3 + CO \rightarrow Fe + CO_2$
I balanced oxygen by putting 1/3 iron(III) oxide formula units on the reactants side.

Student: Why was using Fe_2O_3 to balance oxygen the right thing to do?
Tutor: Since I needed to keep C balanced no matter what, I had to use a formula that contained O but did not contain C. That is why I chose Fe_2O_3.

Student: How did you get 1/3 for the coefficient of Fe_2O_3?
Tutor: Since we had 2 O on the products side, we needed to get 2 O on the reactants side too. On the reactants side, we already had 1 O from CO, so we needed to change the number of Fe_2O_3 to give us 1 O. To do this I multiplied Fe_2O_3 by the number of oxygens I needed, which was 1, and then divided that by the number of oxygens in Fe_2O_3 to get 1/3.

Student: $1\ Fe_2O_3 + 3\ CO \rightarrow 3\ Fe + 3\ CO_2$
Tutor: This is a good idea. Removing any fractions you had in the equation should make it easier for you to balance the equation.

Student: Is the equation balanced yet?
Tutor: No. C and O are currently balanced, but Fe is not.

Student: 1 Fe_2O_3 + 3 CO → 2 Fe + 3 CO_2
Tutor: It looks like the equation is now balanced, great job! There is one more thing that needs to be done before we are through though. You need to remove all of the "1s" that are in the equation.

Student: Fe_2O_3 + 3 CO → 2 Fe + 3 CO_2
Tutor: Good move. The equation now uses proper notation and is balanced. The problem is now completed.

Student: Why did I need to remove the "1s" from the equation?
Tutor: In order for an equation to be considered "in standard form," there must not be any "1s" written out as coefficients. It's just like in algebra; you don't write "1 x", you simply write "x".

Student: What does this balanced equation mean?
Tutor: It tells you that 1 mole of Fe_2O_3 plus 3 moles of CO will react to form 2 moles of Fe plus 3 moles of CO_2.

Student: How many moles of Fe can we make from 2 moles of Fe_2O_3?
Tutor: From the balanced equation we see that there is 1 Fe_2O_3 for 2 Fe, so with 2 moles of Fe_2O_3 we can make 2 x (2 / 1) = 4 moles of Fe.

(End of session.)

Institute of Education Sciences - Grant Application Submissions for 2004

View Application Submit Review

Application Abstract

APPLICATION R305K04008

Project Director Dr. Benny Johnson

Applicant Organization Quantum Simulations, Inc.

Application Title Integrated Software for Artificial Intelligence Tutoring and Assessment in Science

Collaborating Organizations

NAME	CONTACT
Carmichaels Jr. Sr. High School	Kevin Willis, Test Site
Russell County Board of Education	Scott Pierce, Test Site
Wexford, Inc.	Sheila Cassidy, Evaluator
Wexford, Inc.	Selma Sax, Evaluator

Abstract

Integrated Software for Artificial Intelligence Tutoring and Assessment in Science

1. The RFA goal under which the applicant is applying:
Goal One

2. The potential contribution the proposed project will make to the solution of an education problem:
The purpose of this project is to advance the state of the art in science educational software by developing powerful new artificial intelligence (AI) tools for computerized assessment of student work in science, beginning with chemistry, building on and integrating with Quantum's successful interactive intelligent tutoring technology in chemistry. This work addresses a longstanding, clearly articulated need for improved software in chemistry education. With the advent of NCLB, the importance to teachers and students of sophisticated high-quality tools for individualized assessment has never been greater, and the goal of this work is to build advanced assessment capabilities to achieve the same power and sophistication as recently achieved for interactive tutoring.

3. The population(s) from which the participants of the study(ies) will be sampled (age groups, race/ethnicity, SES):
The students involved in evaluation efforts will be high school students

Personal Information

Meeting Information

Form Repository

Travel Arrangements

Peer Review

Change Password

Logout

taking first-year chemistry. Partnerships have been arranged to work with a variety of and schools including at-risk students across a range of rural and urban populations, racial, ethnic and minority groups and low-income backgrounds.

4. The proposed research method(s):
The artificial intelligence development methods proposed have their origins in the ACT theory of Anderson, Corbett, Koedinger, et al., with improvements to AI model-tracing theory developed by Quantum in prior research that extend the range of student errors to which the method is applicable, including unanticipated mistakes. For preliminary formative evaluation Quantum has partnered with Wexford, Inc., an organization expert in scientifically based evaluation research, as an external evaluator.

5. The proposed intervention if one has been proposed:
The intervention proposed is to place a comprehensive set of AI tutoring systems with integrated AI assessment capabilities in the hands of every student in a chemistry class, to use on demand through the Internet during study time at home and at school. Though making a meaningful impact on science achievement is difficult, to succeed it is clear that students must have the opportunity for sustained, quality practice with science concepts during their own study time.

The Quantum intelligent tutoring technology used as the basis for this work provides breakthroughs in key respects that enable it to be effective in one-on-one instructional situations where conventional tutorials typically are not: students can use their own problems, get help on their own mistakes, and ask the tutor questions about the concepts and processes being studied.

The AI technology is primarily intended to be used in a supplementary role to reinforce classroom instruction, by giving students the individual assistance and assessment they need at home to get the most out of the day's lesson. By making integrated AI tutoring and assessment available to students on demand, including at home, the quality and effectiveness of at-home study time will be increased. Preliminary empirical evidence that the intervention will be effective is presented.

Home IES Homepage Help Contact Technical Support

Institute of Education Sciences - Grant Application Submissions for 2004		

View Application Submit Review
Application Detail

APPLICATION R305K04008

Item 1. on ED 424 Organization Name and Address

Name	Quantum Simulations, Inc.
Unit	Quantum Simulations, Inc.
Address	5275 Sardis Road
City	Murrysville
State	PA
Country	United States of America
Zip-Code	15668

Item 2. on ED 424 Project Director Name and Information

First Name	Benny
Middle Initial	G
Last Name	Johnson
Prefix	Dr.
Address	5275 Sardis Road
City	Murrysville
State	PA
Country	United States of America
Zip-Code	15668
Phone	724-733-8603
Fax	724-325-2062
Email	johnson@quantumsimulations.com

Item 3. on ED 424 Authorized Representative Name and Informati

Primary	Yes
First Name	Benny
Middle Initial	G
Last Name	Johnson
Prefix	Dr.
Address	5275 Sardis Road
City	Murrsyville

State	PA
Country	United States of America
Zip-Code	15668
Phone	724-733-8603
Fax	724-325-2062
Email	johnson@quantumsimulations.com

Primary	No
First Name	Rebecca
Middle Initial	-
Last Name	Renshaw
Prefix	Ms.
Address	5275 Sardis Road
City	Murrysville
State	PA
Country	United States of America
Zip-Code	15668
Phone	724-733-8603
Fax	724-325-2062
Email	renshaw@quantumsimulations.com

Item 4. on ED 424 Application Details

4a. D-U-N-S Number	(b)(2) [redacted]
4b. T-I-N Number	25-1808705
4c. CFDA Number	84.305K
4d. Novice	N/A
4e. Federal Debt Delinquency	No
If Yes, Federal Debt Explanation	N/A
4f. Type of Applicant	Private, Profit-Making Organ
If Other, Description:	N/A
4g. Type of Submission	Non-construction Applicatio
4h. Executive Order 12372	N/A
4i. Project Start Date	6/1/2004
Project End Date	5/31/2007
4j. Human Subjects Research	Yes
If yes, Exempt From Regulations?	No
4k. Exemption Number(s)	N/A
4l. Exempt Research Narrative	N/A

4m. Non-Exempt Research Narrative	Research will involve pre- ar interview and survey protoco will be conducted by a third-; also the IRB on this project. ; project director or other mem Simulations, Inc. staff be pre and data collected will be an(way impact student performa evaluation of students.
4n. Human Subjects Assurance Number	FWA00006177
4o. Application Title	Integrated Software for Artifi Tutoring and Assessment in S

Item 5. on ED 424 Estimated Budget

a. Federal Funding
b. Applicant Funding
c. State Funding
d. Local Funding
e. Other Funding
f. Program Income Funding

Personal Information
Meeting Information
Form Repository
Travel Arrangements
Peer Review
Change Password
Logout

Section A on ED 524 Federal Budget

Budget Categories	Project Year 1 (a)	Project Year 2 (b)	Project Year 3 (c)	Project Yea (d)
1. Personnel	$236,250.00	$240,900.00	$255,008.00	$0.
2. Fringe Benefits	$23,625.00	$24,090.00	$25,501.00	$0.
3. Travel	$6,000.00	$1,000.00	$3,850.00	$0.
4. Equipment	$0.00	$0.00	$0.00	$0.
5. Supplies	$3,000.00	$0.00	$0.00	$0.
6. Contractual	$0.00	$0.00	$0.00	$0.
7. Construction	$0.00	$0.00	$0.00	$0.
8. Other	$65,000.00	$70,000.00	$40,000.00	$0.
9. Total Direct Costs	$333,875.00	$335,990.00	$324,359.00	$0.
10. Indirect Costs	$166,125.00	$164,010.00	$175,641.00	$0.
11. Training Stipends	$0.00	$0.00	$0.00	$0.
12. Total Costs	$500,000.00	$500,000.00	$500,000.00	$0.

Section B on ED 524 Non-Federal Budget

Budget Categories	Project Year 1 (a)	Project Year 2 (b)	Project Year 3 (c)	Project (d

1. Personnel	$0.00	$0.00	$0.00
2. Fringe Benefits	$0.00	$0.00	$0.00
3. Travel	$0.00	$0.00	$0.00
4. Equipment	$0.00	$0.00	$0.00
5. Supplies	$0.00	$0.00	$0.00
6. Contractual	$0.00	$0.00	$0.00
7. Construction	$0.00	$0.00	$0.00
8. Other	$0.00	$0.00	$0.00
9. Total Direct Costs	$0.00	$0.00	$0.00
10. Indirect Costs	$0.00	$0.00	$0.00
11. Training Stipends	$0.00	$0.00	$0.00
12. Total Costs	$0.00	$0.00	$0.00

Abstract

Integrated Software for Artificial Intelligence Tutoring and As
Science

1. The RFA goal under which the applicant is applying:
Goal One

2. The potential contribution the proposed project will make tc
of an education problem:
The purpose of this project is to advance the state of the art in
educational software by developing powerful new artificial int
tools for computerized assessment of student work in science,
with chemistry, building on and integrating with Quantum's sı
interactive intelligent tutoring technology in chemistry. This w
a longstanding, clearly articulated need for improved software
education. With the advent of NCLB, the importance to teache
students of sophisticated high-quality tools for individualized
never been greater, and the goal of this work is to build advan
capabilities to achieve the same power and sophistication as re
achieved for interactive tutoring.

3. The population(s) from which the participants of the study(i
sampled (age groups, race/ethnicity, SES):
The students involved in evaluation efforts will be high school
taking first-year chemistry. Partnerships have been arranged tc
variety of and schools including at-risk students across a range
urban populations, racial, ethnic and minority groups and low-
backgrounds.

4. The proposed research method(s):
The artificial intelligence development methods proposed hav
in the ACT theory of Anderson, Corbett, Koedinger, et al., wi
improvements to AI model-tracing theory developed by Quan

research that extend the range of student errors to which the m
applicable, including unanticipated mistakes. For preliminary
evaluation Quantum has partnered with Wexford, Inc., an orga
expert in scientifically based evaluation research, as an extern

5. The proposed intervention if one has been proposed:
The intervention proposed is to place a comprehensive set of A
systems with integrated AI assessment capabilities in the hand
student in a chemistry class, to use on demand through the Int
study time at home and at school. Though making a meaningfu
science achievement is difficult, to succeed it is clear that stud
the opportunity for sustained, quality practice with science con
their own study time.

The Quantum intelligent tutoring technology used as the basis
provides breakthroughs in key respects that enable it to be effe
on-one instructional situations where conventional tutorials ty
students can use their own problems, get help on their own mi
the tutor questions about the concepts and processes being stu

The AI technology is primarily intended to be used in a supple
to reinforce classroom instruction, by giving students the indiv
assistance and assessment they need at home to get the most o
lesson. By making integrated AI tutoring and assessment avail
students on demand, including at home, the quality and effecti
home study time will be increased. Preliminary empirical evid
intervention will be effective is presented.

Collaborating Organizations

NAME	CONT.
Carmichaels Jr. Sr. High School	Kevin Willis, Tes
Russell County Board of Education	Scott Pierce, Test
Wexford, Inc.	Sheila Cassidy, E
Wexford, Inc.	Selma Sax, Evalu

Upload PDF

VIEW CURRENT FILE (394.7 KB)
For best viewing results, please ensure that you have the latest version o