Developing a Software Tool to Help Teachers Write Better Multiple-Choice Questions

A proposal by Morgan C. Benton and Marilyn M. Tremaine

1. Solution to an Educational Problem

Research has shown that students learn best when they follow a three-part process. First, they need to receive and understand what their learning goals are. Second, they need to be able to assess their understanding relative to these learning goals. Finally, they need to be shown or guided in how to achieve their goals following the self-assessment. A teacher's role, therefore, is in setting appropriate goals for students, helping students to perform self-assessment and showing them how to achieve the goals set (Black, Harrison, Lee, Marshall, & Wiliam, 2003). This project proposes a software system to help teachers be more effective in the second stage of this process, self-assessment. The software system will be designed to help teachers write multiple-choice questions that match the educational goals set. The questions will be designed so that they can be used by students for self-assessment. The system will also give teachers feedback on student achievement and underlying reasons for not reaching set goals. This will help an instructor tune teaching to give better guidance for goal achievement.

Research has also shown that teachers are poor at formative assessment. They tend to use questions test a student's recall of facts without examining higher order issues such as whether the student can recognize new examples of a concept or apply the concept or process to an entirely new situation. These types of questions are what we call higher-order questions. They are also the questions that are the most related to learning goals. It is these questions that tease out the misconceptions and wrong paths a student is taking, and through this, provide the teacher and student with feedback on the paths to take to achieve the learning goal. The software system being proposed in this project will help teachers design these higher order questions.

It is much easier to build a software system that supports good question design then it is to get teachers to use the system and change their assessment practices. Thus, although the proposed system supports the development of better assessment tools, there is no indication that it will not simply be used for generating rote memory multiple choice questions. The project therefore proposes a second significant addition to the software system drawing on research that has already been done in getting people to stop smoking (Prochaska & Velicer, 1997). The system will integrate aspects of the transtheoretical model of behavior change (TTM) that will assess a teacher's progress through a change process and provide appropriate interventions given a diagnosis of a person's current stage.

Once the software system is developed, it will be delivered to a test set of science and math teachers in 15 high schools. Teachers will either be given only software that supports electronic test generation, software that supports higher order test generation or software that supports higher order test generation using interventions supplied by the transtheoretical model of behavior change. Comparisons of student learning and of teacher's beliefs and attitudes about self-assessment will be captured in this study and used to measure the efficacy of this approach. Learning gains are expected to follow from an improvement in teachers' assessment practices. Black et al.(2003) report on the results of a meta-analysis of formative assessment interventions and note that effect sizes were some of the largest ever reported for sustained educational interventions. Thus, if the TTM interventions can are shown to change assessment practices, we should see significant improvements in student learning.

Using a software tool for improving formative assessment also has several key productivity advantages for the education community. First the system supports easier exam generation and grading, which can lead to more assessments. Second, use of the system can lead to easier best practices sharing because the generated documents are electronic. Third, use of the system provides an easy route into improving teacher's formative assessment skills because of its non-invasive approach.

2. Research Narrative

Effective learning requires that students understand the learning goal, are able to assess their own understanding relative to that goal, and then have both a means and a motivation to get there. As a partner in learning, the teacher's role is to set appropriate goals, to provide opportunities for students to perform self-assessment, to help students find motivation, and finally to suggest appropriate mechanisms by which students may bridge the gap between current and target understanding(Black et al., 2003). This model suggest that a plan for professional development would help teachers become better goal-setters, better at assessing student ability, better at understanding what motivates students to achieve, and finally better at developing lessons that help students reach their goals. This project aims at the second of these goals: to make teachers better at assessment.

Teachers' assessment competence is a problem(Stiggins, 2001). Research indicates that teachers are not trained thoroughly in assessment methods(Crooks, 1988), are not good at creating tests (Benton, Tremaine, & Scher, 2004; Carter, 1984; Fleming & Chambers, 1983), aren't good at assessing the quality of tests designed by others(Carter, 1984), may not deliver them often enough(Bangert-Drowns, Kulik, & Kulik, 1991), and don't provide appropriate feedback to students based on test results(Brookhart, 1993; Stiggins, Frisbie, & Griswold, 1989; Whitmer, 1983). Specifically with reference to the questions that teachers ask, either on a test or verbally, studies find an overemphasis on questions that assess only rote learning or memory, and do not tap higher-orders of thinking(Black & Wiliam, 1998b). Yet higher-order thinking is exactly what we wish to promote(National Council of Teachers of Mathematics, 2000; National Research Council, 1996).

The rest of this proposal builds upon the above stance on learning and teachers' assessment ability. The following section will lay the theoretical foundation explaining the mechanisms and evidence associated with formative assessment practice, summarizing the pros and cons of three professional development paradigms, and finally integrating the best of those paradigms using the transtheoretical method. Having established a solid theoretical foundation, the next section will describe research model, main research questions, and hypotheses for this intervention. Following that, it will be shown how the research model suggests the software architecture for the proposed tool. The tool will be described in detail. Finally, a plan for how this project will be carried out, including plans for a field trial and evaluation of the tool, data analysis and statistical power, ethical considerations, deliverables, project personnel, and institutional resources will be explained. The proposal concludes with a discussion of how success in this project will lead into a proposal for a larger trial of efficacy and replicability under Goal 3.

2.1 Theoretical Foundation

Two bodies of theory, one on formative assessment, another on professional development, form the foundation for this project. Each will be described thoroughly below. New to the discussion of teacher professional development is the incorporation of the transtheoretical model.

2.1.1 Formative Assessment

Definitions

There is not a generally agreed upon definition for the term *formative assessment*. It is useful to set formative assessment apart from its complement *summative assessment*, which Bloom, Hastings, and Madaus (1971) defined as those tests given at the end of episodes of teaching (units, courses, etc.) for the purpose of grading or certifying students, or for evaluating the effectiveness of a curriculum (p117). Summative assessments generally result in letter or number grades, percentiles or other measures which are often criticized for their subjectivity or lack of relevance to actual learning.

Formative assessment, on the other hand, makes student learning the central goal. It entails teachers asking well-designed questions that make students' understanding apparent, and it requires that teachers attend to student answers and give students feedback which students then use to correct their misconceptions. However, formative assessment is not necessarily defined by specific behaviors. Black et al. (2003) say that

> *It is misleading to say that formative assessment requires a teacher to act in a closely defined way. However, it does demand that the motivation that makes a teacher act in a given way is always to involve the students in the learning and to allow the students to be aware of their learning successes.* (p99)

Hence these authors make the assertion that there is no recipe for the development of formative assessment practice with teachers. Despite this assertion, their reviews of the literature on classroom assessment (Black & Wiliam, 1998a, 1998b) indicate that four areas of practice provide substantial opportunity for improving learning:

- Encouraging teacher use of higher-order questions
- Improving the quality of feedback to students on their performance
- Using peer and self-assessment to encourage students to monitor their own performance
- Formative use of summative tests

The first three of these areas of practice will be discussed now.

Practices of Formative Assessment

When teachers ask higher-order questions, students learn more. By *higher-order* is meant questions that ask students to do more than recall information from memory, i.e. such items as facts, figures, and definitions(Anderson & Krathwohl, 2001). In one recent study it was observed that students rarely learned unless they reached an impasse—a situation when the student gets stuck in answering a question, detects an error, or answers correctly but expresses uncertainty about why the answer is correct(VanLehn, Siler, Murray, Yamauchi, & Baggett, 2003). Effective teaching might then be associated with asking students questions that bring them to an impasse—an impasse that can only be broken through learning. Furthermore, in order to get students to do higher-order learning in line with national standards in math and science (National Council of Teachers of Mathematics, 2000; National Research Council, 1996) it follows that breaking the impasse must require students to apply concepts, evaluate options, think deeply.

Unfortunately, a majority of the questions that teachers ask do not tap higher-order thinking (Stiggins et al., 1989), and their methods of asking questions in the classroom do not allow students sufficient time to formulate a reasoned response(Rowe, 1974). A majority of the questions that teachers ask in classrooms assess nothing other than how well students read and remember facts in a textbook, or the previous day's lesson. The questions are not designed to elicit the depth of students' knowledge or understanding on the topics being covered. Furthermore, teachers tend to solicit an immediate response from a select set of students within any given class who typically respond to questions. In contrast, using a "hands down" strategy which forces all students to participate, and increasing the "wait time" from asking a question to soliciting an answer to as little as five seconds can produce remarkable improvement in student engagement and responsiveness.

Formative assessment supports using well-designed questions not only to encourage deeper thinking, but also to diagnose student misconceptions. A recent study showed that even in one-on-one tutoring sessions, instructors are not very good at monitoring the subjective understanding of their students(Chi, Siler, & Jeong, 2004), although this did not seem to hamper learning in such individualized settings(Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001). Of course the major problem is logistic—how does a classroom teacher, who sees upwards of 100 students per day, monitor and respond individually to all of them? The DIAGNOSER project built a system to allow high school physics teachers to do just that(Thissen-Roe, Hunt, & Minstrell, 2004), built upon the concept of facet-based instruction (Minstrell, 1992). *Facets* are conceptions of the world, or procedures for problem-solving held by students. In any given domain some facets are "correct" interpretations of the world, and others are less so. The role of a teacher in facet-based instruction is to determine which facets are held by students and convince them to replace these with the target facet, or learning goal.

A major bottleneck in the development of systems such as the DIAGNOSER or Cognitive Tutor™ (Blessing, 2003; Koedinger, 1998) is that they are specialized and depend on a detailed cognitive task analysis of the domain being taught. While for some subjects that have received extensive study such as Algebra I, or Newtonian motion, empirical analyses of students' thinking are available, but there are still a large number of subject areas that have not received such attention. It will be some time before such systems based on deep, empirically confirmed understanding of student cognition will be able to be developed, if ever. We propose an alternative system that encourages teachers to develop such deep understanding on their own. Encouraging teachers to develop such knowledge is a part of what Shulman(1986) referred to as pedagogical content knowledge. Recent depictions of the teacher as a learner have emphasized building up such knowledge(Iszák & Sherin, 2003; Sherin, 2002).

A deep understanding of students' thinking gained from asking higher-order questions will allow teachers to develop in the second practice of formative assessment—providing useful feedback. The quality of feedback has been shown to have one of the largest impacts on students' learning(Bangert-Drowns, Kulik, & Morgan, 1991). A problem is that if teachers provide students with a number or letter grade, students will ignore any other written comments on their work, although it is precisely these comments which would help them the most. This has led to the successful implementation of "comment-only" marking as a formative practice(Black et al., 2003). Effective feedback does three things: points out to the student what he or she did well and what things need work, reminds the student of the learning goal, and gives actionable advice on how to bridge the gap. The obvious problem is that providing such high-quality feedback is extremely time-consuming. Software systems seem a likely solution.

Systems such as the DIAGNOSER facilitate the delivery of individualized feedback based on test performance(Thissen-Roe et al., 2004). Commercial systems, e.g. WebCT™, support tying feedback to specific responses, but do not offer guidance as to what type of feedback teachers should provide. The Cognitive Tutor™ provides feedback in real time while students are in the act of problem solving to help them recover from errors or overcome roadblocks in their understanding. Expert systems have been used successfully in healthcare settings to provide tailored instructions and feedback to people for taking medication and changing unhealthy behaviors such as smoking(Hirst, DiMarco, Hovy, & Parsons, 1997; Velicer & Prochaska, 1999). These expert systems allow healthcare professionals to anticipate the kinds of issues patients will face and prepare their responses ahead of time. An interesting feature of the expert healthcare systems is that the output is printed, which means that patients do not need to interact with a computer directly. This is important for educational settings where there is not a computer for every student in every class and teachers must give feedback orally or on paper.

Self-assessment is the third formative practice to be discussed. Feedback is not complete unless students take advantage of it to improve their learning. Toward this goal it is important that students be taught to assess their own progress. Self-assessment and peer assessment, if used appropriately, encourage students to think deeply about success criteria and then review their own performances as measured against these criteria. In addition, if students are involved in their own performance assessment, it can reduce the amount of time that teachers need to spend marking assessments(Black & Wiliam, 1998b). This will free teachers to spend time providing detailed feedback on a smaller number of key assignments. In general, self-assessment practices fall into the category of meta-cognitive knowledge(Anderson & Krathwohl, 2001), and aid students in become self-sufficient, lifelong learners.

A problem with these three practices—higher-order questions, high quality feedback, and teaching self-assessment—is that, a broad range of practices that fit into those categories have been shown to be successful making it hard to measure the effect of adopting formative assessment techniques. Judgments must be made as to whether or not a given behavior represents a formative approach to assessment, and indeed, the same practice can be either summative or formative depending on the context. Luckily, however, there is a relatively stable set of attitudes and beliefs associated with the adoption of formative practice. It is likely that measuring these can provide insight into teachers' adoption of formative practice.

Attitudes and Beliefs Associated with Formative Assessment Practice
While the specific practices that teachers take on when they adopt formative assessment will differ substantially from teacher to teacher, their attitudes and beliefs change in a rather uniform way. Black et al. (2003) summarized that:

> *Specifically what changed for teachers was*
> - *their views about learning,*
> - *their professional priorities,*
> - *their expectations of their students, and*
> - *their feelings about control in their classrooms.*
>
> *(p91, bullets added for emphasis)*

It may that more than anything else, formative assessment is an outward expression of a core belief system about students, how they learn, and the purpose of a school education. An understanding of these beliefs gives us a handle on how to measure change toward adoption.

It is generally agreed that a teacher's number one priority should be the learning of his or her students, yet teachers' behavior often indicates that this is not the case. The literature on teacher stress offers many possible explanations as to what sorts of pressures may lead teachers to put priorities other than learning ahead(Kyriacou, 2001). An oft-cited example is the pressure put on teachers to have their students perform well on standardized tests. Such pressure leads them to drill students on question types that are representative of those on the test, but which may not be the best way to foster understanding. Additional time may be spent on things like guessing strategies, and other techniques designed to improve scores without affecting learning. Regardless of these pressures, it was found that teachers who successfully adopted formative assessment practices had a renewed and strengthened belief that student learning should be the number one priority in the classroom(Black et al., 2003). These teachers were able to put aside their fears of not "covering the material," able to avoid the temptation to "teach to the test," and were able to focus on students' understanding.

In conjunction with their renewed focus on learning, teachers' expectations of their students changed. The connection between teacher expectations and student performance is well-established(Brophy, 1983). Whereas at the beginning of the study the teachers tended to see student ability as relatively fixed, teachers began to believe that student ability was fluid. Teachers who had begun to assume all of the responsibility for their students' learning, began to shift that responsibility back to the students. Armed with new formative teaching practices, the students were ready to take the responsibility for their own learning.

The third thing that changes is teachers' feelings about control in the classroom. When a teacher adopts formative assessment

2.1.2 Professional Development
A review of literature identifies at least three major approaches to teacher professional development. The first, referred to here as the "best practices" approach, seeks to find and develop model methods and curricula, capture them in a distributable form which is then introduced to teachers, generally in a top-down fashion. The second approach, the inquiry-based approach is more bottom-up and begins with teachers in classrooms engaging in intense reflection about their practice and gradually expanding that discussion to encompass other

teachers in their departments, and then schools, and then school systems. The third approach is the standards-based approach which seeks first to identify universal standards for educational content and then work to implement broad reforms that bring teaching, learning, and assessment in line with these standards. All of these methods offer benefits and weaknesses. "Best practices" often work, but when such practices are mandated or thrust upon a teacher, the result is often resistance. In reaction to these feelings, the inquiry approach lets teachers find their own solutions, which again, often work and about which teachers are often enthusiastic. The problem with this approach is that those managing the school system worry, often rightly, about a lack of consistency across the grades. Standards based systems seek to strike a balance between the two previous approaches because they don't mandate specific practices quite so forcefully as the "best practices" approach, and they also ensure consistency because everyone must conform to the standard, achievement of which is assessed with standardized tests. But finally there is a problem with this approach as well, because the standardized tests are often accused of not being an accurate reflection of what students know, and they also encourage teachers to "teach to the test" in a way that fosters rote memorization and hurts thinking. To integrate the best of all of these while avoiding the negatives, the transtheoretical model is introduced here.

Behavioral Change and the Transtheoretical Model

The transtheoretical model claims to be the only model of behavioral change that incorporates a temporal dimension. Whereas other models identify behavioral change as a more or less instantaneous event, i.e. the moment someone quits smoking or starts exercising, only TTM identifies change as a process that happens over time. It identifies six stages (*precontemplation, contemplation, preparation, action, maintenance, termination*), and ten processes (*consciousness raising, dramatic relief, self-reevaluation, environmental reevaluation, self-liberation, social liberation, counterconditioning, stimulus control, contingency management, helping relationships*) associated with lasting change. The processes have been found to be appropriate during different stages. Three concrete measures for identifying in what stage a person resides have been developed which are *decisional balance*, *self-efficacy*, and *situational temptation*. Based upon these measures it is possible to design individualized interventions that have shown to be very effective at achieving favorable measures of variables like *recruitment*, *retention*, *progress*, *process*, and *outcomes*.

Briefly explained precontemplation is the period of time when a person has no plans to change a "problem" behavior. Contemplation is the period when a person intends to change with in six months (self-report). Preparation is within a month of change. Action is when the person take concrete actions, determined by experts in the type of target behavior, that constitute a real shift in behavioral patters (for smokers this means 100% abstention). Maintenance is marked by a reduction in effort to continue action, higher levels of self-efficacy and lower levels of temptation are experienced.

What TTM means for professional development is that teachers will be in different stages with respect to their readiness to change to a target "best practice." As in the inquiry model, they won't change until they are ready to, but the way of getting them there doesn't have to be all internal. Standards can be used effectively in the classroom, and finally unlike standardized tests which are purely summative, formative assessment in the classroom means that students make meaningful progress towards the standards. Whereas in the past it took a psychotherapist or clinician to diagnose a person's stage of change and prescribe an appropriate intervention, in our case we will do it with a software tool. The tool is described in detail in Appendix B.

Stage of change is normally determined by asking people directly whether or not they practice the problem behavior, i.e. Do you currently smoke? Or Do you ever think of quitting smoking? With non-formative assessment, the stage of change metric can't be so simple since we're not focusing on a single behavior, e.g. giving only number grades. What has been found is that knowledge, beliefs and attitudes are very accurate at predicting behavior, and as discussed above, we know very well what the knowledge, beliefs, and attitudes of teachers are that have embraced formative assessment. Therefore the stage of change metrics for this project will be based upon determining people's stage of change based upon their knowledge, attitudes and beliefs.

Decisional balance means people's awareness of the pros and cons of changing. For example, people know that smoking is bad for them, but they also know that when they try to quit they get the jitters and it makes them feel terrible.

2.2 Research Model

Figure 1 (Appendix A) depicts the research model that will guide the software implementation and evaluation. This is not a validated model, nor is it within the scope of this project to take on such a task. The focus of this project is to develop a tool to raise the quality of multiple-choice questions and the likelihood of their formative use in math and science. Further implementation of support systems for other types of assessment items or further work to validate the research model will be indicated based upon the success of this project.

At the core of the research model knowledge, beliefs, and attitudes with respect to formative assessment are believed to be the primary predictors of the implementation of formative assessment practice. The adoption of formative assessment practice will be associated with learning gains. In the dashed boxes are the processes of change, taken from TTM, that support the stages of change. The temporal dimension of this model is depicted in the stages of change stretching across the bottom of the figure.

Research Questions

Before giving the research questions it is important to note that multiple-choice questions do not appear in the model explicitly. This is because the model is agnostic with respect to the actual types of assessment items that are implemented within formative assessment practice. In other words, it's not important whether or not a teacher chooses multiple-choice questions or portfolio-based assignments to perform assessment as long as the teacher clearly states the learning goals, helps students assess their own understanding in relation to the goal, and provides them with a means to bridge the gap. In this project, the selection of a focus on multiple-choice questions has other motivations that will be discussed below in the description of the actual software tool. That being said, the first research question is as follows:

> *RQ1: Does the use of the software tool improve the quality of teachers' multiple-choice questions?*

Following on from the discussion of question quality, the primary concern of this research project is whether or not it is possible to use a tool to improve teachers' ability to write high-quality questions. Again 'high-quality' refers to questions that tap higher-order thinking, and in addition, are able to diagnose the misconceptions held by students. A complementary skill that also may be affected is teachers' ability to evaluate questions authored by others as being either low or high quality. This research question leads to the following two hypotheses:

> *H1:* *Teachers who use the tool will generate more questions that assess higher-order thinking, and fewer which assess only recall.*
>
> *H2:* *Teachers who use the tool will be better able to discriminate between questions that do or do not assess higher-order thinking.*

In addition to improving the quality of questions, an explicit goal of the software tool is to encourage the formative use of the questions in the classroom. This leads to the second research question:

> *RQ2:* *Does the use of the software tool encourage the use of formative assessment practice in the classroom?*

Since 'formative assessment' is not associated with an explicit set of behaviors, it will be necessary to evaluate the practices of teachers who use and don't use the tool and make some judgments as to whether these practices are 'formative' or not. In this case, it is expected to see teachers doing things such as giving at least one short multiple-choice quiz per week, and providing individualized descriptive feedback as opposed to purely numeric feedback in response to students' performance on the questions. This leads to the following hypothesis:

> *H3:* *Teachers who use the tool will adopt more behaviors that can be classified as formative with respect to the use of the questions developed with the tool.*

Thirdly, it is of interest whether or not the TTM-based architecture has any impact upon the adoption and use of the tool. In the field trial the plan is to distribute two versions of the software tool—one that implements features indicated by TTM, and one that only contains the core multiple-choice writing interface. It is expected that the TTM-enhanced tool will foster less attrition, quicker change, and other behaviors such as sharing the tool with colleagues. This leads to the following research question and hypotheses:

> *RQ3:* *Does implementation of TTM-based functionality make a software tool for professional development more successful?*
>
> *H4:* *Teachers who receive the TTM-enhanced tool will be less likely to discontinue use of the tool.*
>
> *H5:* *Teachers who receive the TTM-enhanced tool will be more likely to change their behaviors.*
>
> *H6:* *Teachers using the TTM-enhanced tool are more likely to spread the use of the tool to other teachers.*

Finally, the primary motivation for the implementation of this system is ultimately to improve learning, hence the final hypotheses:

H7: *Teachers' use of the tool will be positively correlated with learning gains of those teachers' students.*
H8: *Students of teachers who use the TTM-enhanced tool will have higher learning gains than those who use only the basic tool.*

The next section will describe the software tool in detail. The section after that will explain the detailed schedule for the tool's implementation and evaluation, including how the above research questions and hypotheses will be measured.

2.3 Description of the Tool

A detailed description of the tool is in Appendix B.

2.4 How the Project will be Carried Out

A timeline for the execution of this project can be seen in Figure X in Appendix A. There are two phases to the project: a software development phase, and a field trial evaluation phase. The software development phase is divided into three parallel development processes and will end after the first year of the project. The second phase begins at the same time as the software development phase and runs the entire two years of the project. The software development phase will be described in the next section. The section after that will describe the field trial, and the section following that will discuss the data to be collected and plans for data analysis. This section ends with a brief analysis of minimum detectable effect sizes for this study and a discussion of contributions that the study will make regardless of whether or not a statistically significant positive effect is detected or not.

2.4.1 How the Tool will be Developed

As described earlier, the software tool is divided into three large modules: the core MCQ Writer, the Intervention Bank, and the Diagnostic Engine. Since these modules can function relatively independently of one another it is possible to develop them in parallel. Referring to the timeline, specific development activities are specified for each module over the course of the summer, fall, and winter of 2005, and spring of 2006. Each of these processes will be discussed briefly now.

Developing the MCQ Writer

At the beginning of June of 2005 two graduate student research assistants will be hired to assist with the analysis and development of the software tool. In addition, four teacher consultants will be recruited to assist with the core tool specification. The teachers will be two high school math and two science teachers from the northern New Jersey area. The teachers will be selected to represent different levels of computer using ability, i.e. novice, average, expert, but everyone will have a generally positive attitude towards computers and software. Toward the end of June the four teachers, two programmers, program director (PD), co-program director (CoPD), and the consultant who is the collaborative systems specialist (CSS) will meet for a one day design session to discuss how best to design the core MCQ Writer module of the software. Prior to the meeting the program director will provide all participants with a very high level description of the main functionality that is to be implemented, and an agenda for the meeting. The CoPD will run this first meeting as her specialty field is scenario-based human-computer interface design. The goal for the first meeting will be to produce a first draft of the overall appearance for the system, and to identify major features of the interface.

During the next month, the PD and two programmers will develop a very rough prototype refining the specification, making architectural choices as they go, and acting upon the recommendations of the teacher consultants and CSS. This stage will require research into how teachers use computers. At the end of July the group will meet again to see, interact with, and respond to the design. The goal for this meeting is to reach a relative consensus on what the main features of the system will be, and also what features will NOT be implemented. During August the PD and programmers will then work to finish a detailed software requirements specification. At the end of August, the group will meet for a third time to review this specification and to make any final recommendations or changes. The goal for the end of this third session and for the summer is to have a completed blueprint for the core MCQ Writer.

Between September and the end of November of 2005, the programmers, with support from the PD, will implement the software design agreed upon by the design group. Also during that time the PD will develop, in consultation with the CoPD, a plan for a first pilot user test of the system to coincide with the period when university instructors are developing their final exams. Between two and five volunteer NJIT instructors will be recruited to participate in the user test of the core system. Criteria for the success of the test will be the usability of the system, in addition to an evaluation of the quality of the test questions that were developed. Question quality will be determined by having instructors bring in old (i.e. from previous semesters) exams and then having judges sort questions into quality categories. Analysis of this data will involve a discussion with the instructor about the questions in light of the quality assessments made by the judges. This pilot evaluation should result in a list of the major flaws in the system that must be fixed prior to its deployment. Depending on the results, a second pilot may be scheduled prior to the midterm exam period in spring 2006.

The primary development activities for spring 2006 will be to integrate the core tool with the Intervention Bank and Diagnostic Engine. Also during spring the tool will be distributed to the four teachers who helped to design it to allow them to try it out and give feedback. The development of the tool is scheduled to end in June 2006, leaving ample time to prepare for final production, packaging, and distribution.

Developing the Intervention Bank
The types of interventions that may appear in the Intervention Bank were described earlier. During the summer of 2005, the PD and CoPD will meet with the formative assessment specialist (FAS) two or three times to plan exactly how many and what types of interventions will be built to be incorporated into the system. It is anticipated that these interventions will involve mostly consciousness raising, and self-evaluation. The interventions will be designed to target teachers who are at varying stages of change moving toward formative assessment. In addition, the PD and CoPD will meet with the CSS to discuss collaborative elements that form the basis for incorporating helping relationships into the system. By the end of August, a

complete plan and specification will be created. The PD has primary responsibility for generating this specification, but will rely on the guidance of the other team members.

Beginning in September 2005, the PD will take the finished interventions specifications and meet with the NJIT Media Services center, who will assist in developing the interventions in digitized form. By December the PD in conjunction with Media Services will have a completed set of digitized multimedia objects that will all be stored in the Intervention Bank. The Intervention Bank will be incorporated into the core MCQ Writer tool during the early spring of 2006 and tested with teachers toward the end.

Developing the Diagnostic Engine
The Diagnostic Engine is the third module to be developed, and is likely to be the most challenging. The primary goal is to develop and validate two short instruments that can be used to determine a teacher's state of change and decisional balance with respect to the adoption of formative assessment practices. As described earlier, knowledge, beliefs, and attitudes will be used as predictors of formative action since the beliefs and attitudes that underlie formative practice are far more uniform than the practice itself. The strategy is to combine a review of the literature on predictors of teachers' behavior change with an iterative process of teacher interviews, followed by prototype instrument development, and testing. Also during this time, ways to gather implicit feedback from teachers within the context of the software system will be developed.

At the same time as this is occurring, a strategy for presenting interventions to the user in response to the evaluation of the user's state of change and decisional balance must be developed. The challenge here is to make the changes subtle enough so that the user is not annoyed, or does not feel that he or she is losing control of the software. This is the module of the software that is expected to change and evolve the most over the duration of the field trial. The Diagnostic Engine will be incorporated with the other two modules during spring 2006. Upon completion the summer of 2006 will be spent preparing for the field trial which will begin with the software in the fall of 2006.

2.4.2 Evaluation Plan for the Software Tool
A field trial of the software comprises its evaluation and the second phase of this project. According to a statistical power analysis (see discussion at the end of section 2.4.3) it was decided to recruit 40 schools for the field trial of the software. High schools will be targeted from the northern and central New Jersey area, one of the most diverse populations in the U.S. Recruiting will occur during the summer of 2005, and baseline data will be collected during the 2005-2006 school year (see data collection and analysis plan below).

The field trial will be conducted as follows:

- Baseline data will be collected on student learning in year 1 of the project without any software tool intervention
- In year 2 of the project, a random set of teachers will be given and trained in the MCQ/TTM-enhanced question-building tool. A second set will be given and trained in an online non-enhanced question-building tool.
- Data will be collected on student learning in year 2 of the project.

- Data will also be collected on teacher’s use of the tools, on the assessment questions employed and on the teachers’ opinions and attitudes towards formative assessment

The following analyses will be done with the data.

- Student learning will be compared for non-tool, non-enhanced tool and enhanced tool usage.
- Judges will score the types of questions that were generated by the teachers and the question types will be compared to the log data on tool usage.
- If the correlation between system usage patterns and question types demonstrates that log data can be used to predict the type of assessment teachers generate, this data will be used to compare student learning with higher-order assessment usage patterns. If not, the actual questions will be used to compare student learning with higher order formative assessment exams.
- The TTM intervention level will also be extracted from the log data. This will also be compared to the level of questions generated by the teachers to assess whether the interventions had an impact on teacher behavior.
- Finally, the diaries and interviews will be evaluated to both assess the impact of the intervention and to capture other changes teachers might have initiated in formative assessment that would not be captured by the log data or exams.

These analyses will address the hypotheses postulated earlier. In particular, we are interested in the effect size of the intervention. Although the subject numbers are low, we should be able to obtain enough information to determine if the software solution for educating teachers in better formative assessment is a viable approach.

The next section discusses each type of data to be captured in more detail. In particular, it builds the relationships between each of the dependent variables in this study in order to support the overall model of how the software tool intervention is related to student learning improvement.

2.4.3 Data Collection and Analysis

In order to test the project’s hypotheses, data will be gathered from a number of sources including software log files, unobtrusive observations of teacher question generation activities, teachers’ evaluation/classification of peer generated questions into higher-order/lower-order groups, teacher interviews, teacher diaries, and pre- and post-test data measuring student learning. In accordance with the three research questions described above we need to collect data that informs us about:

- the type of multiple choice questions written by teachers
- teachers’ use of formative assessment practices
- students’ learning (gains) in relation to teachers (change in) behavior

This section will describe the data to be collected and the relevance of the data to the research questions.

Software Log Files as a Data Source

Log files are a rich and highly detailed data set. They consist of the capture of every keystroke and mouse click that a user makes while using a software program. Computer programs that search for specific behavior patterns of the user can process these files. For example, log files can tell us what types of templates a teacher brings up to design a specific question (e.g., the cognitive question template in Figure X, Appendix B) and whether they answered 'yes' to all items on the question checklist (Figure Y, Appendix B). We can determine if there are high correlations between the type of question created and log file patterns and, if found, use the log files as data that represents the question generating practices of a teacher. These question generating practices can then be compared across the different interventions given our teacher population.

The log file patterns in this study can provide information that is relevant to nearly all of the experiment's hypotheses. If teachers use only the core functions the tool provides—i.e. to write questions, construct, deliver and score quizzes—then it will be possible to gather data with respect to the number and types of quizzes, type and quality of questions, average quiz length, and student performance, etc. The log files will also provide information with respect to teachers' stage of change to determine how well our intervention mechanism is working.

There are, a variety of ethical issues associated with collecting log files. It is necessary that teachers who use the tool be continuously made aware and in control of this data collection and transmission. The design of the software will take great care to protect the privacy, integrity, and anonymity of any data that is stored or transmitted by the system. Users of the software will be asked for permission to use anonymous log file data. They will be told what this entails. In addition, users will be given the opportunity to cancel any data transmission. In no case will data that can be identified with named individuals be collected. A full ethical analysis will be a part of the scenario-based design process.

A significant part of the software analysis phase will be devoted to identifying the most salient data points to be gathered via log files. It will be important to focus on data that informs understanding of the types and quality of the questions that teachers write and use, to collect data that allows comparison between users of the scaffolding and TTM-enabled, and non-scaffolding –TTM-enabled versions of the tool, and to be able to collect student performance data that can indicate differences in the learning between students with teachers using the different versions of the tool. Reflection on the functionality provided by the tool indicates that all of these types of data will be readily available as byproducts of teachers' use of the tool. Even teachers' lack of use of the tool will be a source of data. It is acknowledged that there is a chance that few teachers will choose to enable the data logging functions of the software, but experience with such functionality indicates that this is only a remote possibility. Overall, it will be important to define ahead of time the key streams of data that will be monitored as there is the likelihood that there will be an overwhelming flood of usage data that will be difficult to interpret.

Student Pre and Post Tests as a Data Source
Experience with the design of experiments to measure the impacts of educational interventions indicates that statistical power is greatly increased by having baseline data about teachers and students (Bloom, 2003). Furthermore, other studies indicate that the most appropriate control for any given teacher's performance is that same teacher in the past (Black et al., 2003). For this reason it is proposed that collection of students' test score data begin one year before the software's field trial. Schools and teachers will be recruited during the summer of 2005 and asked at that time to participate in a study that will last two years. It will be explained that

during the first year participation only requires that the researchers can collect representative performance scores from students of participating science and math teachers. These scores will serve as the baseline scores and, in this way, teachers will become their own control. During the second year, all participating teachers will receive a version of the software containing only the core MCQ Writer tool with no scaffolding to support the generation of higher order questions, or the TTM-enhanced MCQ Writer tool with scaffolding. Half of the schools, selected at random, will receive the enhanced tool. If more than one teacher in a school is participating, all teachers will receive the same version of the tool to reduce possible confounding effects. Again, appropriate performance measures from students of participating teachers will be collected during the school year. It will then be possible to compare learning gains of the students in the two school years for each teacher, and also to compare the learning gains of teachers with a basic tool, and those with the enhanced tool. The appropriate pre and post measures will be determined during the summer of 2005 in consultation with Dylan Wiliam, who has experience in this area. Standard methodology for interpreting student learning gains will be used. It is anticipated that these measurements can then be correlated with measures of teachers' stage of change, decisional balance, and actual changes in behavior.

Interviews and Teacher Diaries as a Data Source

In addition to gathering log files and student test scores, two other qualitative data collection methods are planned—the use of diaries and semi-structured interviews. The purpose of these sources of data is to gather insight into teachers' thoughts and behaviors that are not directly observable. Since there will be no direct observation of teachers' classroom practices other than the log files, this will be the only way to find out from teachers what they do in the classroom and how they do it. The data gathered via these means will be used to explain what is being observed in the log files and students' test scores.

The diary study will involve giving a small number of teachers a short set of standard questions that they will answer at specified intervals. On average, the teachers will be asked to fill out the questionnaire once a week with respect to the lesson that they have just taught on that day. Teachers will be asked about what they taught, how they taught it, if this style of teaching represents a change from how they taught it before, and what the impetus for the change was. It is acknowledged that such structured reflection on one's teaching practices is likely to be a catalyst for change on its own, therefore not all of the teachers will participate in the diary study. A small subset will be chosen at random, and an equal number of teachers will be chosen from both the enhanced and non-enhanced groups. At scheduled points during the school year, the research team will read the diaries that have been written so far and use them to evaluate the progress of change. These analyses may result in modifications to the change intervention. These analyses will also aid in planning the questions for the semi-structured interviews to occur near the end of the field trial.

The final form of data collected will be semi-structured interview data. Based on intermediate analyses of the log file data, the diary studies and the student performance data, a questionnaire will be designed to ask questions that will fill the gaps in our understanding about teachers' use of the software tools, their reasons for changing or not changing their behaviors, and their knowledge, beliefs, attitudes about multiple-choice questions and formative assessment in general. We will interview a subset of the teachers. They will be chosen based on their representative and unique behavior based on the log files. Additionally, since teachers will not be barred from sharing the software with colleagues, we may choose to interview some teachers who were not a part of the original sample. Since teachers will be required to activate the software upon installation, we will be able to gather information on how many teachers outside of the original sample have chosen to try the software.

Data Analysis and Statistical Power

To summarize the data collection strategy, the four broad methods will be via user logs, pre and post tests with students, a diary study with a subset of teachers, and semi-structured interviews conducted at the end of the field trial. The analysis of much of this data will be straightforward, but for some types of data it is not possible to tell at this point what the best methodology will be. Expert consultants will be hired to advise the project during design and after collection of how to avoid pitfalls and increase the chances of having meaningful results.

As for the statistical power of the study, our study design is based largely on Bloom's (2003) guidelines for minimum detectable effect sizes. Assuming a sample of 40 schools, 20 students per class, 25% attrition over two years (leaving 30 schools), and the existence of baseline data for comparison Bloom suggests that there is a reasonable chance of detecting effect sizes on the order of .30 to .36 or larger. Given data presented by Black and Wiliam (1998b) indicating that typical effect sizes for the implementation of formative practices run from .4 to .7, there appears to be a moderate chance of detecting significantly positive effects with this intervention. It must be stressed though, that even if the study does not yield a detectable positive effect on teachers' assessment practices, it will still add a great deal to our knowledge of how to implement and deploy software-based interventions for professional development and assist IES and other organizations in making informed decisions about how to select software-based professional development interventions.

2.4.4 Ethical Considerations

The protection of the participants in this project is the highest priority. All plans and procedures involving participating schools, teachers, and students will receive thorough and appropriate professional review via NJIT's Institutional Review Board. Explicit IRB approval for all aspects of the study will be acquired prior to making contact with any potential participants. The directors of this project are familiar with and have been thoroughly trained in the ethical treatment of human participants. Any questions regarding the IRB process at NJIT, or with regard to the researchers' standing with NJIT's IRB can be directed to:

Dr. Dawn Hall Apgar
Institutional Review Board Chair
New Jersey Institute of Technology
University Heights
Newark NJ 07102-1982
(973)642-7616, dawn.apgar@njit.edu

2.5 Project Deliverables

At the conclusion of this project the following deliverables will be produced:

- **MCQ Writer**
  This is the software tool. It will have been tested and revised. It will be in a state ready to be deployed to schools on a wider basis for a larger scale field trial.
- **Metrics on teachers' state of change and decisional balance**
  This study will produce preliminary evidence supporting or rejecting the validity of the research model presented here. If teachers' knowledge, attitudes, and beliefs are in fact a good predictor of behavioral change, then these metrics will prove to be very valuable in evaluating a wide range of professional development interventions for teachers.
- **Evidence regarding appropriateness of TTM for professional development**
  While the use of the transtheoretical model for professional development situations has been suggested (Smith, 2000), it has not been studied or tried extensively, and not at all in the domain of teacher training. This study will help evaluate the viability of using this method.

With the exception of the software tool, these deliverables will be in the form of papers to be submitted to journals or professional conferences. Expected contributions are in the fields of teacher professional development, formative assessment practice, design of computer systems/interfaces to support behavioral change, and the application of the transtheoretical model.

3. Personnel

Morgan C. Benton is a Special Lecturer and doctoral candidate in Information Systems at the New Jersey Institute of Technology. At NJIT, Morgan is departmental Webmaster and has designed and built interactive systems to support teachers for the past three years. He teaches web programming and computer ethics. More recently he was selected to participate in a Summer Research Internship at the Educational Testing Service (ETS). While there he worked with Aurora Graf and Isaac Bejar on systems for automatic item generation. Morgan was also recently selected to participate in Carnegie-Mellon University's Summer School for Intelligent Tutoring Systems. Prior to coming to NJIT, Morgan spent five years teaching English as a second language to Japanese junior high students in rural northern Japan. If funded, Morgan will devote 100% of his time to this project.

Dr. Marilyn M. Tremaine will serve as the underlying guide for the interface and experiment designs in the project. She is Chair of the Department of Information Systems and Director of the degree programs in Human-Computer Interaction at the New Jersey Institute of Technology and Rutgers University. She is internationally known for her work in designing and evaluating new interface designs. She has chaired the major international profession society and served on key journal editorial boards and as conference chair for all of the major conferences in her field. She has published over 125 articles in books, journals and conferences with several of her papers having "classic" status and still being read and cited 20 years after publication.

Dr. Dylan Wiliam, the Director of the learning and Teaching Research Center at ETS will serve as a consultant for the project on issues related to formative assessment and has agreed to contribute five days per year to the study. Dr. Wiliam is internationally acknowledged as an authority on formative assessment, and has published more than 150 books, journal articles, and chapters in edited works on education. He has taught both mathematics and science in public schools in the United Kingdom, has been involved extensively in the pre-service and in-service education of schoolteachers and was, for five years, Dean of the School of Education at King's College London (the fifth largest university in the United Kingdom).

Dr. Gerry Stahl, Associate Professor in the College of Information and Technology at Drexel University will serve as a consultant for the project on issues related to math and collaborative learning and will also contribute five days per year. Dr. Stahl conducts research in computer-supported collaborative learning (CSCL) and HCI. His new book, *Collaborating with Technology: Mediation of Group Cognition* is forthcoming from MIT Press. He is Executive Editor of the *International Journal of Computer-Supported Collaborative Learning* (ijCSCL). He is the Principle Investigator of the *Virtual Math Teams Project*, a large 5-year research effort in collaboration with the Math Forum@Drexel. He served as Program Chair for the international CSCL '02 conference, Tutorials Chair for CSCL '03 and Workshops Chair for CSCL 05.

4. Resources

4.1 Technical Skills
Locating and hiring competent programmers to assist in the building of this software will not be a problem. NJIT is routinely ranked at or near the top of Yahoo! Internet Life Magazine's list of the "Most Wired" universities in the U.S. NJIT's College of Computing Sciences is home to sizable graduate programs at the master's and doctoral level in both Information Systems and Computer Science. Furthermore, there is a strong culture within the college of students becoming involved with research at all levels.

4.2 NJIT Media Services
NJIT is a global leader in providing online and distance education. Greatly responsible for that success is the NJIT Media Services Center, which specializes in the preparation of multi-media educational materials. Not only is the Media Services Center equipped with and capable of providing professional recording and production facilities and services, but also the center has members on staff who specialize in classroom assessment, and understand how best to produce technology for learning. This center is capable of assisting in the development of digitized tutorials and videos that will be embedded into the software tool. Their staff will also be able to offer assistance with regards to the content of the materials to be developed. Consideration for their time and effort has been incorporated into the budget for this proposal.

4.3 Schools
Northern New Jersey is one of the most diverse parts of the U.S. with respect to race, ethnicity, and socio-economic status. NJIT is located in downtown Newark and maintains a strong connection to the schools surrounding community. Not only does NJIT actively recruit students from local high schools, but NJIT students do volunteer work in technology labs around the city and in outreach programs with the students in those schools. Our initial efforts to contact schools in the area have met with enthusiastic response (see the e-mails and letter of support in Appendix A). Enough teachers have already volunteered to take part in the software design sessions planned for the early stages of the project. Using these and other connections, we expect to have little trouble recruiting a broad range of classrooms from varying socioeconomic and ethnic backgrounds upon which to test the tool.

5. Plans for Goal 3 Proposal

Contingent upon the favorable results of this development project, there is interest in pursuing a further grant under Goal 3 to assess the effectiveness and replicability of this project. Ideally if this software works as hoped, successful dissemination and implementation of this program for professional development will be extremely cost effective, as little or no staff time will be needed to do in-service workshops or training. Furthermore a model based on free distribution coupled with encouragement to pass the software along to colleagues may increase the effectiveness. Adoption can happen at the individual, departmental, or school level potentially removing some social or political barriers to its use. We are optimistic that some or all of these goals may be accomplished.

References

Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Educational Objectives* (Complete Edition ed.). New York: Longman.

Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C.-L. C. (1991). Effects of Frequent Classroom Testing. *Journal of Educational Research, 85*(2), 89-99.

Bangert-Drowns, R. L., Kulik, J. A., & Morgan, M. T. (1991). The Instructional Effect of Feedback in Test-like Events. *Review of Educational Research, 61*(2), 213-238.

Benton, M. C., Tremaine, M. M., & Scher, J. M. (2004, August 6-8). *Computer Aids for Designing Effective Multiple-Choice Questions.* Paper presented at the Americas Conference on Information Systems, New York, New York.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for Learning: Putting it into practice*.Berkshire, England: Open University Press.

Black, P., & Wiliam, D. (1998a). Inside the Black Box: Raising Standards Through Classroom Assessment. *Phi Delta Kapan, 80*(2), 139-149.

Black, P., & Wiliam, D. (1998b). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy and Practice, 5*(1), 7-75.

Blessing, S. (2003). *The Cognitive Tutor™: Successful Application of Cognitive Science*.Pittsburgh, PA: Carnegie Learning®.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (Eds.). (1971). *Handbook on the Formative and Summative Evaluation of Student Learning*.New York, NY: McGraw-Hill.

Bloom, H. S. (2003). *Sample Design for an Evaluation of the Reading First Program*.New York, NY: MDRC.

Brookhart, S. (1993). Teachers' grading practices: Meaning and Values. *Journal of Educational Measurement, 30,* 123-142.

Brophy, J. E. (1983). Research on the Self-Fulfilling Prophecy and Teacher Expectations. *Journal of Educational Psychology, 75*, 631-661.

Carter, K. (1984). Do teachers understand the principles for writing tests? *Journal of Teacher Education, 35*(6), 57-60.

Chi, M. T. H., Siler, S. A., & Jeong, H. (2004). Can Tutors Monitor Students' Understanding Accurately? *Cognition and Instruction, 22*(3), 363-387.

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from Human Tutoring. *Cognitive Science, 25*(4), 471-533.

Crooks, T. J. (1988). The Impact of Classroom Evaluation on Students. *Review of Educational Research, 58*(4), 438-481.

Fleming, M., & Chambers, B. (1983). Teacher-made Tests: Windows on the Classroom. In W. E. Hathaway (Ed.), *Testing in the Schools: New Directions for Testing and Measurement* (pp. 29-38). San Francisco: Jossey-Bass.

Hirst, G., DiMarco, C., Hovy, E., & Parsons, K. (1997). *Authoring and Generating Health-Education Documents That Are Tailored to the Needs of the Individual Patient.* Paper presented at the User Modeling: Proceedings of the Sixth International Conference, UM97., Vienna.

Iszák, A., & Sherin, M. G. (2003). Exploring the Use of New Representations as a Resource for Teacher Learning. *School Science and Mathematics, 103*(1), 18-27.

Koedinger, K. R. (1998, June 5-6). *Intelligent Cognitive Tutors as Modeling Tool and Instructional Model.* Paper presented at the NCTM Standards 2000 Technology Conference, Arlington, VA.

Kyriacou, C. (2001). Teacher Stress: directions for future research. *Educational Review, 53*(1), 27-35.

Minstrell, J. (1992). Facets of students' knowledge and relevant instruction. In R. Druit, F. Goldberg & H. Niedderer (Eds.), *Research in physics learning: Theoretical issues and empirical studies* (pp. 110-128). Kiel: University of Kiel, Institute for Science Education.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*.Reston, VA.

National Research Council. (1996). *National Science Education Standards*.Washington, D.C.: National Academy Press.

Prochaska, J. O., & Velicer, W. F. (1997). The Transtheoretical Model of Health Behavior Change. *American Journal of Health Promotion*, *12*(1), 38-48.

Rowe, M. B. (1974). Wait time and rewards as instructional variables, their influence on language, logic, and fate control. *Journal of Research in Science Teaching, 11*(2), 81-94.

Sherin, M. G. (2002). When Teaching Becomes Learning. *Cognition and Instruction, 20*(2), 119-150.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Smith, W. R. (2000). Evidence for the Effectiveness of Techniques to Change Physician Behavior. *CHEST, 118*(2), 8S-17S.

Stiggins, R. J. (2001). The Unfulfilled Promise of Classroom Assessment. *Educational Measurement: Issues and Practice, 20*(3), 5-15.

Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside High School Grading Practices: Building a Research Agenda. *Educational Measurement: Issues and Practice, 8*(2), 5-14.

Thissen-Roe, A., Hunt, E., & Minstrell, J. (2004). The DIAGNOSER project: Combining assessment and learning. *Behavior Research Methods, Instruments, & Computers, 36*(2), 234-240.

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why Do Only Some Events Cause Learning During Human Tutoring? *Cognition and Instruction, 21*(3), 209-249.

Velicer, W. F., & Prochaska, J. O. (1999). An Expert System Intervention for Smoking Cessation. *Patient Education and Counseling, 36*, 119-129.

Whitmer, S. P. (1983). *A descriptive multi-method study of teacher judgement during the marking process* (No. Research Series No. 122). East Lansing, MI: Michigan State University, Institute for Research on Teaching.

Morgan C. Benton

Information Systems Department
New Jersey Institute of Technology
University Heights
Newark, NJ 07102-1982
USA

973-596-5291 (off)
973-596-5777 (fax)
morgan.benton@njit.edu

Morgan Benton brings years of design experience in social and technical systems to his role as Lecturer of Information Systems at New Jersey Institute of Technology. Recent systems he has designed and built include an interactive attendance accountability system, a personalized grade delivery system that supports rich student decision-making, and a plug-in to the WebCT™ course management system to aid instructors in managing student documents. His recent research has examined how students respond to online quizzes focusing on variables of question quality and delivery mode. Morgan has diverse training in Leadership, Sociology and Computer Systems which supports his skills in designing socio-technical systems. He also has 5 years of experience teaching English as a second language in Japanese junior high schools, and 4 years teaching practice an undergraduate technical curriculum.

Professional Preparation

New Jersey Institute of Technology	Information Systems	MS 2002
University of Richmond	Leadership Studies, Sociology, & Physics	BA 1996

Appointments & Professional Experience

Summer 2004 **Research Intern**
Assessments and Scoring Research Center, Research & Development
Educational Testing Service, Princeton, NJ

2001-present **Lecturer in Web Design and Computer Ethics**
Information Systems Department
New Jersey Institute of Technology, Newark, NJ

2001-present **Webmaster and Web Developer**
Information Systems Department
New Jersey Institute of Technology, Newark, NJ

Spring 2001 **Web Usability Consultant**
Merrill Lynch Investment Managers, USA

1998-2001 **English Teacher**
Hiroo Town Board of Education, Hiroo, Hokkaido, Japan

1996-1998 **English Teacher**
Niikappu Town Board of Education, Niikappu, Hokkaido, Japan

1993-1996 **Undergraduate Research Assistant**
Assisted in publication of *An American Challenge*, a study on social trends in the Appalachian region commissioned by CORA (Commission on Religion in Appalachia)

Publications

Benton, M. C., Tremaine, M. M., & Scher, J. S. (2004). Computer Aids for Designing Effective Multiple-Choice Questions, *Proceedings of the 2004 Americas Conference on Information Systems (AMCIS)*, August 6-8, New York, NY, USA.

Benton, M. C. (2002). Standards 101: A tutorial for IT Managers. *Journal of the Standards Engineering Society*, Jan/Feb.

Benton, M. C., Kim, E., & Ngugi, B. K. (2002). Bridging the Gap: From Traditional Information Retrieval to the Semantic Web, *Proceedings of the 2002 Americas Conference on Information Systems (AMCIS)*, August 8-11, Dallas, TX, USA.

Awards and Distinctions

AMCIS 2004 Doctoral Consortium Participant

Carnegie Mellon University 2004 Summer School on Intelligent Tutoring Systems Participant

Sartain Award for the Outstanding Graduate in Sociology, University of Richmond

Kessler Award for the Outstanding Member of the Junior Class, University of Richmond

Scholarships at University of Richmond: University Scholar, Bonner Scholar, Virginia Baptist Scholar

Thesis Advisor—Dr. Marilyn M. Tremaine

Marilyn Mantei Tremaine

Information Systems Department
New Jersey Institute of Technology
University Heights
Newark, NJ 07102-1982

973-596-5824 (off)
973-596-5777 (fax)
marilyn.tremaine@njit.edu

Dr. Marilyn Tremaine is currently Professor and Chair of Information Systems. Prior to her NJIT appointment she served as a Full Professor in the Computer and Information Systems Department at Drexel University. Before moving to the U.S. Dr. Tremaine was an Associate Professor of Computer Science at the University of Toronto. She has also been Vice President of Research and Development for a software startup company and Senior Research Scientist at the EDS Center for Applied Research. She is known internationally for her work in collaborative systems and human factors software lifecycle management and is one of the top twenty cited researchers in the field of HCI. Her current research focuses on the design and development of multimodal interfaces to be used in collaborative settings.

Dr. Tremaine has been active in her professional community serving as ACM SIGCHI Vice-Chair of Communications, Chair of the Advisory Council, Vice-Chair of Finance and most recently, SIGCHI Chair. She also chaired the CHI'86 Conference on Human Factors in Computing, the CSCW'92 Conference on Computer Supported Cooperative Work, the ASSETS 2000 Conference on Assistive Technology, and is currently Technical Program Chair for Group 2003. Dr. Tremaine also serves on six technical journal editorial boards and four-to five conference program committees annually.

Professional Preparation

University of Southern California	Communication Theory	PhD 1982
University of Southern California	Communication Theory	MS 1978
University of Wisconsin	Math and Physics	BS 1979

Appointments & Professional Experience

2001-present **Professor and Department Chair**
Information Systems Department
New Jersey Institute of Technology, Newark, NJ

2000 **Visiting Distinguished Professor**
National University of Singapore, Singapore

1998-2000 **Full Professor**
College of Information Science and Technology
Drexel University, Philadelphia, PA

1988-1998 **Associate Professor**
Department of Computer Science
University of Toronto, Toronto, Ontario

1986-1988 **Senior Research Engineer**
Center for Machine Intelligence, Electronic Data Systems Corp.

1985-1986 **Vice President Product Development, Softworks, Ann Arbor, MI**

1982-1986 **Assistant Professor**
Michigan Business School, University of Michigan

1979-1982	**Lecturer** Michigan Business School, University of Michigan
1977-1979	**Visiting Scholar** Department of Psychology, Carnegie-Mellon University
1975-1977	**Research Intern** Xerox Palo Alto Research Center

Distinguished Professional Positions and Awards

Chair – ACM SIGCHI
Technical Program Chair – Graphics Interface'97 , Group 2003
Visiting Fellow – Corpus Christi College, Cambridge University
Conference Chair – CUU 2003, ASSETS 2000, CSCW'92, CHI'86
Editorial Board – ACM Transactions on CHI, ACM Interactions, Behaviour and Information Technology, Interacting with Computers, Computer-Supported Cooperative Work, and the ABLEX HCI Series
Distinguished Service Award – ACM SIGCHI - 1992
Winner Apple Design Contest – 1993
Pennsylvania's Governors School of Excellence Best Teacher Award – 1999
University of Toronto Computer Science Department Best Teacher Award - 1994

Recent Refereed Publications

1. Velez, M. C., Tremaine, M., Sarasevic, A., Doronhonceanu, B., Krebs, A., Marsic, I. Who's in Charge Here? Communicating Across Unequal Computer Platforms. *ACM Transactions on CHI,* December 2004.
2. Ngugi, B., Tremaine, M. and Recce, M. Fighting Identity Fraud with the Addition of Biometric Techniques. (2004) In *Proceedings of the American Society of Information Systems 2004 Conference*, August 5-8, 2004, New York, NY, 10 pages.
3. Zhang, S. and Tremaine, M. Incidental Information Exchange: A Measure for Assessing the Affect of Mediated Communication on Work Relationships. (2004) In *Proceedings of the American Society of Information Systems 2004 Conference*, August 5-8, 2004, New York, NY, 10 pages
4. Wu, D. and Tremaine, M. Knowledge Worker Adoption of Time Management Tools: Satisfaction and Perceived Effectiveness. (2004) In *Proceedings of the American Society of Information Systems 2004 Conference*, August 5-8, 2004, New York, NY.
5. Chen, X., Chung, J., Lacsina, P. and Tremaine, M. Mobile Browsable Information Access for the Visually Impaired. (2004) In *Proceedings of the American Society of Information Systems 2004 Conference,* August 5-8, 2004, New York, NY, 10 pages.
6. Benton, M., Tremaine, M. and Scher, J. Computer Aids for Designing Effective Multiple Choice Questions. (2004) In *Proceedings of the American Society of Information Systems 2004 Conference,* August 5-8, 2004, New York, NY, 10 pages.
7. Zhang, P., Carey, J., Te'eni, D. and Tremaine, M. Integrating Human-Computer Interaction Development into SDLC: A Methodology. (2004) In *Proceedings of the American Society of Information Systems 2004 Conference,* August 5-8, 2004, New York, NY, 10 pages.
8. Kaur, M., Tremaine, M., Huang, N., Wilder, J., Gacovski, Z., Flippo, F., and Mandravadi, C. S. Where is It? Event Synchronization in Gaze-Speech Input Systems. *Proceedings*

of the ACM 5th International Conference on Multimodal Interfaces, MMI'2003, ACM Press, New York, NY, November 2003, 151-158.

9. Adamnovich, S., Merians, A., Boian, R., Tremaine, M., Burdea, G., Recce, M. and Poizner, H. A Virtual Reality-Based Exercise System for Hand Rehabilitation Post-Stroke. *Proceedings of the IEEE Second International Workshop on Virtual Rehabilitation*, September 2003, 74-81
10. Whitworth, E., Lewis, J., Boian, R., Deutsch, J. and Tremaine, M. Formative Evaluation of a Virtual Reality Telerehabilitation System for Lower Extremity. *Proceedings of the IEEE Second International Workshop on Virtual Rehabilitation*, September 2003, 13-20.
11. Wu, D., Im, I., Tremaine, M., Instone, K. and Turoff, M. A Framework for Classifying Personalization Schemes Used on e-Commerce Websites. *In Proceedings of the 36th Hawaii International Conference on Systems Sciences – HICSS 2003,* Big Island, HI, January 2003, CDROM, 10 pages.
12. Lacsina, P., Chen, X. and Tremaine, M. Designing Non-Visual Bookmarks for the Mobile Knowledge Worker. *Proceedings of the American Conference on Inforamtion Systems 2003,* August 2003, CDROM and AIS Digital Library, 8 pages.
13. Marsic, I., Krebs, A. M., Dorohonceanu, B. and Tremaine, M. Designing and Examining PC to Palm Collaboration. *In Proceedings of the 36th Hawaii International Conference on Systems Sciences – HICSS 2003,* Big Island, HI, January 2003, CDROM, 10 pages.
14. Merians, A.S., Jack, D., Boian, R., Tremaine, M. M., Burdea, G.C., Adamovich, S., Recce, M. and Poizner, H. Virtual Reality-Augmented Rehabilitation for Patients Following Stroke. *Physical Therapy, Vol 82*, No. 9, September 2002, 898-915.
15. Williams, C. and Tremaine, M. M. SoundNews: An Audiobrowsing Tool for the Blind. *In Proceedings of the International Conference on Universal Access in Human-Computer Interaction UAHCI 2001*, Elsevier, Amsterdam, August 2001
16. Kaur, M., Tremaine, M., Wilder, J. and Hung, G. Integration of Gaze and Speech for Multimodal Human-Computer Interaction. *In Proceedings of the Third International Conference on Information, Communications and Signal Processing – ICICS 2001,* Singapore, Singapore, October 2001.

Collaborators

Prof. Joseph Wilder	Rutgers University	Prof. William A. S. Buxton Alias Research
Prof. Ivan Marsic	Rutgers University	Prof. John Mylopoulos, U. of Toronto
Prof. Grigore Burdea	Rutgers University	Prof. Ric Holt, University of Waterloo
Prof. Manish Parashar	Rutgers University	Prof. Richard Kazman, Carnegie-Mellon U.
Prof. Maria Klawe	Princeton University	Prof. Judith Olson, University of Michigan
Prof. Barry Wellman	University of Toronto	Prof. Toby Teorey, University of Michigan
Prof. Hirohiko Mori	Masaka U. - Japan	Prof. Hiroshi Ishii, MIT
Prof. Ronald Baecker	University of Toronto	Prof. Michel Beaudouin-Lafon, U. Paris-Sud
		Prof. Wendy Mackay, INRIA, France

Advisor - Prof. Allen Newell – Carnegie- Mellon University

Dylan Wiliam

Educational Testing Service
Rosedale Road
Princeton, NJ 08541
USA

609-734-1370 (office)
609-734-5010 (fax)
dwiliam@ets.org

Dylan Wiliam is a senior researcher on the Evidence-Centered Teaching in Algebra and Its Impact on Student Learning. Dr. Wiliam received his Ph.D. from London University in 1993. He is the senior research director of the Learning & Teaching Research Center in ETS Research & Development in Princeton, NJ. He was a teacher of mathematics before joining King's College London in 1984, and he continued teaching at the college level until 2003, when he joined ETS. He served as dean and head of the School of Education at King's College for five years before advancing to assistant principal there. He is the author of more than 150 articles, books, and papers on education, including several foundational pieces about formative assessment, some with Paul Black, his colleague at King's.

Dr. Wiliam has been able to combine a deep understanding of teachers, classrooms and schools with a rigorous approach to examining impact on student learning. At ETS, he provides R&D leadership in the area of formative assessment in schools, including requisite attention to the professional development of teachers.

Professional Preparation

Durham University	General Science (class III)	BS 1976
Open University (class I)	Mathematics & Science	BA 1983
South Bank Polytechnic	Mathematics Education (Distinction CNAA)	MS 1985
London University	Education	PhD 1993

Appointments & Professional Experience

2003-present **Senior Research Director**
Center for Learning & Teaching Research, Research & Development
Educational Testing Service, Princeton, NJ

2001-2003 **Assistant Principal**
King's College, University of London

1998-2003 **Professor of Educational Assessment**
King's College, University of London

1996-2003 **Dean and Head of the School of Education**
King's College, University of London

1994-1998 **Senior Lecturer in Mathematics Education**
King's College, University of London

1989-1991 **Seconded as Academic Coordinator**
Consortium for Assessment and Testing in Schools (CATS)

1986-1994 **Lecturer in Mathematics Education**
King's College, University of London

1984-1985	**Research Fellow: Graded Assessments in Mathematics** Chelsea College (later King's College, London)
1977-1980	**Teacher of Mathematics, Scale 1** Christopher Wren School (ILEA); scale 2 from January 1980
1976-1977	**Tutor in Mathematics and Physics** St. Cloud private residential VIth Form College

Selected Publications

Wiliam, D., Black, P.J., Harrison, C., Lee, C., & Marshall, B. (2003). *Assessment for Learning: Putting it into Practice*, pp. 160. Open University Press.

Wiliam, D. (2003). The impact of educational research on mathematics education. In A. Bishop, M. A. Clements, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Second International Handbook of Mathematics Education*, pp. 469-488. Dordrecht, Netherlands: Kluwer Academic Publishers.

Lester Jr., F. K., & Wiliam, D. (2002). On the purpose of mathematics education research: making productive contributions to policy and practice. In L. D. English (Ed.), *Handbook of International Research in Mathematics Education*, pp. 489-506. Mahwah, NJ: Lawrence Earlbaum Associates.

Wiliam, D., & Black, P. J. (2002). *Standards in public examinations*, pp. 48 King's College London Department of Education and professional studies.

Boaler, J., Wiliam, D., & Brown, M. L. (2002). Students' experiences of ability grouping—disaffection, polarization, and the construction of failure. *British Educational Research Journal*, 27(4), pp. 631-648.

Wiliam, D. (2000). The meanings and consequences of educational assessments. *Critical Quarterly*, 42(1), pp. 105-127.

Wiliam, D. (1998). A framework for thinking about research in mathematics and science education. In J. A. Malone, W. Atweh, & J. R. Northfield (Eds.), *Research and Supervision in Mathematics and Science Education*, pp. 1-18. Mahwah, NJ: Lawrence Earlbaum Associates.

Collaborators

Black, P. J.—King's College London
Boaler, J.—Stanford University
Harrison, C.—King's College London
Lee, C.—King's College London
Lester Jr., F. K.—University of Indiana
Marshall, B.—King's College London

Gerry Stahl

College of Information
Science & Technology
Drexel University
Philadelphia, PA 19104

215-895-0544 (office)
215-895-2494 (fax)
gerry.stahl@drexel.edu
www.cis.drexel.edu/faculty/gerry

Gerry Stahl teaches, publishes and conducts research in human-computer interaction (HCI) and computer-supported collaborative learning (CSCL). His new book, *Collaborating with Technology: Mediation of Group Cognition* is forthcoming from MIT Press. He is Executive Editor of the *International Journal of Computer-Supported Collaborative Learning* (ijCSCL). He is the Principle Investigator of the *Virtual Math Teams Project*, a large 5-year research effort in collaboration with the Math Forum@Drexel. He served as Program Chair for the international CSCL ’02 conference, Tutorials Chair for CSCL ’03 and Workshops Chair for CSCL 05. He teaches undergrad, masters and PhD courses in HCI, CSCW and CSCL.

Professional Preparation

Massachusetts Institute of Technology (MIT)	Humanities & Science (Math & Philosophy)	BS 1967
University of Heidelberg	Continental Philosophy	1967-68
University of Frankfurt	Social Theory	1971-73
Northwestern University	Philosophy	MA 1971
Northwestern University	Philosophy	PhD 1975
University of Colorado	Computer Science	MS 1990
University of Colorado	Computer Science	PhD 1993
University of Colorado	Computer Science	Postdoc 1996-99

Appointments & Professional Experience

2002-present **Associate Professor**
College of Information Science & Technology
Drexel University, Philadelphia, PA

2001-2002 **Visiting Research Scientist**
BSCW Development Team, CSCW Department, FIT
GMD and Fraunhofer Institutes, Bonn, Germany

1999-2001 **Assistant Research Professor**
Department of Computer Science & Institute of Cognitive Science
University of Colorado, Boulder, CO

1996-1999 **Post Doctoral Research Fellow**
Center for LifeLong Learning and Design
University of Colorado, Boulder, CO

1993-1996 **Director of Software R&D**
Owen Research Inc., Boulder, CO

1990-1993 **Graduate Research Assistant**
College of Environmental Design
University of Colorado, Boulder, CO

1990-1991	**Intern Interface Developer** US West Advanced Technology, Denver & Boulder, CO
1989-1990	**Instructor & Teaching Assistant** Department of Computer Science University of Colorado, Boulder, CO
1984-1989	**Director** Community Computerization Project, Philadelphia, PA

Relevant Publications

Stahl, G. (2003). Building collaborative knowing: Elements of a social theory of learning. In J.-W. Strijbos, P. Kirschner & R. Martens (Eds.), *What we know about CSCL in higher education*. Amsterdam, NL: Kluwer. Retrieved from http://www.cis.drexel.edu/faculty/gerry/publications/journals/oun/oun.pdf.

Stahl, G. (2000). Collaborative information environments to support knowledge construction by communities. *AI & Society, 14*, 1-27. Retrieved from http://www.cis.drexel.edu/faculty/gerry/publications/journals/ai&society/.

Stahl, G. (2002). *Groupware goes to school.* Paper presented at Groupware: Design, Implementation and Use -- CRIWG 2002, 8th International Workshop on Groupware, La Serena, Chile. Retrieved from http://www.cis.drexel.edu/faculty/gerry/publications/conferences/2002/criwg/Stahl_CRIWG_Paper.pdf

Stahl, G. (2002). *The complexity of a collaborative interaction.* Paper presented at the Fifth International Conference of the Learning Sciences (ICLS 2002), Seattle, WA. Retrieved from http://www.cis.drexel.edu/faculty/gerry/publications/conferences/2002/icls/ICLS_Stahl.pdf.

Stahl, G., Sumner, T., & Repenning, A. (1995*). Internet repositories for collaborative learning*: Supporting both students and teachers. Paper presented at Computer Support for Collaborative Learning (CSCL '95), Bloomington, Indiana. Retrieved from http://www.cis.drexel.edu/faculty/gerry/publications/conferences/1990-1997/cscl95/cscl.htm.

Stahl, G. (Editor) (2002). *Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of CSCL 2002*. January 7-11. 755 pages. Boulder, Colorado, USA. Hillsdale, NJ: Lawrence Erlbaum Associates. Retrieved from http://isls.org/cscl/cscl2002proceedings.pdf.

Stahl, G. (2002). Rediscovering CSCL. In T. Koschmann, R. Hall & N. Miyake (Eds.), *CSCL 2: Carrying forward the conversation* (pp. 169-181). Hillsdale, NJ: Lawrence Erlbaum Associates. Retrieved from http://www.cis.drexel.edu/faculty/gerry/publications/journals/cscl2/cscl2.pdf.

Stahl, G. (2000). *A model of collaborative knowledge-building.* Paper presented at the Fourth International Conference of the Learning Sciences (ICLS 2000), Ann Arbor, MI. Retrieved from http://www.cis.drexel.edu/faculty/gerry/publications/conferences/2000/icls/icls.pdf.

Stahl, G. (2001). WebGuide: Guiding collaborative learning on the web with perspectives. *Journal of Interactive Media in Education, 2001*(1). Retrieved from http://www-jime.open.ac.uk/2001/1.

Stahl, G., Sumner, T., & Owen, R. (1995). Share globally, adapt locally: Software to create and distribute student-centered curriculum. *Computers and Education. Special Issue on Education*

and the Internet, 24(3), 237-246. Retrieved from http://www.cis.drexel.edu/faculty/gerry/publications/journals/c&e/.

Funded Research Activities

- 2003-2008: "*Collaboration Services for the Math Forum Digital Library*" (PI) $2,300,000; sponsor: NSF IERI.
- 2003-2005: "*Collaboration Services for the Math Forum Digital Library*" (PI) $450,000; sponsor: NSF NSDL.
- 2000-2001: "*New Media to Support Collaborative Knowledge-Building: Beyond Consumption and Chat*" (PI) $19,752; sponsor: Lab for New Media at CU and the Omnicom Corporation.
- 1997-2000: "*Allowing Learners to be Articulate: Incorporating Automated Text Evaluation into Collaborative Software Environments*" (primary author and primary software developer; PIs: Gerhard Fischer, Walter Kintsch and Thomas Landauer) $678,239; sponsor: James S. McDonnell Foundation.
- 1997-2000: "*Conceptual Frameworks and Computational Support for Organizational Memories and Organizational Learning*" (co-PI with Gerhard Fischer and Jonathan Ostwald), $725,000; sponsor: NSF.
- 1999-2000: "*Interoperability among Knowledge Building Environments*" (PI) $9,124; sponsor: Center for Innovative Learning Technology / SRI.
- 1998-1999: "*Collaborative Web-Based Tools for Learning to Integrate Scientific Results into Social Policy*" (co-PI with Ray Habermann) $89,338;sponsor: NSF.

Budget Justification

Below is a narrative explaining each of the budget line items on Ed. Form 524.

1. ED FORM 524 1. Personnel

The following people will be employed on the project:

Program Director (PD)—Morgan C. Benton

As Program Director, Morgan will hire and direct the student programmers on the project. He will set up meetings with the grant's two key consultants and set up the workshops with the high school teachers who will aid in the design of the software system. In addition, he will establish the contacts with the 40 high schools and set up the testing process with the high school teachers. As the chief graduate student on the project, he will design and lay out the architecture of the software system and develop the staged interventions. He will design the multimedia tutorials for the software system and provide the teacher training. In addition, he will perform the learning and attitude collection and analyze the data.

Morgan has a faculty position of lecturer at NJIT because he has a family to support while being a Ph.D. student. As such, his position requires him to teach four courses a term. His contract is such that he cannot buy out of a portion of the courses he teaches, but rather must completely supplement his salary if he chooses not to teach. He will therefore devote 100 percent of his time to this project.

Co-Program Director, Interface Design Specialist (CoPD)—Marilyn M. Tremaine

As Co-Program Director, Dr. Tremaine will be actively involved throughout the project's lifecycle. As a senior faculty member and Morgan Benton's thesis advisor, she will help conduct the scenario-based design that will be used to develop the functionality of the system and give guidance on the user interface design and usability testing phases of the software system. She will use her position as department chair and senior faculty member to make contacts and help open doors to facilitate the conduction of the research. She will also advise on experiment design, data collection and data analysis.

The support amount listed is one month of summer salary for this ongoing contribution to the research. Approximately 15 percent of her time will be devoted to this project throughout the year.

Formative Assessment Specialist (FAS)*—Dylan Wiliam

During year 1, Dr. Wiliam will assist in the design and creation of the interventions that will be incorporated into the Intervention Bank module of the software.

During year 2, his expertise will be used to help design the field test, and interpret the analysis of the data collected, particularly with respect to the determination of the degree to which teachers adopted formative assessment practices.

Dylan Wiliam is one of the leading world experts in formative assessment interventions. His insights will help us to avoid serious misconceptions in the design of our software tool and in our data interpretation.

Collaborative Systems Specialist (CSS)*—Gerry Stahl
During year 1, Dr. Stahl will provide guidance for the design of the collaborative features to be incorporated into the software. The collaborative features will both be used for the teacher interventions and for providing peer feedback from other students using the assessments developed.

During year 2, Dr. Stahl will give us advice in developing the field test and in interpreting the data collected. He will be particularly useful in helping us assess the impact of the collaborative portion of the tools being built.

Gerry Stahl has had considerable experience working with high school teachers in the development of a collaborative math project across the United States. He will be especially useful in helping us assess the quality of the formative assessments generated with our tool.

Staff Programmers (SP)
Two graduate students will be hired each on a half-time basis for one year beginning in Summer 2005 through the end of Spring 2006. They will assist the Program Director in requirements gathering and be chiefly responsible for system implementation, testing, debugging and deployment. The stipend for the graduate students is based upon the allowable federal guidelines for employing graduate research assistants. A considerable amount of programming will need to be done to build this system. Much of this programming can be done through guided course projects, but more senior level personnel will be needed to coordinate the code development.

Teacher Consultants (TS)*
Four teacher consultants will play a central role in the design of the interface for the software system. We will use scenario-based design techniques with them to capture the functional requirements of the system and the conceptual models they use to think about assessment measures.

Role	**Cost**		
	Year 1	**Year 2**	**Total**
PD	$59,332	$61,409	$120,741
CoPD	$13,905	$14,392	$28,297
SP	$20,400	$0	$20,400
		Total:	$169,438

*NOTE: The two research consultants and four teacher consultants will be paid honoraria that are not subject to fringe benefits. The description of and justification for these honoraria will be contained in section 524.8 below, however, it seemed more appropriate to summarize their duties here.

2. ED FORM 524 2. Fringe Benefits

Fringe benefits cover both legally required benefits (FICA, workmen's compensation, etc.) and other benefits (insurance, retirement, etc.), are set by the university, and are calculated at a fixed percentage of base salary as follows:

Period	Year 1	Year 2
Academic Year	27.3%	25.9%
Summer	8.7%	6.9%
Graduates	8.7%	9.0%

Resulting in the following amounts for this grant:

Year 1: $16,423
Year 2: $13,981

Fringe benefits are not paid to consultants.

3. ED FORM 524 3. Travel

Travel costs have been budgeted as follows:

The Program Director and Co-Program Director will make two trips (one each project year) to Washington D.C. for the IES annual meeting estimated at $500/person/year.

An additional $4000 per year is budgeted for the Program Director and Co-Program Director each to travel to one additional conference estimating the cost of attending a conference at $1500-2000 per conference. Potential conferences are the AERA Annual Meeting, CHI (Computer-Human Interfaces), CSCL (Computer-Supported Cooperative Learning), ED-MEDIA (World Conference on Educational Multimedia, Hypermedia & Telecommunications).

Travel for the research consultants will be from Princeton, NJ and Philadelphia, PA to Newark NJ, a round trip of 100 and 200 miles, respectively. That amounts to 1500 miles annually for five trips, reimbursed at the federal rate of $0.375/mile comes to $562.50 per year.

Teacher consultants will be recruited from the northern and central New Jersey area and their travel is expected to average 50 miles round trip, which comes to 600 miles or $225 for the three meetings.

4. ED FORM 524 4. Equipment

There will be considerable programming efforts involved with this project. The plan is to purchase three workstation class laptop computers to support the development efforts. These machines will be distributed to the Program Director and to the two Staff Programmers among whom all of the programming efforts will be distributed. The cost of the machines is estimated to be $2500 per machine for a total of $7500.

Data integrity will be a high priority for this project. In order to protect against system crashes or other unfortunate incidents that may cause data loss, an external mass storage device will be purchased for each system. Currently an 80GB USB external hard drive for data backups costs about $125; hence the budget for these devices will be $375.

5. ED FORM 524 5. Supplies

The total budget for supplies is $4450 which covers software, office supplies, and meeting costs as described below.

Software must be purchased for the computer systems, and is estimated to cost $3000:

- **Macromedia Studio MX** which contains **Dreamweaver, Flash,** and **Fireworks** will be necessary to support the development of interactive multimedia and online components of the software.
- **Morae** (www.techsmith.com) is a professional quality software system for conducting user testing on computer systems. It will also be used to record and produce video tutorials on both the use of the software and how to develop formative multiple-choice questions.
- **EndNote, Adobe Acrobat,** and **Microsoft Office** will be used for writing and producing research reports.
- **SPSS** will be used for data analysis.
- **Microsoft Visio** and **Visual Studio.Net** will be used as the design specification and programming environment.
- In addition, utilities for data hygiene—i.e. virus scanning, spyware detection, file compression—include **Norton SystemWorks, PowerArchiver,** and **AdAware.**

Where available, university site licenses will be used to obtain software at no cost to the project. For other software, academic pricing discounts will be used. Estimated software costs are as follows (doesn't include taxes and shipping):

Title	Unit Cost	Qty	Total	
Macromedia Studio MX	$249.00	3	$747.00	
Morae	$1,103.00	1	$1,103.00	
EndNote 8	$93.95	1	$93.95	
Acrobat 7.0	$89.95	1	$89.95	
Microsoft Office Pro	$168.95	3	$0.00	(site license)
SPSS	$599.95	1	$599.95	
Visio	$139.95	3	$0.00	(site license)
Visual Studio.Net	$85.95	3	$0.00	(site license)
Norton SystemWorks	$49.95	3	$149.85	
PowerArchiver	$17.50	3	$52.50	
AdAware SE Plus	$26.95	3	$80.85	
		Total:	$2917.05	

In addition to software, the cost of office supplies such as paper, digital media (e.g. blank CDs), shipping, etc. will be budgeted at $500 per year.

In addition, to cover the costs of supplies for the three software design sessions with the teacher consultants, an additional $450 is included. This is based on $15 per person, for ten people (PD, CoPD, 2 programmers, 4 teachers, FAS, CSS) times three meetings.

6. ED FORM 524 6. Contractual

There are no costs in this category.

7. ED FORM 524 7. Construction

There are no costs in this category.

8. ED FORM 524 8. Other

There are three items in the ‘other’ category: honoraria for research consultants, honoraria for teacher consultants, and media services production costs.

We have allowed honoraria of $500 per day for each of the two research consultants, each of whom will spend five days per year working on the project. This comes to a total of $5,000 per year or $10,000.

The four teacher consultants will participate in three, one-day design sessions to be held, one each, in June, July and August of 2005. Each teacher will be paid $250 per day as an honorarium. The budget for this expense is therefore $3,000.

Interactive and multimedia elements for the software system will be developed in conjunction with the NJIT Media Services department who will provide the facilities, personnel, and expertise in recording and producing those media in digital form. The effort required to produce these materials is estimated to be about the same as the development cost for a normal semester long university course. Typically this involves 14 hours for recording the materials plus another 14 for production, totaling 28 hours at the media services rate of $100/hour. When distributing the final software package, the Media Services department will also handle production on CD-ROM media at a cost of $4 per CD. Budgeting for 100 CDs this comes to $400. Therefore the total media services budget is $3200.

10. ED FORM 524 10. Indirect Costs

Indirect costs are based upon an Indirect Cost Rate agreement between NJIT and the U.S. Department of Health and Human Services. The covered period begins on 02/26/2004 and ends 06/30/2006. NJIT’s rate agreement reimburses 51% of MDTC. The totals for years one and two are as follows:

Year 1 Total: $70,825.00
Year 2 Total: $51,430.00

11. ED FORM 524 11. Training Stipends

There are no costs in this category.

Appendix A
Figures and Letters of Support

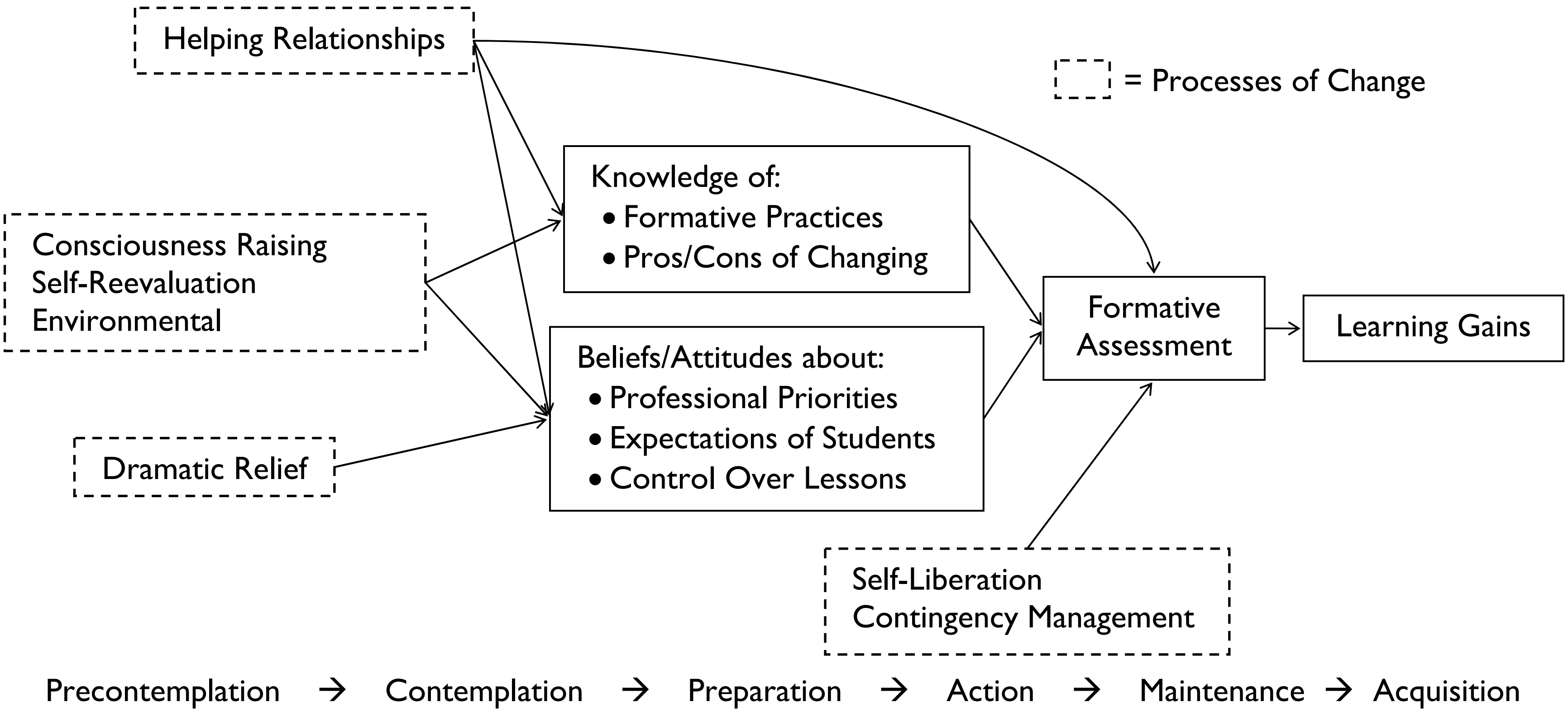

Figure 1—Research Model Integrating the K/B/A Framework with the Transtheoretical Model of Change in a Formative assessment Context

Figure 2—Timeline for Teacher Quality Grant

Month	**Software Development—Phase 1**			Evaluation in Schools Phase 2
	MCQ-writer	**Intervention Bank**	**Diagnostic Engine**	
2005 Jun	Write specification for MCQ writer.	Develop list of plans, specs for interventions	Interviews w/ teachers pros/cons processes	Recruit Schools and Teachers for Field Trial Determine appropriate baseline measures
Jul				
Aug				
Sep	Programming of core question writing tool	Recording, digitizing, building modules for software	Develop and test instruments, design data gathering for software	Collect baseline data on teacher and student performance in classes 2005-2006 school year
Oct				
Nov				
Dec	First pilot—F05 finals NJIT profs			
2006 Jan	User testing, bug fixing, pilot with HS teachers	Integrate Interventions and Diagnostics into core tool		
Feb				
Mar				
Apr	Debugging and testing final packaging and delivery strategies for field trial			
May				
Jun				
Jul				
Aug				
Sep				Begin Field Trial
Oct				
Nov				
Dec				
2007 Jan				
Feb				
Mar				
Apr				
May				End Field Trial
Jun				Final report to IES

SAINT JOSEPH HIGH SCHOOL

BROTHERS OF THE SACRED HEART

November 17, 2004

Morgan C. Benton
Department of Information Systems
New Jersey Institute of Technology
University Heights
Newark, NJ 07102-1982

Dear Morgan:

I am please that you asked me to participate in your project and gladly write this letter of support.

I am currently a teacher of computer science at St. Joseph High School in Metuchen, NJ. I have 15 years of teaching experience at the junior high, high school and undergraduate levels. This semester I am teaching, Independent Study in Computer Science (4 students), AP Computer Science (18 students) , Introduction to Computer Science – Java (25 students), Desktop Publishing (48 students) and Computer Applications (22 students) . I also moderate the Computer League and Investment Clubs. Members of the Computer League participate in the ACSL Contests, the NJIT High School Programming Contest and the TCNJ High School Programming Contest. Teams of students in the Investment Club participate in the Stock Market Game sponsored by the NJ Council on Economic Education.

I have found that many of the multiple choice questions provided by textbook publishers are ambiguous. Many times they ask students to find the "best" answer from 5 suitable answers. While this allows for some very interesting class discussions, I would prefer spending the time providing greater depth in a topic. I feel that your project will improve the writing of multiple choice questions across the curriculum. It will also allow me to better evaluate my students.

I would be willing to participate in your project by writing test questions using your software. I would also be willing during the summer months to help with the software design.

I look forward to working with you.

Sincerely,

Dorinda Sparacio

Dorinda Sparacio
Computer Science

145 PLAINFIELD AVENUE • METUCHEN, NJ 08840-1099
(732) 549-7600 • FAX (732) 549-0664

Subject: RE: NJIT Graduate Student Requests Your Help
From: "Safko,Gregory " <safkogr@winslow-schools.com>
Date: Thu, 11 Nov 2004 08:26:19 -0500
To: "'morgan.benton@njit.edu'" <mcbenton@njit.edu>

Hello, Morgan,

I would be glad to share ideas and expertise with you. I appreciate what you are going through academically and time-wise, because I am also like you; I am a full time teacher (with 2 pre-school daughters), and I am also going after my PhD in Information Systems at Nova Southeastern University.

As my schedule and time will permits, I will be glad to swap expertise. I know how grueling the coursework, idea paper, approvals, and dissertation process can be, so I think we have eachother's sympathy.

Greg

PS: I too find myself each year having ambiguous multiple choice questions sneak past me and make it onto exams, so I appreciate the concern.

-----Original Message-----
From: Morgan C. Benton [mailto:morgan.benton@njit.edu]
Sent: Thursday, November 11, 2004 7:50 AM
To: safkogr@winslow-schools.com
Subject: NJIT Graduate Student Requests Your Help

Dear Mr. Safko,

Let me introduce myself. My name is Morgan Benton and I'm a Ph.D. student in the Information Systems Department at NJIT. I'm writing to solicit the support of you and other science teachers in your school on a grant that I am writing to the US Department of Education, Institute of Education Sciences. The purpose of the grant is to develop new strategies for professional development for K12 teachers in math and science (http://www.ed.gov/programs/edresearch/applicant.html#r84305m). Specifically, I plan to develop in conjunction with teachers a software package that would help science teachers write multiple-choice questions.

What I'm asking for at this time is for you or another science teacher at your school to write a letter of support for me indicating that you are interested in my project and would be willing to participate if it gets funded. If this grant is funded, I will need the help of a small team of teachers, say three or four (in total, not from one school), to be actively involved in the software development process. In phase two, I will need the help of a greater number of teachers, from a broader range of schools, who will try out the software to determine whether or not it meets the success criteria we've set for it. I'm asking for participation from a number of schools in New Jersey and would be thrilled to have your support.

In case you're wondering how I got your address, Dr. Michael Baltrush in the Computer Science Department was kind enough to share it with me as your school participated in last year's programming contest. My request is not related to the contest from either last year or this year. I promised Dr. Baltrush that I would not harass anyone on his list, so please accept my apology if you did not wish to be contacted. You are the first person in your school that I am contacting, and I realize that you may not be the most appropriate contact person. In case you are not, I would be grateful if you would either inform me of a better person to contact, or forward this message to that person for me.

Subject: RE: NJIT Graduate Student Requests Your Help
From: Elizabeth Buniak <ElizabethBuniak@popejohn.org>
Date: Thu, 11 Nov 2004 16:50:52 -0500
To: mcbenton@njit.edu

Dear Morgan,

Thank you for your e-mail and considering my input for your important project. I will be happy to write a letter of support as I am truly very interested in this area of assessment as well. I share your concerns in generating fair diagnostic tools for the same reasons that you list based on your teaching experiences. Please count me in.

I will visit the site below to gain more information about the project. Let me know how to direct the letter of support.

Best regards,
Liz Buniak

From: Morgan C. Benton [mailto:morgan.benton@njit.edu]
Sent: Thu 11/11/2004 7:49 AM
To: Elizabeth Buniak
Subject: NJIT Graduate Student Requests Your Help

Dear Ms. Buniak,

Let me introduce myself. My name is Morgan Benton and I'm a Ph.D. student in the Information Systems Department at NJIT. I'm writing to solicit the support of you and other science teachers in your school on a grant that I am writing to the US Department of Education, Institute of Education Sciences. The purpose of the grant is to develop new strategies for professional development for K12 teachers in math and science (http://www.ed.gov/programs/edresearch/applicant.html#r84305m). Specifically, I plan to develop in conjunction with teachers a software package that would help science teachers write multiple-choice questions.

What I'm asking for at this time is for you or another science teacher at your school to write a letter of support for me indicating that you are interested in my project and would be willing to participate if it gets funded. If this grant is funded, I will need the help of a small team of teachers, say three or four (in total, not from one school), to be actively involved in the software development process. In phase two, I will need the help of a greater number of teachers, from a broader range of schools, who will try out the software to determine whether or not it meets the success criteria we've set for it. I'm asking for participation from a number of schools in New Jersey and would be thrilled to have your support.

In case you're wondering how I got your address, Dr. Michael Baltrush in the Computer Science Department was kind enough to share it with me as your school participated in last year's programming contest. My request is not related to the contest from either last year or this year. I promised Dr. Baltrush that I would not harass anyone on his list, so please accept my apology if you did not wish to be contacted. You are the first person in your school that I am contacting, and I realize that you may not be the most appropriate contact person. In case you are not, I would be grateful if you would either inform me of a better person to contact, or forward this message to that person for me.

Let me explain my idea a bit more fully. I am a teacher, too, and every semester I dread

Appendix B
Further Description of the Intervention

This section will describe the architecture of the tool as well as the design rationale behind that architecture, including the rationale for focusing on multiple-choice questions. The architecture of the software tool comes directly from the research model proposed in Figure 1. There are three elements to the architecture: the core MCQ Writer, an Intervention Bank, and a Diagnostic Engine. The MCQ Writer represents the main functionality the tool is designed to support—this module will scaffold teachers' efforts to design multiple choice questions that tap higher-order thinking and which diagnose misconceptions held by students. The Intervention Bank is a repository of different types of behavioral interventions suggested by TTM (contained in the dashed boxes in Figure 1) that will be incorporated into the tool dynamically based upon the teacher's current state of change. Lastly, in order to determine the teacher's state of change the Diagnostic Engine will gather information from users about their knowledge, beliefs, and attitudes with respect to formative use of multiple-choice questions.

The Core—MCQ Writer

This module is the core of the proposed tool. The goal of the other two modules, i.e. the Intervention Bank and the Diagnostic Engine, is to increase the chance that teachers will use this core module. The function of the MCQ Writer is to support teachers' writing and evaluation of multiple-choice questions. The specifics of this support will be described below, but first it is important to justify the choice of multiple-choice questions as the focus for this project. The selected response item—multiple-choice questions, matching, true-false etc.—is the only new type of assessment item to have been developed in the 20th century(Stiggins, 2001). As with other item types, multiple-choice questions have strengths and weaknesses and there are tradeoffs associated with their use.

Why Multiple-choice Questions?

Perhaps the main problem is that multiple-choice items are notoriously difficult to construct well(McKeachie, 2002). Common pitfalls(Benton et al., 2004; Haladyna, Downing, & Rodriguez, 2002) include, for example, questions that:

- only assess knowledge at the level of recall
- have multiple or missing correct answers
- use overly verbose or confusing language
- confuse examinees with the use of words like *not* or *except*
- are susceptible to cheating because of overuse
- have ineffective distractors
- make improper use of *all* or *none of the above*
- try to assess understanding of more than one skill or concept
- don't include the main idea in the question stem
- include obvious hints as to the correct answer
- assess knowledge in a way that is dependent upon other questions on the quiz

These pitfalls were identified and validated empirically, where possible, in a taxonomy of guidelines for appropriate question construction(Haladyna & Downing, 1989a, 1989b). Research on teacher training and the effectiveness of teacher constructed tests indicates that understanding and adoption of these guidelines is not widespread(Stiggins, 2001). Furthermore, there are additional problems in how teachers use these questions and provide feedback(Epstein et al., 2002).

Despite these problems multiple-choice questions have advantages and remain popular with teachers. Perhaps the greatest advantage to teachers is that multiple-choice questions are relatively easy to score and grade. Considering that teachers may spend up to one-third of their professional time involved with the creation, delivery and scoring of tests(Stiggins, 2001), anything that can reduce their workload may help to reduce their stress(Kyriacou, 2001). MCQs are particularly amenable to delivery with computers or bubble sheets which makes them technologically attractive. For large-scale tests, the binary nature of responses makes statistical analysis of test results easier. Most importantly, however, when constructed well, multiple-choice questions can provide a teacher with rich diagnostic information as to the state of students understanding of a subject. Well-constructed questions can and do assess higher-order thinking. A short quiz of well-planned questions at the beginning of a lesson can allow a teacher to adjust the pace and focus of the lesson to student needs, or can enable the instructor to group students so as to maximize the learning effects of peer interaction. On tests it is possible to generate automated feedback that is tailored to each student based on his or her responses.

Practically speaking, we know that teachers will continue to use multiple-choice questions regardless of their problems. Since we know that multiple-choice questions can be used effectively in formative assessment, since we know what the problems are, since we know how to train teachers to reduce these problems, and since multiple-choice questions are suited to computer-assisted test creation and delivery, multiple-choice questions were chosen to be the focus of this project.

Functionality of the MCQ Writer

The core functionality of the MCQ Writer is quite simple and comparable to that found in other computerized test management systems. Teachers will be able to enter, store, and organize multiple choice questions. Teachers can then use the entered questions to create tests that have flexible delivery options, i.e. computer delivered, printed on paper, etc. For computerized delivery, the system will also support automated scoring and the option to attach specific feedback to each answer choice and each question. Also included will be basic item analysis features such as calculation of discrimination and quartiles. The core functionality of the tool is intentionally simple so that any teacher can quickly begin using the tool.

What will make this tool stand apart is the optional support teachers will have to guide them through the question writing process. Although the definitive specifications for this functionality will be part of the system design at this point the following functionality is anticipated: content templates, examples of high-quality questions, interactive and/or video tutorials, alignment with objectives, question quality checklists, personalized feedback reports for students, and rudimentary text difficulty analysis for vocabulary and cohesion. More detailed descriptions of these features are in Appendix B. In following with the tenets of the transtheoretical model, the use of these features will be strictly optional as proposing actions to a person not ready for behavioral change can actually hurt their prospects for change. How and when the use of this functionality will be regulated by the diagnostic engine. First however, the processes of change that will be supported are described.

The Intervention Bank—Supporting Processes of Change
TTM has found that people in different stages of change use different processes to change their behavior and require different types of support. From the ten processes described those most amenable to incorporation into a software tool were selected for implementation into this software package. Again, this list will not be finalized until the full design specifications for the software are complete, but described below are the anticipated features.

Helping relationships are important to behavioral change throughout the change process, but become more so as the process moves into the preparation and action stages. Consequently means for communication and collaboration with colleagues and experts engaged in the change process will be an important part of the functionality. There are several ways this can be implemented. Directly from the tool teachers will be able to join e-mail lists or online discussion forums. Such forums are expected to cover topics ranging from technical questions about the software to the sharing of personal experiences related to classroom assessment. This will allow for the exchange of information, and provide emotional outlets for teachers involved in the change process. Another collaborative feature will be a (possibly anonymous) peer-review system for the questions that teachers write. Using this system teachers can have their questions reviewed by peers and share questions with one another. It may at some point evolve into a shepherding process as is used in the software patterns community. Not all teachers will choose to take part in helping relationships, but the functionality will exist to support those who do.

Consciousness raising will be one of the most central interventions incorporated into the system. This will include an array of different educational materials designed to be consumed at a range of levels from casual perusal to intense study. At the casual end, users may find a box in one corner of the screen with factoids such as, "Did you know that increasing the wait time after you ask a question in class to five seconds or more can greatly increase student learning? Click here to find out more…" or perhaps a quiz questions such as "What is the optimal frequency for quizzes in class? Daily, once or twice a week, once or twice a term—click here to check your answer…." Such casual messages are meant to pique interest in a non-threatening way, and to provide links to lengthier bits of information based upon research that can increase teachers' knowledge about the practices and pros of using formative assessment. At the more intense end, teachers can take advantage of video mini-lectures or download and read key research papers related to formative assessment. In addition, detailed examples of good questions will permeate the help documentation. A major goal of these functions is to increase teachers' knowledge of formative practices and of the pros of adopting them. This is in line with moving people towards the TTM action stage.

Another key process involved with change is **reevaluating oneself and one's environment**. One of the key elements of learning is being able to assess where one is in relation to a learning goal. In this case the goal is the implementation of formative assessment practice. Teachers may be given the opportunity to take a "formative practices inventory" which will point out to them areas in their practice that are more or less in line with a formative approach to teaching. The tone of this inventory will be informational, not judgmental. Teachers will also have the ability to enter in test questions from their previous exams and receive an analysis of the quality of those questions. The analysis will not be entirely automated but rather will take advantage of the functionality already implemented for judging the quality of new questions in the MCQ Writer.

Environmental reevaluation involves becoming aware of the impacts of one's actions on others. This functionality may take the form of provocative questions designed to focus the teacher's attention on how students respond in reaction to changes in the teacher's behavior. Reevaluation processes are associated with the contemplation and preparation stages of change and most likely shouldn't be introduced until the teacher has expressed some baseline interest in changing.

Dramatic relief refers to change processes that leverage people's emotions to achieve change. Typically these processes use testimonials or affective advertisements. The use of such functionality in a professional development tool raises ethical questions and must be approached with care. While it is expected that the interactions teachers have with others involved in helping relationships will entail emotional exchanges, incorporating tailored ads into a tool may be seen as manipulative and may cause more harm than good. While not being ruled out entirely, the cost/benefit analysis and implementation of such functionality will be weighed carefully.

Implementing processes to support **self-liberation** means that at the crucial juncture when teachers are ready to take action they need to have several choices available to them to channel their proactive energies. Having at least three choices of action has shown itself to be the most motivating scenario. In the case of this software and example of such choice may be the choices of different question templates, test delivery methods, or ways to provide feedback to students. Another component of self-liberation involves public statements of commitment to the change, such as New Year's resolutions. A possible feature to support such a statement might be the recommendation to form a pact with students to always conform to certain formative practices. The tool might supply a sample contract. Such a process would involve the teacher actually explaining to his or her class explicitly what kind of behaviors are going to change in the class and would create a social system that supports and reinforces the teacher's decision.

Finally, **contingency management or reinforcement** processes involve rewards or incentives to the teacher for maintaining efforts to change behavior. With respect to these processes, software is rather limited. While monetary rewards are not possible, public recognition for making changes is possible. Perhaps tests or metrics to determine a teacher's knowledge, implementation, and commitment to formative practice could be included that, when successfully passed, would result in the teacher's name being posted on a public website, trigger a letter being written to a key supervisor, or simply print out a certificate of completion or accomplishment. Another possibility is to recommend to the teachers that they find ways to reward themselves for sticking to their change goals. Implementing these features will require a bit of creativity and will be explored.

Having now discussed the plans for implementing the functionality support these seven processes of change, the question of timing their delivery needs to be addressed. According to TTM, if interventions are not stage-matched then they will have little or possibly even negative effects. The diagnostic engine will be designed to help manage the timing of interventions.

The Diagnostic Engine—Timing Interventions

A key part of the transtheoretical model is that behavioral change interventions need to be delivered so as to be in sync with the needs of the person who is changing. In principle, the procedure for doing this is straightforward. First the stage of change and/or decisional balance of the person is assessed. Second, an array of stage-matched interventions is prescribed. In practice, it may be challenging to determine the stage of change and decisional balance of teachers thinking of implementing formative practices. As discussed above, there is no concrete set of formative practices about which we can ask teachers. Rather, we are relying on teachers' knowledge, beliefs, and attitudes to predict the teacher's current state of change. The plans for developing and incorporating these metrics is discussed in section 2.4.1.