Themes in Education 1:2, 115-136, 2000

Andes: An Active Learning, Intelligent Tutoring System for Newtonian Physics

Kay G. Schulze[1], Robert N. Shelby[2], Donald J. Treacy[2], Mary C. Wintersgill[2], and Kurt VanLehn[3]
schulze@usna.edu; vanlehn@cs.pitt.edu

[1] Computer Science Department & [2] Physics Department
United States Naval Academy, USA
[3] Learning Research and Development Center
University of Pittsburgh, USA

ABSTRACT

A trend that has been emerging more strongly in the last decade has been to question the role and content of physics in contemporary higher education. Although the majority of students should study physics, most of them will not become professional physicists or engineers. Efforts have been made within the physics education community to make the subject more accessible to the average student. In addition, it is well known that students appear to learn more when personal tutoring is available as they attempt to solve homework problems. The artificial intelligence community has developed systems that try to tutor students in mathematics, chemistry, and computer programming.

Andes is an intelligent tutoring system in classical physics that has been in development by researchers at the Learning, Research, and Development Center (LRDC) at the University of Pittsburgh and the United States Naval Academy since 1996. It allows students to solve physics problems in an environment that provides visualization, immediate feedback, procedural help, and conceptual help. It currently tutors students in static forces, translational and rotational kinematics, translational and rotational dynamics, work and energy, and linear and angular momentum. The system encourages the students to draw free body diagrams, when appropriate, and enter the equations necessary to solve the problem. It allows them to ask for hints, both in terms of what they have done incorrectly as well as when they have reached an impasse and do not know how to proceed.

Andes is comprised of an expert model and a student environment. The expert model, developed by the researchers at the Naval Academy, incorporates the physics domain knowledge and pedagogical strategies that are the foundation of Andes. The student environment, developed by the researchers at the LRDC, incorporates the interface (called the Workbench), the student assessment, and the help system.

After having used Andes with a limited number of students at the Naval Academy over a three-year period, it was used in a formal study in the fall of 1999. The statistical and anecdotal results of this study are reported in this paper.

INTRODUCTION

In recent years there has been a continuing effort to make physics more accessible to the average student in survey courses at the college level [1-3]. This effort has focused not only on the instructor/student interaction in the classroom but also on the interaction of the student with the material being taught, including problem solving activity in the context of doing homework. Physics homework has traditionally been very time consuming and frustrating for students. Unsupported homework may even be a source of many of the incorrect algebraic, symbol-pushing habits that are so widespread and harmful [4, 5].

It is well known that students appear to learn more when personal tutoring is available as they attempt to solve homework problems, so many computer programs have been developed that try to tutor students doing physics homework. Many major textbooks provide a tutoring system adapted to their text, and these systems usually have the students simply enter their answers to the problems. The system determines if the answer is correct, and recognizes certain incorrect answers and provides suggestions that are specific to them. The problem with these systems is that they provide very little guidance to the students as they are trying to solve the problem. Hence, they offer little advantage over the printed answers in the back of the book. Students still find their homework frustrating, and it is certainly a less pedagogically effective experience than it could be.

A few systems have been developed that go to the other extreme. They prompt students for each step of the problem and constrain the ways that the step can be done. For instance, if the problem involves applying Newton's laws, then the tutor might first ask, "what major law would be appropriate for solving this problem?," and then provide a menu which includes "Newton's Second Law" as one of its options. If the student picks anything other than Newton's second law, he receives feedback and the chance to try again. Once he has chosen Newton's second law, he is prompted for the information necessary to apply the law. He might first be prompted for the physical object that the law will be applied to, then for the forces that are involved, then for the object's acceleration, and so on. Although such extensive guidance and constraint reduce the frustration students feel when solving problems, it is doubtful that it increases learning, and it almost certainly reduces transfer of knowledge. When students try to solve a problem without the aid of the tutor, they are not able to construct a solution because the tutor is no longer prompting them.

Studies of human tutors suggest that they may start with relatively heavy prompting and constraint, but then gradually reduce the level of guidance provided to the student [6]. In the context of intelligent tutoring systems, the prompting and constraints are called "scaffolding" and the process of removing them is called "fading" [7]. State of the art, commercial physics tutoring systems provide no scaffolding for the most part, while a few provide heavy scaffolding and no fading, but none approximate human tutors very well. One attempt to provide a compromise between the two extremes is to construct a system that contains both a no-scaffolding tutor and a heavy-scaffolding tutor and have the student alternate between them. Reif and Scott [8] have developed such a tutor with promising results.

The Andes tutoring system takes a different approach and tries to emulate the activity of a human tutor who exercises a moderate level of interaction. Instead of prompting each step, it allows the student take any action he wants. However, if the action is wrong (e.g. an incorrect equation or a vector drawn in the wrong direction), then Andes turns the erroneous entry red. Correct entries are turned green. The lack of prompting is intended to improve the probability that the skills developed using Andes will successfully transfer to the context of solving problems on paper. In fact the Andes screen is designed to simulate many of the characteristics of working on a sheet of paper. However, Andes' immediate feedback prevents students from straying onto pedagogically unproductive and frustrating paths. Immediate feedback is controversial, because it can become a crutch, however, most studies have shown that its overall impact is positive [9-11].

In order to encourage students to try to understand what their error was, Andes does not immediately provide a complete, specific correction, but has a multi-level help system that is discussed below. Andes is not the first tutor to take this approach. The approach is called "model tracing tutoring" and it has been successfully applied to elementary algebra, geometry and introductory computer programming [8, 9, 11-16]. However, Andes is the first model tracing tutor (MTT) for science. The source of power for model tracing tutoring (as well as the source of its name) is the *expert model*. It is an artificial intelligence (AI) problem solver that can generate all the acceptable solution paths for the problem. Thus, as long as the student enters equations and other steps that are part of one of the solutions encompassed by the expert model, he will continue to receive positive feedback. When he enters a step that is not part of an expert solution, the system tries to discover which correct step is closest to the one that he entered and offers help on the basis of that solution path.

THE EXPERT MODEL

The knowledge in the expert model and pedagogical strategies which guided the tutor were provided by three physics professors at the United States Naval Academy (USNA) who have more than 90 years of combined experience. Since there are no teaching assistants at the Naval Academy, considerable experience in both teaching and tutoring a wide variety of students has been accumulated.

The knowledge base (expert model) contains approximately 600 rules comprised of two types: planning and physics rules. The purpose of the planning rules is to guide the strategy by which the problem can be solved. The problem goal is initially identified and subsequently sub-goals are established which prompt the execution of the individual steps in the solution. The physics rules contain the underlying physics concepts. An example of each type of rule can be found in Figure 1.

(a) If

the problem goal is to find the magnitude of the angular acceleration
and a body exists
and the moment of inertia is known for the body

then

make a subgoal to find the net torque for the body

(b) If

the subgoal is to find the vector properties (magnitude and direction) of a linear momentum vector at time t
and the mass of the body exists
and a velocity vector for the body exists at time t
and the magnitude of the velocity vector exists

then

the magnitude of the linear momentum vector equals the mass of the body times the magnitude of the velocity vector.

Figure 1. (a) A planning rule and (b) a physics rule

In addition to the rules, a series of facts, known as the fact base is generated for each problem. The given conditions of a problem are the initial facts in the fact base and are used as input to the knowledge base. As a matter of policy, the input matches the written problem statement as closely as possible with a minimum of interpretation. Thus, given the same problem statement, the solution generated by Andes should match what a student would produce once mastery of the material has been achieved. In addition to determining all the equations necessary to solve a problem, the knowledge base also produces a graph indicating the solution paths. The equations and solution graph are then used by Andes to assess student actions and to provide appropriate feedback and hints.

When the execution cycle of the knowledge base occurs, the current facts in the fact base are examined and it is determined which of the rules have antecedents (the "if" portion) that match the current facts. All the rules for which the antecedents match the facts are prioritized based on a depth strategy. The highest priority rule is then fired or executed and the consequent (the "then" portion) is added to the fact base. In some cases, the probability that a rule will fire is manipulated by adding a "salience" statement to the rule, which can cause the priority for the rule to be increased or decreased. The statement that is added to the fact base may be a subgoal guiding what step in the solution should occur next or it may be a physics fact such as an equation or the value of one of the variables.

Pedagogical Issues

One of the primary goals in developing the knowledge base was to provide a tutoring environment that allowed the same type of student/tutor interaction that might be experienced in a one-on-one session with a professor. Although our aim was to make pedagogical decisions that would be applicable to any type of problem at the highest possible level, this was not always feasible. Pedagogical issues that were particular to specific problem types had to be handled individually [17]. For example, the objects in a static force problem are subject to a gravitational field. This fact was included in the initial facts, even though none of the problem statements explicitly state this when the planet is earth.

Inexperienced teachers often do not introduce components when students are solving one-dimensional problems. Although students find this easier at first, they frequently have difficulties making the transition to solving two and three-dimensional problems. Since these difficulties are avoided when all problems are solved using components, the decision was made very early in the development of Andes that the knowledge base would solve problems using component equations, even for one-dimensional problems. This design also provides consistency within Andes.

Many researchers have verified that active learning on the part of the student increases understanding and retention [18, 19]. To reinforce this, Andes encourages the student to draw free body diagrams when appropriate, and to draw and label vectors, etc. To help students develop effective solution paths, variables must be declared before they can be used in equations, and appropriate conventions for naming variables are encouraged. For example, forces are named Fq, where q is a letter associated with the nature of the force, e.g. a weight vector would be Fw, and a tension vector would be Ft.

It was considered important to provide flexibility in the allowed solution paths, as well as to ensure that the system solves the problem in the same manner that the students are taught. Andes demonstrates this flexibility by solving a problem using any reasonable method the student might use. This concept exhibits itself in many ways. One example of flexibility occurs when the student has a choice as to which body to choose when he is solving the problem. For example, Figure 2 demonstrates a problem where the student can find a solution using a compound body or two separate bodies. Andes will allow the student to make either choice.

Another example of flexibility is in the choice of an axis system for a free body diagram. For example, if a problem, such as the one in Figure 2, has a force whose direction is not parallel to the horizontal axis or a ninety degree rotation from it, the student can choose the x-axis to be either horizontal or in the direction of the force. When this situation occurs, the knowledge base will solve the problem using all reasonable axis choices. For a more detailed description of how this is implemented see [20].

A final example of flexibility occurs when it is possible to solve a problem in different ways.

Two blocks, one on top of the other, slide together down a frictionless plane inclined at 37 degrees. The mass of the lower block (block 1) is 30.0kg, and the mass of the top block (block 2) is 25.0 kg. Assume there is no friction between the two blocks.
What is the acceleration of the two-block system?

Answer:

What is the magnitude of the normal force acting on the top block due to the lower block?

Answer:

How many forces are acting on the top block?

Answer:

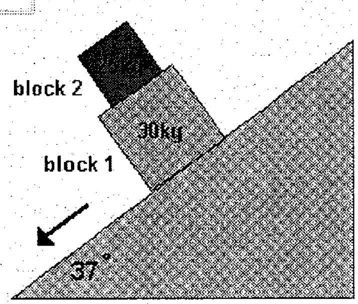


Figure 2. An Andes problem with two potential axes and two potential body choices

A kangaroo can jump about 2.50 m straight up. What is its take-off speed?

Answer:

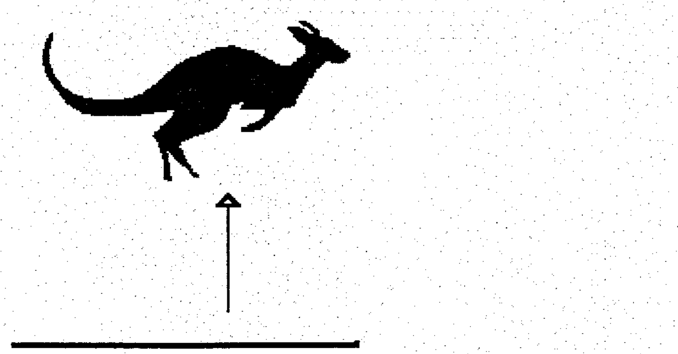


Figure 3. An Andes problem with multiple solution paths

Figure 3 is an example of a translational kinematics problem that asks the student to solve for the take-off speed of the kangaroo. That can be found using either equation (1)

$$(1)\ V_{2\_y}^2 = V_{1\_y}^2 + 2*A_{(1,2)\_y} * Disp_{(1,2)\_y}$$

or the student can use equation (2) to find the time $t_{(1,2)}$

$$(2)\ V_{2\_y} = V_{1\_y} + A_{(1,2)\_y} * t_{(1,2)}$$

and then substitute that value into either equation (3) or equation (4).

$$(3)\ Disp_{(1,2)\_y} = V_{1\_y} * t_{(1,2)} + 0.5 * A_{(1,2)\_y} * t_{(1,2)}^2$$

$$(4)\ Disp_{(1,2)\_y} = 0.5 * [V_{1\_y} + V_{2\_y}] * t_{(1,2)}$$

In this problem, none of the vectors have x-components and hence it can be solved using only y-components. In the above equations V_{1_y} represents the y-component of the velocity of the kangaroo when it takes off (initial time); V_{2_y} represents the y-component of the velocity of the kangaroo when it reaches its maximum height (final time); $Disp_{(1,2)\_y}$ represents the y-component of the displacement between the initial and the final times; and $t_{(1,2)}$ represents the length of that time interval.

Even though Andes supports vector components, this is not to say that students are required to use them if they in fact are not necessary to obtain the solution. For example, if a student is solving the one-dimensional, translational dynamics problem as shown in Figure 4, he can define the mass of the package, the acceleration vector of the package, the tension vector acting on the package, and the weight vector of the package.

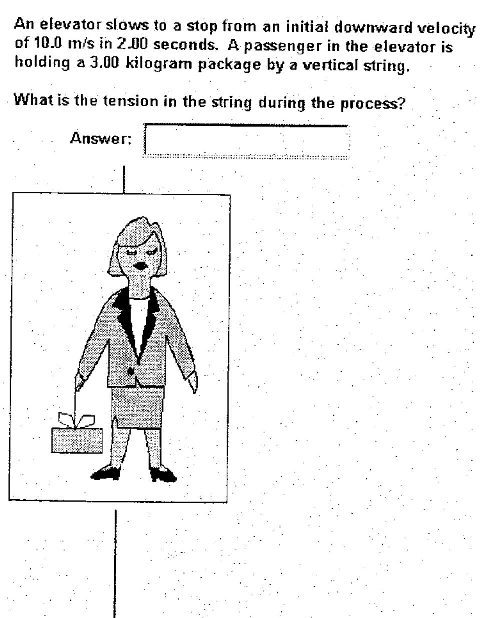


Figure 4. An Andes problem that does not require the use of vector components

In doing so, he has defined a variable *mp* representing the mass of the package, variable *A* representing the magnitude of the acceleration, a variable *Ft* representing the magnitude of the tension vector, and another variable *Fw* representing the magnitude of the weight vector. Then once he has defined a set of axes, Andes automatically provides the variables A_x, A_y, Ft_x, Ft_y, Fw_x, and Fw_y to represent the vector components of the three vectors. Since in this problem there are no x-components on any of the vectors, the knowledge base generates the equations

$$A\_y = A,\ Ft\_y = Ft,\ Fw\_y = -Fw$$

representing the relationships between the y-components of the vectors and their magnitudes. The student can then generate either the equation

$$Ft - Fw = mp * A$$

or

$$Ft\_y + Fw\_y = mp * A\_y$$

and either will be considered correct.

All vectors in the knowledge base are stored as three separate facts and are tied together by a common variable name that Andes binds to the variable name chosen by the student when he draws/defines the vector. The justification for three facts is that a vector can exist even prior to the selection of an axis and hence its magnitude and direction may not be immediately known. The initial vector fact contains the type of vector, e.g. acceleration, the body the vector is acting on, the object or interaction, if any, that caused the vector to exist, the variable name representing the vector, and the time the vector exists. Once a vector exists in the fact base, a planning rule will execute, creating a subgoal to find the vector's properties. This will eventually generate the magnitude fact and direction fact, hence completing the three facts necessary to completely define the vector. Variable names are generated to represent both the direction and the magnitude. Once the vector is completely defined and an axis has been chosen, another planning rule will execute to find the components of the vector. Unless the problem is specifically a three-dimensional problem, Andes only generates the x-component and the y-component for the vector.

Because of this philosophy, magnitudes and directions must be known before components can be obtained. This could present a problem when the problem goal is to find the magnitude of the vector. This was solved by designating the quantity as an unknown. For example, if the problem goal is to find the magnitude of a normal vector and the normal vector exists in the fact base, then the magnitude is asserted to have a value of *unknown1* and it is added to the fact base. Hence, the components can be generated using this unknown magnitude.

STUDENT ENVIRONMENT

The student environment, developed by the researchers at Learning, Research, and Development Center (LRDC) at the University of Pittsburgh, provides four capabilities. First, it allows the student to solve problems in such a way that the tutor can monitor virtually every step of the solution process. Second, it provides immediate feedback on each step, by coloring it red if it is wrong and green if it is correct. Third, it provides help if the student asks for it. Fourth, it provides an assessment of the student's knowledge of physics. This section discusses each of these four capabilities.

Andes Workbench

Most MTTs have a high bandwidth graphical user interface (GUI). A GUI has a high bandwidth if it allows students to display most of their reasoning, typically by requiring them to enter more information than they normally would if they were working on a sheet of paper [21]. For instance, students might have to define variables, as in Andes [22], or provide coherent labels for columns in tables or axes in graphs [23, 24]. Some GUIs even have students maintain a goal tree [13-15]. Many MTT designers claim that requiring students to enter this information increases learning, and there is some evidence supporting this claim [15, 25].

Andes' interface, referred to as the Workbench, provides an interactive environment in which the student works a physics problem. Figure 5 depicts the Workbench which consists of several panes and multiple tools. As the students enter objects on the Workbench, they are provided with immediate, visual, colored feedback.

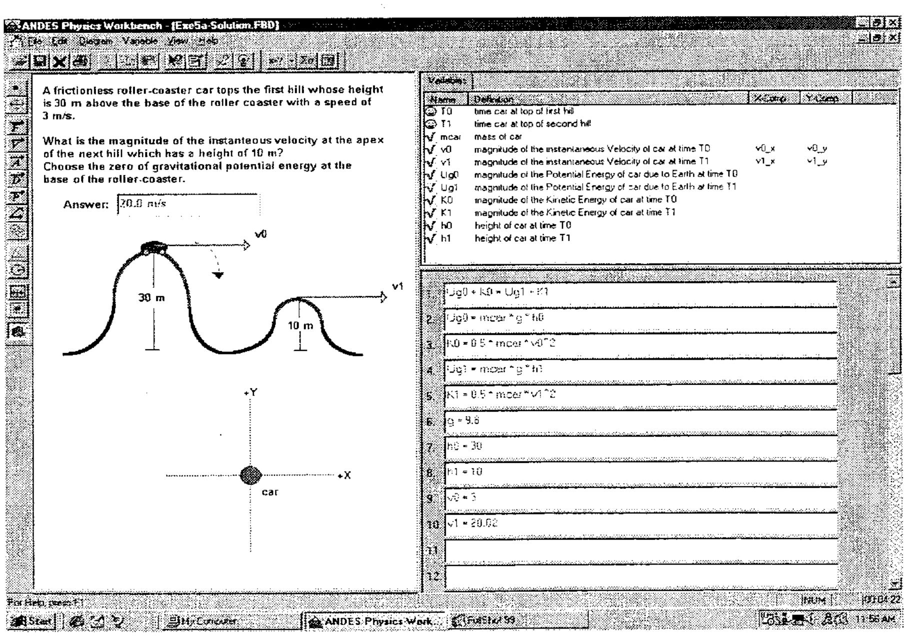


Figure 5. Andes Workbench - the interactive interface for Andes

The toolbar across the top contains the typical buttons to open, close, print, cut and paste. It has a *comment* button the student can use to provide annotated comments. In addition, it has a *help* button that provides help strictly with using the interface, a *help* button that can be selected to receive help on why an object has been turned red, and a generic *hint* button that the student can select when he has no idea what to do next. There is a *solve* button that will solve an equation for a specified variable, a *Greek keyboard* button that will allow the student to type Greek characters, and a *calculator* button.

The toolbar running down the left side of the Workbench contains buttons to aid in drawing vector quantities. There are buttons that allow the student to construct a body, an axis, a force vector, an acceleration vector, etc. Any buttons that are irrelevant to the specific problem are disabled and appear greyed out.

The main window of the Workbench is divided into three panes. The leftmost pane contains a statement of the problem along with a pictorial representation. The student can use the empty space within this pane to construct a free body diagram and draw any vector quantities that may aid in his visualization of the problem. As he constructs the diagram, he is required to assign variable names. For example, when he selects the body tool and clicks in the pane, a dialog box appears that requires him to identify the body and to name it. The variable name he selects then automatically appears in the upper right Variable Window. If he chooses an appropriate body, it will turn green in the left pane, and the variable name in the Variable Window will receive a green check mark and a notation, in English, as to what the variable represents is displayed to the right of it. When he constructs a vector, such as *V* for the velocity vector, the Variable Window will contain the variable *V* representing the magnitude of the vector, and if he has chosen an axis, the variables *V_x* and *V_y* representing the components of the vector. The bottom right pane acts as a sheet of lined paper on which he can enter equations. Students receive immediate feedback upon entering equation in this pane since incorrect equations are turned red and correct ones are turned green. When he has solved the problem, the student enters his numerical answer in the Answer box in the left pane.

Immediate Feedback

Many studies have indicated that giving immediate feedback to students while they are solving a problem can greatly increase their learning [9-11, 16]. Immediate feedback means not just telling the student when the final answer is correct, but telling him after each step in the derivation of an answer whether that step is correct. It appears that immediate feedback is effective for two reasons. First, it makes it much easier for students to determine which missing or incorrect beliefs caused their errors, and thus increases the probability they will change those beliefs. Second, immediate feedback prevents students from wandering down unproductive paths, thereby saving them time and preventing them from losing motivation.

For immediate feedback to be effective, it is best if it is provided quickly. Andes gives rapid feedback by looking up the student's step in one of two data structures. If the student's step is not an equation, Andes looks it up in the solution graph. If it is an equation, Andes converts it to a normal form and looks it up in a large table. This table contains not only the primitive equations from the solution graph, but also all equations that can be obtained by algebraic combinations of the primitive equations (See [26] for details of the treatment of equations).

The Help System

There are three help components in Andes [27-29]. The first is a simple help system for the interface itself, similar to the ones provided in most Windows applications. The student can also choose to run a video demonstrating how to interact with the Workbench. The other two types of help are used much more frequently because they give students hints about how to solve the particular physics problem on which they are working.

One type of physics help is for incorrect (red) entries. The student receives it by selecting an incorrect entry (object or equation) and then clicking on the *what's wrong?* button in the top toolbar (The icon is an *x* concatenated with a ?.) Andes will give a hint, and then allow the student to correct the entry. If the student attempts to correct it and again receives red feedback, then he can ask for further help by again selecting the entry and clicking on the *what's wrong?* button. Andes keeps track of how many hints it has given for each incorrect entry, so it knows to give the student the next hint in the hint sequence. Most of the hints also contain an *explain further* button that will cause the next hint in the sequence to be printed. For example, if the student draws the weight vector vertically upward, Andes will mark in incorrect (red). If the student asks *what's wrong* with it, Andes will pop up a window below the diagram containing the hint "Think about the direction of Fw". The student can then select an *explain further* option and receive the hint "Remember that gravity exerts a *force* straight down". He can then ask for a further explanation, at which point Andes will state that "The direction of Fw is straight down". The final hint is "The direction of the gravitational force on an object is straight down toward the center of the planet". If a hint has an underlined word in it, such as *force* in this example, he can click on the word and a short definition will pop up. In this case, clicking on force will produce "The agent of change of motion of an object".

The other physics help component provides hints to the student when he has reached an impasse and has no idea how to proceed. This help is obtained by clicking on the *what's next?* button, which has a lightbulb icon and is located on the top toolbar. For example, if the student has correctly constructed a body named 'b' and drawn the weight vector, named 'Fw', then his Variable Window will contain the variable 'mb' representing the mass of the body and 'Fw' representing the magnitude of the weight vector. If he has no idea what he should do at this point, he can click on the *what's next* button in the top toolbar and receive the hint "Write an equation relating 'g', 'Fw' and 'mb'. To receive another hint, he can select the *explain further* option in the hint or click again on the *what's next* button. This will cause Andes to hint "Write the equation Fw = g * mb". The third hint in the sequence is "The magnitude of the weight of the body near a planet is equal to the mass of the body times the gravitational acceleration of that body".

The basic idea of the hint sequence is to provide increasingly specific information that will allow the student to determine a correct step to enter. The most difficult part of this process is determining what is the correct step on which to base a hint. See [26, 29] for a discussion of how Andes does this. Once it has determined the correct step, the hint sequence is generated from a list of templates.

Currently the last hint in the hint sequence tells the student exactly what correct step to enter. Such hints are called "bottom-out" hints in most tutoring systems. They are included for two reasons. First, they act like a "just in time" example. Students often study examples in order to figure out what to do; this just makes the process more efficient. If the students study the step given to them by the bottom-out hint, they can learn what to do the next time. Unfortunately, the evaluation (discussed later) suggested that not all students actually studied the suggested step. They simply copied it onto the Workbench without thinking about it. The second reason that Andes gives bottom-out hints is to prevent a student from wasting time and becoming overly frustrated. That is, if even a direct statement of the relevant physics rule does not help the student perform the step, then this student clearly needs something special. In order to avoid wasting time, he can continue solving the problem and try to obtain the special help he needs in some other context-a different problem, the instructor's office hours, or even a lecture. Unfortunately, the evaluation (discussed later) suggests that many students abuse the bottom-out hint by routinely asking for it without first trying to take advantage of the earlier hints.

Assessment

When a student enters a correct step, it is likely that he did the reasoning needed to generate that step, and thus is familiar with the physics rules used in that line of reasoning. Although there is a small probability that he just made a lucky guess, every time Andes sees a correct step that could be generated by a rule, it should raise the likelihood that the student knows this rule. There is a system in Andes using a mathematically sophisticated technology (Bayesian networks) that was designed to do this probabilistic reasoning efficiently. The goal was for Andes to have a fine-grained assessment of the student's knowledge. For each of the approximately 600 rules in the knowledge base, a number indicating the probability the student knows that rule can be assigned.

Currently, the assessment is used in Andes for only one purpose. When Andes is trying to decide on which step to give help, the assessment sometimes plays a role [26, 29]. However, the assessment could be used in future work to help Andes pick an appropriate problem for the student to work on, or to allow students to advance through the curriculum of a self-paced physics course.

AN EVALUATION OF ANDES

Andes was evaluated at the United States Naval Academy in the fall of 1999 [30]. The experiment had two conditions: an experimental group that used Andes to do their homework and a control group, which did their homework with pencil and paper. The control group consisted of 161 students and five different instructors and the Andes group consisted of 173 students and four different instructors. In all cases the instructors were experienced and their classes have a historical record of high achievement. Students were tested both before the instruction and afterwards in order to measure their learning gains.

Methodology

The subjects represented a sample of Naval Academy students, all of whom must take a year of calculus-based physics during their sophomore year. Engineering majors account for 36% of these students, Mathematics and Science majors account for 25%, and Humanities and Social Science majors account for 39%.

Andes was distributed via the Internet, with students being given the URL for a site residing on a server maintained by the Computer Services Division at USNA. Most of the instructors who participated in the study designated a selection of twenty to thirty problems as homework assignments for which hard-copy solutions were required. The additional problems within each problem set were available for students to practice their problem-solving skills or to prepare for tests. Students uploaded log-files after each problem solving session which were identified only by the "call name" the students had self-selected.

Although many tutoring systems are evaluated in labs where there are proctors present to help students who find the tutor's help confusing, Andes was used by students in their dormitory rooms. In earlier studies, a help desk was staffed in order to provide walk-in Andes help, but it was almost never used, so it was not made available during this experiment. Although in some cases students brought Andes problems to the instructor's office hours, for the most part, they worked without any help other than that provided by Andes.

The pre-test was a local variant of an internationally used test of physics concepts (the Force Concepts Inventory) called the Physics Diagnostic Test (PDT) [31]. The post-test was a free response exam constructed to favor the control group. It consisted of four problems that were modeled on homework problems assigned to the control group from the textbook. Some of the situations presented in the examination were unfamiliar to the Andes students. This was done to avoid biasing the material tested in favor of students who had used Andes. The students were allowed to work on the exam for one hour. Both the Andes group and the control group were given written instructions in advance, indicating their work would be evaluated according to a rubric. The grading regime included 40% for the correct application of principles as demonstrated by correct symbolic equations, 30% for appropriate drawings/free-body diagrams, 20% for the correct definition of symbols or the use of standard symbols, and 10% for the correct numerical answer. Two instructors graded the examinations, each of whom selected two problems and graded all the responses to those particular problems.

Results

The post-test was graded on a 400 point scale according to the rubric issued to the students prior to the test. Table 1 indicates the results for the Andes group and the control group together with the standard deviations for each group. Also included are the results of the Physics Diagnostic Test (PDT). The overall result is that the Andes group performed statistically better than the control group on the free response test. The analysis showed that there was only a 3.6% probability that the means for the two groups were the same within statistical accuracy.

Table 1. The results of an examination given to students who used Andes and to a control group

	Number	Exam Mean (400 total)	Exam SD of data	PDT Mean (max score 30)	PDT SD of data
Andes	173	293.4	43.0	13.8	5.2
Non-Andes	161	282.0	62.1	13.7	5.2
Total	334	288.4	53.4	13.8	5.2

When only the lower half of each group was considered, as shown in Table 2, the probability that the means for the Andes and the control groups were the same decreased to 0.76%. The standard deviation of the individual scores on the examination was larger for the control group. This may simply indicate that students in the Andes group had at least some idea of how to begin solving the

Table 2. The results of an examination given to the lower half of the students

	Number	Exam Mean (400 total)	Exam SD of data	PDT Mean (max score 30)	PDT SD of data
Andes	87	281.3	49.6	9.7	2.6
Non-Andes	83	265.1	60.5	9.7	2.6

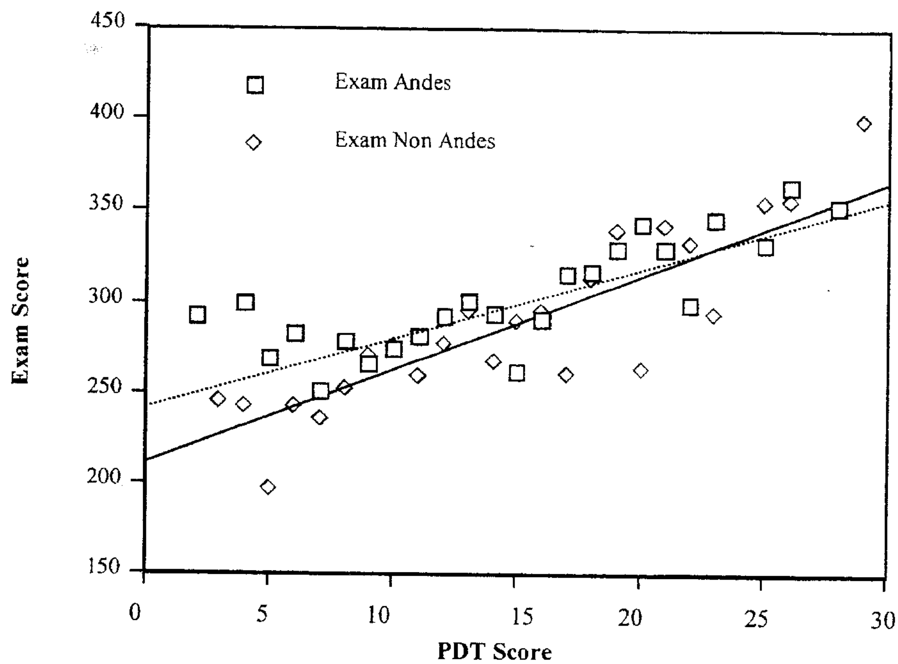
problems and so there were fewer individuals in the group with very low exam scores. The Physics Diagnostic Test scores for the two groups were very similar, as were the standard deviations. This is a strong indication that the two groups of students were conceptually comparable.

The data may be presented in another fashion. If the students are divided into three groups based on their major (engineering, math/science, and humanities/ social science), there are measurable differences among these groups as indicated in Table 3. It is noteworthy that the Andes students performed at a higher level for all of the groups. One point that should be emphasized is that the engineering majors were taking a Statics course concurrently with Physics and the material in that course was similar to a portion of the Physics course.

Table 3. A comparison of examination averages for the Andes students and for the control group broken down by general categories of student major

Major	Number	Exam Mean	PDT Mean	PDT SD
Engineering: Andes	52	316.2	15.5	5.3
Engineering: Non-Andes	55	304.7	14.8	5.0
Total Engineering	107	310.3	15.2	5.1
Math/Science: Andes	38	302.9	14.4	6.1
Math/Science: Non-Andes	42	290.0	14.6	5.6
Total Math/Science	80	296.1	14.5	5.8
Humanities/Social Science: Andes	83	277.6	12.7	4.5
Humanities/Social Science: Non-Andes	65	256.9	12.1	4.8
Total Humanities/ Social Science	148	268.5	12.4	4.6

The result of plotting the mean examination score versus the PDT score for both the Andes and the control groups is shown in Figure 6. The data points were obtained by averaging examination scores for students grouped according to their PDT scores. The data presented are weighted according to how many students had a particular PDT score and thus contributed to the average for that point. The lines are generated as least-squares best fits to the weighted data. It is clear that the exam means for students with low PDT scores are better for the students who used Andes compared to those who did not. If the students who scored less than the



Dashed Line - Andes: $y = 3.825x + 241.468$ $r^2 = 0.621$

Solid Line - Non-Andes: $y = 5.139x + 211.629$ $r^2 = 0.660$

Figure 6. The variation of the examination averages for Andes (squares) and Non Andes (diamonds) students with Physics Diagnostic Test score

class mean (13.9) on the initial administration of the PDT are treated as a group, then the exam means for the Andes and non-Andes groups are 70.3% and 66.3% respectively.

Student opinions of Andes were sampled towards the end of the semester including a question regarding a student's willingness to use Andes voluntarily. The responses to the survey demonstrated two general conclusions concerning student interactions with the tutor. The first is that the students who did a substantial number of Andes problems liked the tutor and students who did not do many problems did not like the tutor as well. It is not clear which of these factors caused the other. The second is that it is necessary to have a group of challenging problems included in a problem set for the students to appreciate what the tutor could accomplish and to justify the time spent initially learning how to interact with the Workbench.

Anecdotal Results

Having worked with Andes for five semesters, it has become apparent that no step in a physics problem solution is so simple that it won't be difficult for a large number of students. For a variety of reasons, even quite simple problems will cause some students to ask for help at every step of the solution. For a limited number of students, uncertainty about even the first step in a solution is an insurmountable difficulty, which can be overcome with a timely hint. Giving effective hints or help is very difficult. Human tutors use a variety of clues to assess the level of understanding that a student possesses at a particular point in a solution and chooses what type of hint to provide in ways that are very complex and difficult to model. It has become obvious that in most cases, hints that give guidance as to the general form of the solution are, in the long run, more effective than giving specific, possibly numerical hints which correct errors in individual equations.

The study indicated that some students were either so ill prepared or so unwilling to make an intellectual effort that they asked for help repeatedly, at every stage of the solution. Often they would proceed directly to the bottom-out hint and copy the suggested step onto the Workbench with barely a pause. This tendency to abuse the hint sequence was exacerbated by an unfortunate design choice. When the bottom-out hint suggested entering an equation, it transformed the suggested equation to make it as similar as possible to the incorrect equation for which the student had solicited help. This design was motivated by the assumption that having two similar equations to look at would make it easier for students to determine exactly what they did wrong. Unfortunately, students tended to use numbers in their equations (as discussed below) instead of algebraic expressions. That is, instead of writing T+W=m*a, they would substitute given information for as many variables as possible, and thus write, for example, 29.4=4*a. If their equation was incorrect and Andes presented a revised version in a bottom-out hint, it would also substitute numbers. For instance, if the correct equation was T-W=m*a (note the negative sign), then Andes would substitute numbers and suggest that the student write 10.5=4*a. Using numbers in the bottom-out hints made it virtually impossible for the student to determine what his error was, so he could learn virtually nothing from such hints.

Many of our students are initially reluctant to define variables or to use consistent notation. Particularly with the early problem assignments, which tend to be relatively easy and deal with topics students often feel they understand well, having to define all the variables used in an equation can seem tedious and unproductive. The procedure students must use in Andes to define vectors has the advantage of providing both a visual and a written representation, as well as requiring a complete definition of a variable before it can be introduced into an equation. One lesson learned from this study is that the majority of students, particularly the weaker ones, are aided by enforced structure and immediate feedback. The fact that this occurs in real time, as the problem solution is being generated is far more effective and less burdensome for an instructor compared with trying to enforce these behaviors by grading policies on homework assignments.

An issue of pedagogical concern arose from students' reluctance to write symbolic equations and solve them before substituting numerical values for the variables. Much of the resistance to this mode of problem solving comes from their inability (or lack of confidence in their ability) to solve problems which require algebraic manipulation. It was concluded that the primary goal of Andes was to tutor physics and that while it is desirable to give students practice in algebra, it is not essential. In response to student requests, a symbolic manipulator was built into the Workbench so that if a student has entered sufficient information in terms of the number of equations and the values of known quantities, then Andes will solve for a specified unknown. The results of this change were surprising in that the existence of this "solve" function caused a whole new subset of students to regard the Andes tutor as helpful. This is intriguing, because it has evidently introduced a significant number of students to the concept of evaluating how many equations are required to permit a solution to be obtained.

One of the greatest benefits of using Andes is that it forces students to organize their thoughts and procedures and to improve their accuracy. Andes will not accept an equation that is "almost correct", but demands that the drawings, definitions, and expressions be completely accurate. Many students find this to be frustrating, having been accustomed to relying on a human grader to interpret a partially correct or slightly inaccurate solution as being "close enough" particularly when the correct numerical answer appears at the end of the solution. Andes established these requirements to help students form robust approaches to problem solving and as such, should be transferable to a pencil and paper environment, with the algebraic manipulation done by a hand-held calculator.

FUTURE DIRECTIONS

The current version of Andes only encourages students to draw free body diagrams and coordinate axes when appropriate and allows them to use numerical values in equations at any point. Having used Andes for five semesters, we have learned that students don't like to be required to define variables or to use consistent notation. This has led to poor solution clarity. Future versions will, at the instructor's discretion, remove this flexibility and require the student to draw free body diagrams and coordinate axes when appropriate, and to draw and label all vectors. Instructors may also be able to require that students identify the physics principle they are trying to use and Andes will provide appropriate feedback concerning the principle. This feedback will have two forms. The first is whether it is the correct principle to apply and the second is feedback concerning when use of the principle has been completed. Similarly, students may be required to enter symbolic equations and delay using numerical values until the last step in the solution. This should reduce negative feedback due to incorrect algebraic equations and Andes will be able to provide more robust hints. Lastly, changes need to be made to stop the abuse of hints, but we have not yet determined how best to accomplish this.

Effective intelligent tutors should encourage deeper learning. The authors believe that natural language processing is necessary for encouraging such learning. A new project, Atlas, is developing natural language based enhancements for model tracing tutors that are modeled after human tutorial dialog, and are intended to encourage deeper learning [32]. It is hoped that Atlas can be incorporated into Andes.

ACKNOWLEDGEMENTS

This research is sponsored by the Office of Naval Research (ONR) Cognitive Sciences Division grant numbers N0001498WR30124, N0001499WR30104, N0001400WR20228, and N00014-96-1-0260.

REFERENCES

1. Redish, E. F. and Steinberg, R. N., Teaching physics: Figuring out what works, *Physics Today*, 52, 24-30 (1999)
2. Hake, R. R., Interactive engagement vs traditional methods: A six-thousand-student student survey of mechanics test data for introductory physics courses, *American Journal of Physics*, 66, 64-74 (1998)
3. Redish, E. F., The implications of cognitive studies for teaching physics, *American Journal of Physics*, 62(6), 796-803 (1994)
4. Halloun, I. A. and Hestenes, D., Common sense concepts about motion, *American Journal of Physics*, 53(11), 1056-1065 (1985)
5. Larkin, J., The role of problem representation in physics, In D. Gentner and A. Stevens (Eds.), *Mental Models*, 75-98, Hillsdale, NJ: Lawrence Erlbaum Associates (1983)
6. McArthur, D., Stasz, C. and Zmuidzinas, M., Tutoring techniques in algebra, *Cognition and Instruction*, 7(3), 197-244 (1990)
7. Collins, A., Brown, J. S. and Newman, S. E., Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics, In L. B. Resnick (Ed.), *Knowing, Learning and Instruction: Essays in Honor of Robert Glaser*, 543-594, Hillsdale, NJ: Lawrence Erlbaum Associates (1989)
8. Reif, F. and Scott, L. A., Teaching scientific thinking skills: Students and computers coaching each other, *American Journal of Physics*, 67(9), 819-831 (1999)
9. Anderson, J. R., Corbett, A. T., Koedinger, K. R. and Pelletier, R., Cognitive tutors: lessons learned, *The Journal of the Learning Sciences*, 4(2), 167-207 (1995)
10. Kulik, J. A. and Kulik, C. L. C., Timing of feedback and verbal learning, *Review of Educational Research*, 58, 79-97 (1988)

11. Schooler, L. J. and Anderson, J. R., The disruptive potential of immediate feedback, In M. Piattelli-Palmarini (Ed.), *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, 702-708, Hillsdale, NJ: Lawrence Erlbaum Associates (1990)
12. Anderson, J. R., Boyle, C. F. and Reiser, B. J., Intelligent tutoring systems, *Science*, 228, 456-462 (1985)
13. Koedinger, K. and Anderson, J. R., Reifying implicit planning in geometry: Guidelines for model-based intelligent tutoring system design, In S. P. L. a. S. J. Derry (Ed.), *Computers as Cognitive Tools*, 15-45, Hillsdale, NJ: Lawrence Erlbaum Associates (1993)
14. Reiser, B. J., Beekelaar, R., Tyle, A. and Merrill, D. C., Gil: Scaffolding learning to program with reasoning-congruent representations, In L. Birnbaum (Ed.), *The International Conference of the Learning Sciences: Proceedings of the 1991 Conference*, 382-388, Charlottesville, NC: Association for the Advancement of Computers in Education (1991)
15. Singley, M. K., The reification of goal structures in a calculus tutor: Effects on problem solving performance, *Interactive Learning Environments*, 1, 102-123 (1990)
16. Mark, M. A. and Greer, J. E., The VCR tutor: Effective instruction for device operation, *The Journal of the Learning Sciences*, 4(2), 209-246 (1995)
17. Schulze, K. G., Wintersgill, M. C., Shelby, R. N. and Treacy, D. J., The pedagogical considerations in constructing the Andes knowledge base (2000) (submitted for publication)
18. Cross, P., Teaching for learning, *AAHE Bulletin*, 39(8), 3-7 (1987)
19. Strauss, M. and Fulwiler, T., Writing to learn in large lecture classes, *Journal of College Science Teaching*, 19(3), 158-163 (1989/1990)
20. Schulze, K. G., Correll, D., Shelby, R. N., Wintersgill, M. C. and Gertner, A. A., CLIPS problem solver for Newtonian physics force problems, In C. Giarratano and G. Riley (Eds.), *Expert Systems Principles and Programming, 3rd Edition and CDROM*, Expert Systems/VIRTVONF/ PAPER14/paper14.pdf, 1-7, Boston, MA. PWS Publishing Company (1998)
21. VanLehn, K., Student modeling, In M. Polson and J. Richardson (Eds.), *Foundations of Intelligent Tutoring Systems*, 55-78, Hillsdale, NJ: Lawrence Erlbaum Associates (1988)
22. Gertner, A. S. and VanLehn, K., Andes: A coached problem solving environment for physics, In G. Gauthier, C. Frasson and K. VanLehn (Eds.), *Proceedings of ITS 2000*, 254-263, New York: Springer-Verlag (2000)
23. Aleven, V., Koedinger, K. R., Sinclair, H. C. and Snyder, J., Combating shallow learning in a tutor for geometry problem solving, In B. P. Goettle, H. M. Halff, C. L. Redfield and V. J. Shute (Eds.), *Intelligent Tutoring Systems, Proceedings of the Fourth International Conference*, 364-373, Berlin: Springer-Verlag (1998)
24. Aleven, V., Koedinger, K. R. and Cross, K., Tutoring answer-explanation fosters learning with understanding, *Artificial Intelligence in Education*, 199-206, Amsterdam: IOS Press (1999)
25. Merrill, D. C. and Reiser, B. J., Scaffolding effective problem solving strategies in interactive learning environments, In A. R. a. K. Eiselt (Ed.), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, 629-634, Hillsdale, NJ: Lawrence Erlbaum Associates (1994)
26. Gertner, A. S., Providing feedback to equation entries in an intelligent tutoring system for physics, In B. P. Goettl, H. M. Halff, C. L. Redfield and V. J. Shute (Eds.),

Proceedings of the Fourth International Conference on Intelligent Tutoring Systems, Berlin: Springer-Verlag (1998)

27. Schulze, K. G., Shelby, R. N., Treacy, D. J. and Wintersgill, M. C., Andes: A coached learning environment for classical Newtonian physics, In J. Chambers (Ed.), *Proceedings of the 11th International Conference on College Teaching and Learning*, 151-162, Jacksonville, FL The Center for the Advancement of Teaching and Learning (2000)
28. Schulze, K. G., Shelby, R. N., Treacy, D. J. and Wintersgill, M. C., Andes: An intelligent tutor for classical physics (2000) (submitted for publication)
29. Gertner, A. Conati, C. and VanLehn, K., Procedural help in Andes: Generating hints using a Bayesian network student model, *Proceedings of the 15th National Conference on Artificial Intelligence*, 106-111, Cambridge, MA: The MIT Press (1998)
30. Shelby, R. N., Treacy, D. J., Schulze, K. G. and Wintersgill, M. C., A statistical assessment of a field trial of Andes (2000) (in preparation)
31. Hestenes, D., Wells, M. and Swackhamer, G., Force concept inventory, *The Physics Teacher*, 30, 141-158 (1992)
32. VanLehn, K., Freedman, R., Jordan, P., Murray, C., Osan, R., Ringenberg, M., Rose, C., Schulze, K., Shelby, R., Treacy, D., Weinstein, A. and Wintersgill, M., Fading and Deepening: The next steps for Andes and other model-tracing tutors, In G. Gauthier, C. Frasson, and K. VanLehn (Eds.), *Proceedings of ITS2000*, New York: Springer-Verlag (2000) (in press)

ΣΥΝΟΨΗ

Μία τάση που αναδείχθηκε έντονα την περασμένη δεκαετία ήταν η διερεύνηση του ρόλου και του περιεχομένου της Φυσικής στη σύγχρονη ανώτερη εκπαίδευση. Αν και οι φοιτητές πολλών ειδικοτήτων θα πρέπει να σπουδάζουν τη Φυσική, οι περισσότεροι από αυτούς δεν πρόκειται να γίνουν επαγγελματίες φυσικοί ή μηχανικοί. Η εκπαιδευτική κοινότητα της Φυσικής έχει κάνει προσπάθειες με στόχο να καταστήσει το αντικείμενο προσιτό για το μέσο φοιτητή. Επιπρόσθετα, είναι γνωστό ότι οι φοιτητές εμφανίζονται να μαθαίνουν περισσότερα, κατά την επίλυση προβλημάτων στις εργασίες τους, με ατομική διδασκαλία. Η τεχνητή νοημοσύνη έχει αναπτύξει συστήματα που στοχεύουν στη διδασκαλία των μαθηματικών, της χημείας, και του προγραμματισμού υπολογιστών.

Το Andes αποτελεί ένα έμπειρο σύστημα διδασκαλίας της κλασσικής Φυσικής που αναπτύσσεται από ερευνητές στο Κέντρο Μάθησης, Έρευνας και Ανάπτυξης (LRDC) στο Πανεπιστήμιο Pittsburg και στη Ναυτική Ακαδημία των Ηνωμένων Πολιτειών, από το 1996. Επιτρέπει στους μαθητές να επιλύουν προβλήματα φυσικής σε ένα περιβάλλον που παρέχει οπτικοποίηση, άμεση ανάδραση, διαδικαστική και εννοιολογική βοήθεια. Χρησιμοποιείται από τους φοιτητές στις ενότητες των στατικών δυνάμεων, μεταφορικής και περιστροφικής κινηματικής, μεταφορικής και περιστροφικής δυναμικής, έργου και ενέργειας, ορμής και στροφορμής. Το σύστημα προτρέπει τους φοιτητές να σχεδιάζουν διαγράμματα ελευθέρων σωμάτων και να εισάγουν τις εξισώσεις που είναι

απαραίτητες για την επίλυση του προβλήματος. Δίνει τη δυνατότητα να ζητήσουν υποδείξεις, τόσο για τις λανθασμένες επιλογές τους όσο και για τις περιπτώσεις εκείνες, που έχουν φθάσει σε κάποιο σημείο και δεν γνωρίζουν πώς να προχωρήσουν.

Το Andes περιλαμβάνει ένα έμπειρο μοντέλο και το περιβάλλον του φοιτητή. Το έμπειρο μοντέλο αναπτύχθηκε από τους ερευνητές στη Ναυτική Ακαδημία και συνδυάζει τις γνώσεις της Φυσικής και τις παιδαγωγικές στρατηγικές που αποτελούν τη βάση του Andes. Το περιβάλλον φοιτητή αναπτύχθηκε από τους ερευνητές στο LRDC και συνδυάζει το περιβάλλον διεπαφής (που λέγεται Workbench), την αξιολόγηση του φοιτητή και το σύστημα βοήθειας.

Αφού χρησιμοποιήσαμε το Andes σε ένα περιορισμένο αριθμό φοιτητών στη Ναυτική Ακαδημία για μια περίοδο τριών ετών, ακολούθησε η χρήση του σε μια τυπική μελέτη, στα τέλη του 1999. Στην παρούσα εργασία παρουσιάζονται τα στατιστικά αποτελέσματα της μελέτης αυτής.