

An Interactive and Intelligent Learning System for Physics Education

Lakshman S. Myneni, N. Hari Narayanan, Sanjay Rebello,
Amy Rouinfar, and Sadhana Puntambekar

Abstract—Students tend to retain naïve understandings of concepts such as energy and force even after completing school and entering college. We developed a learning environment called the Virtual Physics System (ViPS) to help students master these concepts in the context of pulleys, a class of simple machines that are difficult to assemble and use in the real world. Several features make the ViPS noteworthy: it combines simulation and tutoring, it customizes tutoring to address common misconceptions, and it employs a pedagogical strategy that identifies student misconceptions and guides students in problem solving through virtual experimentation. This paper presents the ViPS and describes studies in which we evaluated its efficacy and compared learning from the ViPS with learning from constructing and experimenting with real pulleys. Our results indicate that the ViPS is effective in helping students learn and remediate their misconceptions, and that virtual experimentation in the ViPS is more effective than real experimentation with pulleys.

Index Terms— Physics, evaluation/methodology, intelligent tutoring systems, computer-assisted instruction, education

1 INTRODUCTION

IT is hard to imagine a modern education without computers. The use of computers can be beneficial for teachers and learners. Intelligent Tutoring Systems (ITS) exemplify this, by tracking a student's progress and providing tailored feedback and hints along the way. By collecting information on a particular student's performance, an ITS can make inferences about a student's strengths and weaknesses, and can suggest additional work. This paper presents the design and evaluation of an intelligent simulation and tutoring system called the Virtual Physics System (ViPS) for exploring and learning physics concepts within the context of a particular class of simple machines.

One goal of middle school science instruction is to inculcate deep knowledge of fundamental physics concepts such as energy, force, work, and mechanical advantage in students through hands-on work with simple machines like inclined planes and pulleys. However, learning about simple machines, especially pulley systems, is in itself a challenging task for many students. In addition, teachers face the difficulty of helping students abstract what is learned in the context of hands-on work to a more general understanding of physics concepts. We address this

problem through a two-pronged approach: 1) by making it easier for students to construct, simulate, and experiment with simple machines in a virtual environment; and 2) by integrating a tutoring component with the simulation component. We chose pulley systems not only because students generally find them harder to understand than simpler machines like inclined planes, but also because complex pulley setups (e.g., those involving compound pulleys with multiple grooves or many movable pulleys) are so difficult to correctly build and experiment with in the real world within the limited class time available that teachers tend to limit hands-on activities to very simple setups. Furthermore, there are experimental setups such as those with no friction that are impossible to construct and test in the real world.

Our research is part of a larger project to investigate the teaching and learning of physics concepts in middle schools. It is difficult for a teacher to track the progress of students individually in a class with many students. A teacher may not always know who is having difficulty during a group hands-on activity, may not be able to tell why a student is having difficulties, and may not have enough time to look into every students' needs in a large class. A tutoring system coupled with an experimentation and simulation environment, on the other hand, will be able to track each individual student's problem solving activities, such as the construction of a simple machine, the running of it, and solving problems based on such simulations, and provide individualized feedback. This is the primary motivation behind the development of the ViPS.

- L.S. Myneni and N.H. Narayanan are with the Interactive and Intelligent System Research Laboratory, Department of Computer Science and Software Engineering, Auburn University, Auburn, AL 36849.
E-mail: {mynenls, naraynh}@auburn.edu.
- S. Rebello and A. Rouinfar are with the Physics Education Research Group, Department of Physics, Kansas State University, Manhattan, KS 66506.
E-mail: srebello@phys.ksu.edu, amy.rouinfar@gmail.com.
- S. Puntambekar is with the Learning Sciences Program, Department of Educational Psychology, University of Wisconsin, Madison, WI 53706.
E-mail: puntambekar@education.wisc.edu.

Manuscript received 26 Feb. 2012; revised 31 Dec. 2012; accepted 17 June 2013; published online 9 July 2013.

For information on obtaining reprints of this article, please send e-mail to: lt@computer.org, and reference IEEECS Log Number TLTSI-2012-02-0028. Digital Object Identifier no. 10.1109/TLT.2013.26.

2 BACKGROUND

Tutoring is known to improve student learning. For instance, Reiser et al. [19] reported that students working with private tutors could learn given material four times

faster than students who attended traditional classroom lectures, studied textbooks, and worked on homework alone. When a human tutor is not available, the next best option may be an Intelligent Tutoring System (ITS). An ITS is a computer-based instructional system that has knowledge bases for instructional content and teaching strategies. It attempts to acquire and use knowledge about a student's level of mastery of topics to dynamically adapt instructions. Anderson and Skwarecki [1] reported that an ITS is a cost-effective means of one-on-one tutoring to provide novices with the individualized attention needed to overcome learning difficulties. ITS are not only being used in academia to augment classroom teaching, but have also penetrated various industries where companies are using these systems to train employees to perform their job functions. ITS have been built for various domains such as mathematics, medicine, engineering, public services, computer science, and so on, and have been shown to be effective [8], [9], [18], [20], [25]. The potential of ITS for helping students learn is widely recognized.

Many researchers have described the affordances and limitations of problem solving using physical manipulatives and computer simulations in science education research [5]. For example, Tan and Biswas [21] found, in an experiment comparing students who used a simulation with a control group of students, that students who learned from the simulation were more motivated, engaged, and scored higher in a post-test. Zacharia and Anderson [27] investigated the effects of interactive computer-based simulations, presented prior to inquiry-based laboratory experiments, on students' conceptual understanding of mechanics. They found that the use of simulations improved students' ability to generate predictions and explanations of the phenomena in the experiments. Triona et al. [23] investigated how physical and virtual manipulatives affected student learning about mousetrap cars. Students used either physical or virtual manipulatives to design their cars. The physical and virtual treatments showed the same effectiveness in helping students design cars. Finkelstein et al. [7] looked at how students learned about circuits differently with virtual or physical manipulatives. The simulations used by the students were similar to the physical materials, except that the simulations showed electron flow within the circuit, which the physical materials could not. They reported that the students who had used virtual manipulatives, i.e., the simulations, scored better in an exam and were able to build physical circuits more quickly than students who used physical manipulatives. Zacharia et al. [28] looked at physical and virtual manipulatives in the context of heat and temperature. One group of students used physical manipulatives, while the other group of students used physical manipulatives followed by virtual manipulatives. Students who worked with physical followed by virtual manipulatives performed better in a conceptual test than students who only used the physical manipulatives. The authors' conclusion was that one reason for the addition of simulation increasing student learning was that simulations could be manipulated more quickly than physical setups.

Our research combines the two strands of intelligent tutoring and virtual manipulation by designing, developing, and testing a system, ViPS, with both capabilities. The

system employs the instructional technique of coached problem solving [24]. Problem solving is embedded within a pedagogical approach with three stages: identification of misconceptions by means of a pre-test, guided problem solving to address misconceptions, and assessment of learning by means of a post-test. These stages correspond with the three stages of Minstrell's facet-based instruction [16]: identifying what students do and do not know prior to instruction, benchmark instruction in the classroom to initiate change in understanding and reasoning, and diagnostic assessment embedded within instruction. We build upon these ideas by adapting them to remediation of misconceptions through problem solving in an ITS. We designed the system's interfaces in accordance with the cognitive theory of multimedia learning [13]. Furthermore, the ViPS is designed to detect and help address the following misconceptions regarding pulleys that students commonly exhibit.

The ViPS detects which of these misconceptions a student has by asking the student to solve a set of problems at the beginning. The problem solving involves answering questions about pulley setups after constructing and running them in the simulation environment. Based on this, the ViPS constructs a student model. This model, that is continually updated throughout the tutoring session, is used for generating additional problems for the student to experiment with, and for providing hints and other kinds of automatic feedback based on the students' knowledge state. The ViPS combines a virtual experimentation environment [2] that allows simulations [8] with a tutoring component based on the notion of coached problem solving [9] to address student misconceptions.

3 DESCRIPTION OF THE ViPS

The ViPS provides a student with an interactive simulation and tutoring environment in which pulley setups can be created and simulated. Components required for pulley setups can be created and manipulated using a drag and drop interface. Students are asked by the ViPS to solve problems in this environment by creating and running pulley simulations. As a student is working toward a solution, the system keeps track of his or her actions and provides feedback to help the student make progress.

The architecture of the ViPS, implemented in Java, is shown in Fig. 1. It consists of a graphical user interface that manages interaction with students, a simulation module that creates and simulates the pulley setups built by students, a feedback module that generates appropriate messages for the students, a knowledge evaluator that evaluates the knowledge of the student, a tutor module that tutors the student for misconceptions, a student model that includes the history of student interactions and various measures of student performance, a domain knowledge model that represents domain knowledge, a database of problems, and a procedural knowledge model that represents student solution paths within individual problems.

3.1 Graphical User Interface

The graphical user interface is responsible for all the interactions with the students. This interface is divided

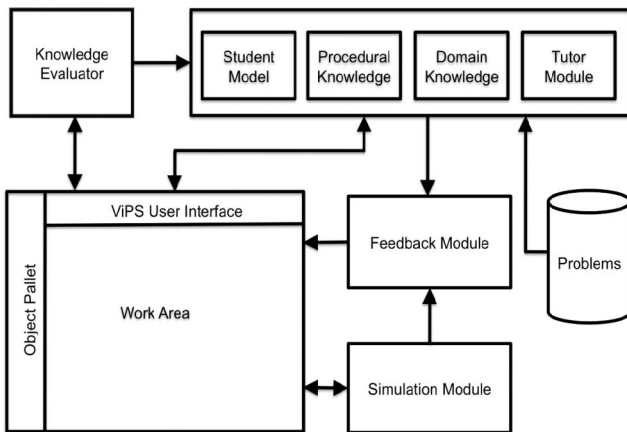


Fig. 1. Architecture of the ViPS.

into two main parts: a tabbed work area for creating pulley setups and solving problems, and an object pallet for selecting the components required to create a pulley setup. A snapshot of the interface can be seen in Fig. 2. Using this interface, students can create a pulley setup by dragging the required components from the object pallet onto the work area and clicking on the thread button. Students can also interactively manipulate various parameters of the components, like the size of a pulley, value of the load, and so on. A problem is given to a student in the form of textual and pictorial representations (see Fig. 3). The student is asked to solve the problem by creating the setups required to answer the question, running the simulations (see Fig. 4), and comparing the simulation outputs of the setups created. The problems in the ViPS were designed and checked by experienced physics educators.

3.2 Tutor Module

The tutor module is responsible for overseeing the process of tutoring a student for the misconceptions he/she might have, and it is also responsible for overseeing the process of student problem solving by using the information generated by the student model to select and present appropriate problems. It uses a decision algorithm to determine the level of coaching to be provided, and interacts with the

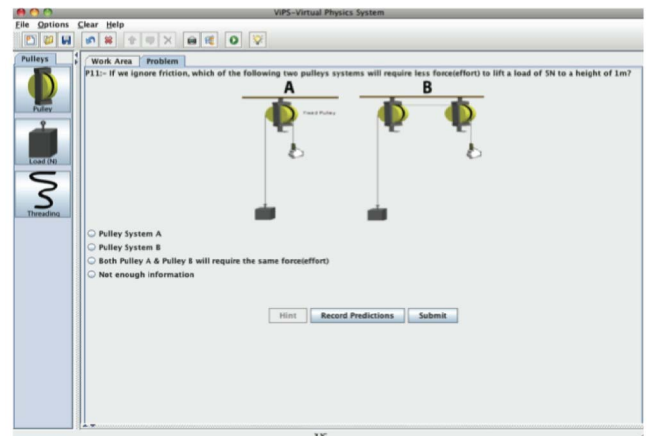


Fig. 3. The ViPS problem view.

feedback module to generate appropriate hints. Our process of designing content for the tutor adhered to the principles stated in the theory of multimedia learning [10], and the tutoring done by the tutor module is based on the Zone of Proximal Development (ZPD) component of Vygotsky's theory of learning [26].

Mayer's cognitive theory of multimedia learning [13] proposes five principles of how to design multimedia to help students understand scientific explanations.

1. *Multiple representation principle.* It is better to present an explanation in words and pictures than solely in words.
2. *Contiguity principle.* When giving a multimedia explanation, present corresponding words and pictures contiguously rather than separately.
3. *Split-attention principle.* When giving a multimedia explanation, present words as auditory narration rather than visual on-screen text.
4. *Individual differences principle.* The foregoing principles are more important for low-knowledge than high-knowledge learners, and for high-spatial rather than low-spatial learners.
5. *Coherence principle.* When giving a multimedia explanation, use few rather than many extraneous words and pictures.

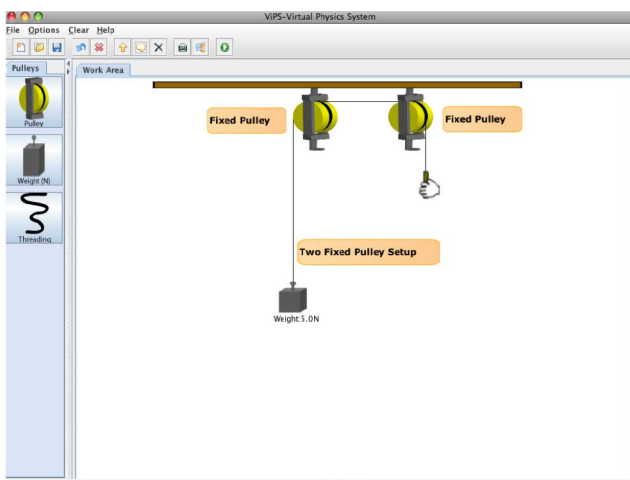


Fig. 2. The ViPS work area.

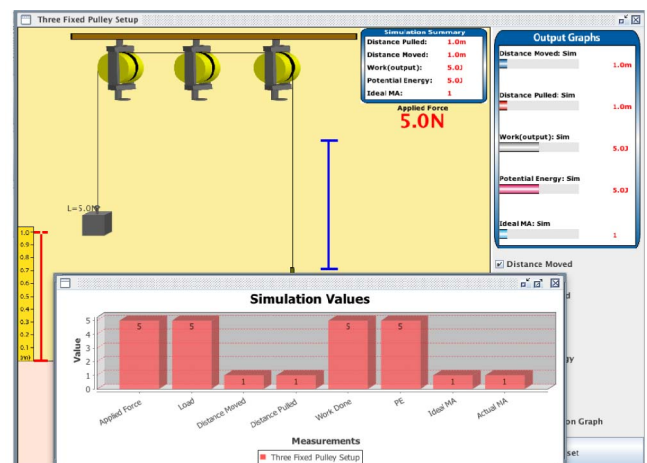


Fig. 4. The ViPS simulation window.

TABLE 1
Different Misconceptions Addressed by the ViPS

	Definition
Misconception 1	The more pulleys there are in a setup, the easier it is to pull to lift a load.
Misconception 2	The longer the string in a pulley setup, the easier it is to pull to lift a load.
Misconception 3	Pulling upwards is harder than pulling downwards.
Misconception 4	Having more pulleys in a pulley setup reduces the amount of work.
Misconception 5	Size (radius) of pulleys in a pulley setup affects the amount of work.
Misconception 6	Improper understanding of force and work.

The graphical interface of the ViPS has been designed to present pictorial and textual information during simulation and tutoring to students in accordance with the first, third, and fifth principles of the cognitive theory of multimedia learning (see Fig. 7).

Vygotsky’s theory is one of the foundations of constructivism. The ZPD is an important component of this theory. The ZPD is the distance or gap between a student’s ability to perform a task with guidance of the teacher and his/her ability to do it independently. According to Vygotsky, learning occurs in this zone. Tutoring by ViPS occurs in the ZPD because it is initiated only when the student is unable to independently solve a set of problems, and it then assists the student to solve these problems under the system’s guidance.

Fig. 6 shows the tutoring process followed by the tutoring module to remediate any misconceptions students might have with respect to pulley systems.

The interaction between the tutor module and the student begins with the student attempting a “preknowledge test” evaluated by the knowledge evaluator. This test helps the ViPS detect any misconceptions the students might have about pulley systems at the outset. After detecting and recording misconceptions that are present, the tutor module helps the student resolve these misconceptions by asking them to solve particular misconception-related problems for each detected misconception. Depending on whether the student solves these problems correctly (or not), tutoring for that misconception is not (is) provided, as explained below. If the student does not exhibit any of the six misconceptions listed in Table 1 at the outset, no problems or tutoring will be given to the student.

For each misconception detected by the preknowledge test, the tutor’s decision as to whether to tutor a student or not about that misconception depends on the student’s response to the problems specific to that misconception that he or she has been given to solve. For each problem, the student has to first enter a prediction (P), later his answer (A), and finally answer to a follow-up (FU) question. Based on these three answers, each of which could be correct (T) or wrong (F), the student’s performance on the problem is classified into one of the six categories R+, R, R−, W−, W, or W+ (see Fig. 5; e.g., R+ if all three answers are correct). The ViPS concludes that the student successfully solved a problem (marked T in Table 2) if the outcomes are R+, R−, or R, else it is concluded that the student failed to solve the problem (marked F in Table 2). The tutor module presents two problems per misconception, and a third problem depending on the outcomes of the first two problems (Table 2, rows 2-5), to determine whether a student indeed

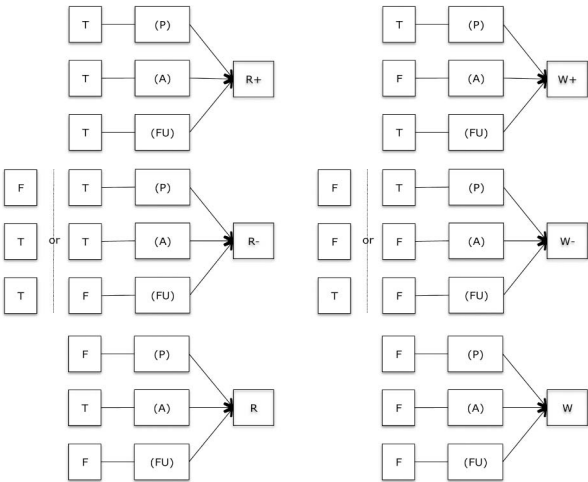


Fig. 5. Students’ problem solving performance classification.

or R, else it is concluded that the student failed to solve the problem (marked F in Table 2). The tutor module presents two problems per misconception, and a third problem depending on the outcomes of the first two problems (Table 2, rows 2-5), to determine whether a student indeed

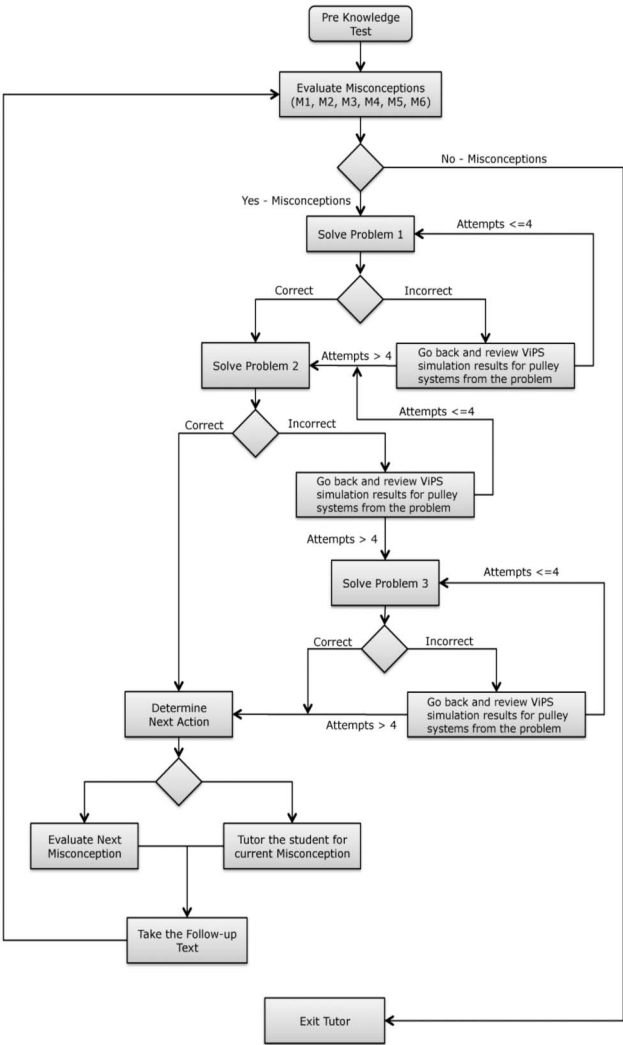


Fig. 6. The ViPS tutoring process flowchart.

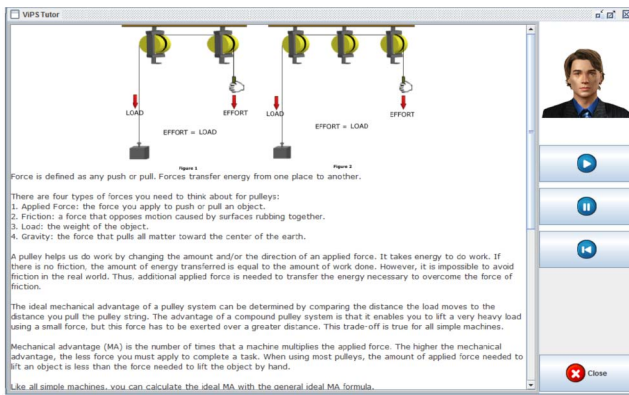


Fig. 7. The ViPS tutor.

has that particular misconception (detected from the preknowledge test). The problem outcomes are used to decide whether to tutor the student for that misconception or move on and evaluate the next misconception detected from the preknowledge test using another set of three problems (see Table 2).

For example, if the student solves the first two problems correctly, then she/he is determined not to have the corresponding misconception, so the tutor will move on to the next misconception (Table 2, row 1). The student has four attempts to solve each problem. If for some reason, the problem is not solved within the four attempts, the system assumes that the student does not have the necessary knowledge to solve the current problem and moves ahead to the next problem. If she/he solves the first problem correctly but errs in the second one, the tutor will present a third problem and depending on its outcome will either move to the next misconception (Table 2, row 2) or start tutoring actions (see Fig. 7) to remediate the current misconception (Table 2, row 3). After all the misconceptions are addressed by the problem solving and guidance sessions of the tutor module, the student can exit the tutor module with enhanced knowledge about pulley systems.

The tutor module produces hints during problem solving. For every problem, the students starts out by entering their prediction and later their answer. If the answer is correct then the tutor module will display a follow-up question to be answered, and depending on the student's answer to this follow-up question, will present another problem or initiate tutoring. The follow-up

TABLE 2
Tutor Action Decision Table

Problem 1	Problem 2	Problem 3	Action
T(R+, R-, R)	T(R+, R-, R)	N/A	Next Misc
T(R+, R-, R)	F(W+, W-, W)	T(R+, R-, R)	Next Misc
T(R+, R-, R)	F(W+, W-, W)	F(W+, W-, W)	Tutor Action
F(W+, W-, W)	T(R+, R-, R)	T(R+, R-, R)	Next Misc
F(W+, W-, W)	T(R+, R-, R)	F(W+, W-, W)	Tutor Action
F(W+, W-, W)	F(W+, W-, W)	N/A	Tutor Action

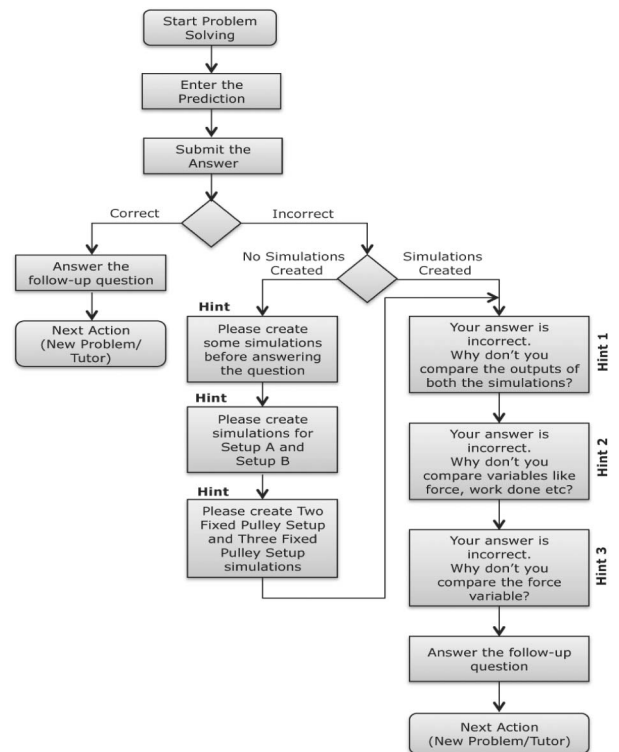


Fig. 8. Dynamic hint generation in the tutor module.

question is used to determine whether a student understood the physics concept behind the problem rather than simply guessing an answer. If the answer is wrong, the tutor module initiates a hint sequence by checking whether the student created the experimental setups and ran the simulations required to answer the current question. If the student did not create any simulations, the tutor presents a high-level hint ("please create some simulations before answering the question"). Thereafter, it starts to provide more specific hints. Fig. 8 illustrates this hint generation process for a particular problem that required the student to compare two fixed pulley setups called A and B. Note that the hints are worded politely, using words such as "please" and "why don't you" due to recent research [14] showing that polite prompts are more effective, especially for low prior knowledge students.

3.3 Knowledge Evaluator

When a student first initiates the ViPS, a preknowledge test, in the form of problems to solve (see Fig. 3), is given. Once the student finishes the test, his/her answers are evaluated by the knowledge evaluator to estimate the student's initial knowledge level and to identify the misconceptions he/she might have so that a subsequent sequence of problems can be generated for the student to solve in tutoring sessions. Thereafter, for each identified misconception, the student is given a set of problems to solve. Depending on his/her success or failure in solving these problems, he/she may be provided with a tutoring session designed to remediate the misconception. A follow-up knowledge test is given to the student after the completion of each tutoring session, and the answers are evaluated by the knowledge evaluator to determine the student's postknowledge level and the status

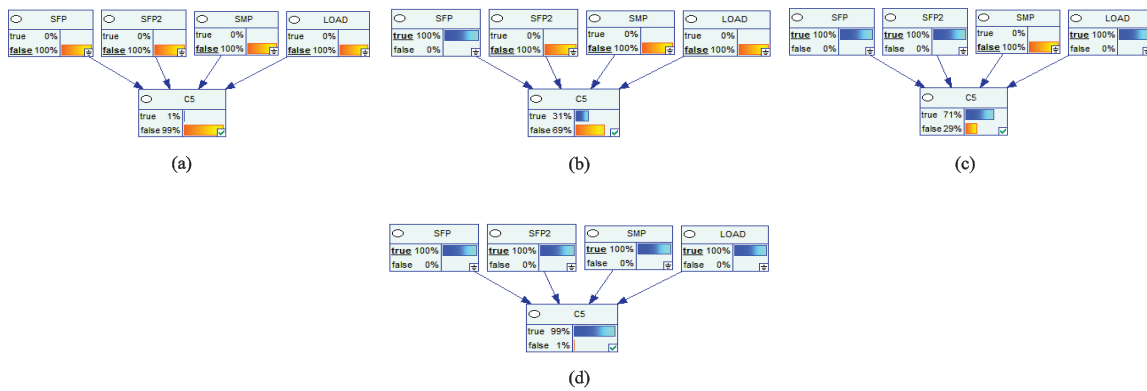


Fig. 9. Pulley setup inference using the Bayesian network

of the corresponding misconception. After the student exits the tutor module, a postknowledge test is given to evaluate the status of all detected misconceptions. The results from these are used by the knowledge evaluator to determine whether a student retained the knowledge acquired from tutoring through the end of the sessions.

3.4 Simulation Module

The simulation module is responsible for simulating the setups created by a student. In particular, it provides a platform for running simulations of setups that are difficult or impossible to create in the physical world, such as running a simulation with zero friction or running a simulation with quintuple pulleys. The outputs generated by the simulation include graphs and real time values of variables like force, work done, potential energy, friction, and mechanical advantage (see Fig. 3).

A student uses the simulation module to run the different pulley setups he/she creates during problem solving. The domain knowledge regarding possible or valid pulley setups is represented in the form of a Bayesian belief network. This network is used by the ViPS to 1) find all possible setups that can be created using components that an individual student has assembled on the work area, 2) find components for creating a valid setup that are missing from the work area, and 3) generate dynamic hints regarding pulley setups to help the student.

The ViPS generates all possible setups that the student may possibly have in mind, based on the components that the student created in the work area. This setup inference process is illustrated by the following example. Initially, the probability of, say, a single compound pulley (SCP) setup with extra pulley is zero (see Fig. 9a). If the student creates a single fixed pulley (SFP) in the work area, the probability of component SFP is updated to 1, and this results in an increase of the probability of setup SCP (see Fig. 9b). There is a further increase in the probability of SCP if the student then adds a second pulley and a load (SCP probability increases from 0.31 to 0.71) (see Fig. 9c). The probability of SCP increases to 0.99 upon the addition of a movable pulley to the existing setup by the student (see Fig. 9d). At this point, all (and only) the components needed for SCP have been assembled by the student in the work area.

It is possible that the components assembled by the student do not lead to a unique pulley setup, and instead

can be used to produce several possible setups. If this happens, the ViPS infers and displays a list of possible setups (see Fig. 11) based on the probabilities of creating each setup as determined by the Bayesian network, and ranked by an algorithm that we developed. This algorithm uses four attributes to rank order possible setups:

1. the number of components needed by a setup that are missing from the work area;
2. the number of grooves in each pulley in the setup;
3. the total number of components in the setup; and
4. the number of times this setup was created by the student previously.

Then the student is asked about which of these setups most closely matches his or her intention. Based on the students' selection, the simulation module generates dynamic hints to guide the student toward the completion of the intended setup in the work area.

As illustrated in Fig. 10, a student who knows exactly the setup he/she wants will not go through this process. The hints are beneficial to students who have an idea of what they want, but need some guidance. A student who does not know what he/she wants may try to game the system; however, since the student is engaging in this experimentation to answer questions posed by the system, it is likely that his/her answers will reveal a lack of knowledge, which will subsequently be remediated by tutoring.

The interaction between the simulation module and the student begins with the student clicking on the thread button after he/she has assembled the needed components in the work area.

This results in the simulation module evaluating the components to see if a unique setup can be constructed, as described in Section 3.4. If not, it initiates a dynamic hint sequence. The hints generated during the setup inference process of a double compound pulley setup are shown in Fig. 11. When the student initiates the process of threading with the components in the work area as shown in Fig. 11 (fixed double pulley, quadruple movable pulley, quintuple movable pulley, and weight), with which no unique setup is possible, the system first displays the most likely setup (single fixed pulley setup using double pulley) based on the highest probability computed by the Bayesian network and ranked using the rank-order algorithm (see Section 3.4). If the student does not select this setup, the system then displays the next possible setups as shown in Fig. 11. Let us

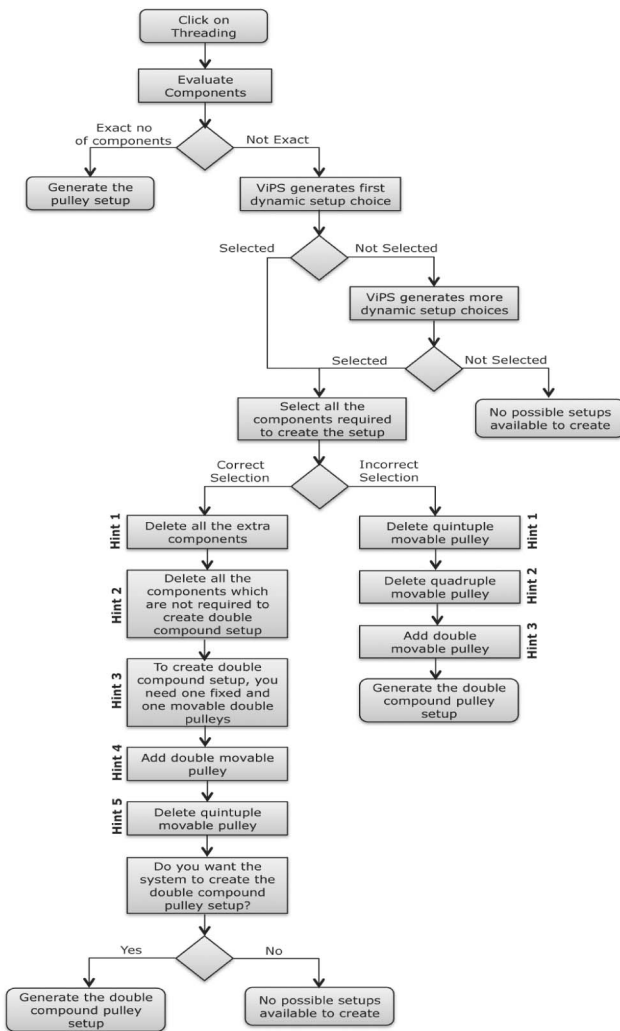


Fig. 10. Dynamic hint generation in a simulation module.

assume that the student selects the “double-compound pulley setup”. The system asks the student, through a dialog box, to select all the components required to create this setup. The student’s selection is used by the system to determine his/her current knowledge, so that it can tailor hints based on the student’s knowledge. If the student was successful in identifying all components required to create the double-compound pulley setup, the system starts the hint sequence at a higher level (e.g., “delete all the extra components”). If not, more specific hints (e.g., “delete the quintuple movable pulley”) are given. Fig. 10 illustrates this process.

3.5 Feedback Module

The feedback module is responsible for generating feedback messages for the students. The feedback provided in the ViPS is proactive [17] and incremental. It starts out with a generic feedback and then proceeds to more specific feedback, as the student progresses through setup creation or problem solving. The module produces the following four types of feedback:

1. The student creates a setup by dragging components onto the work area and clicking the thread button. If the ViPS determines that including all components

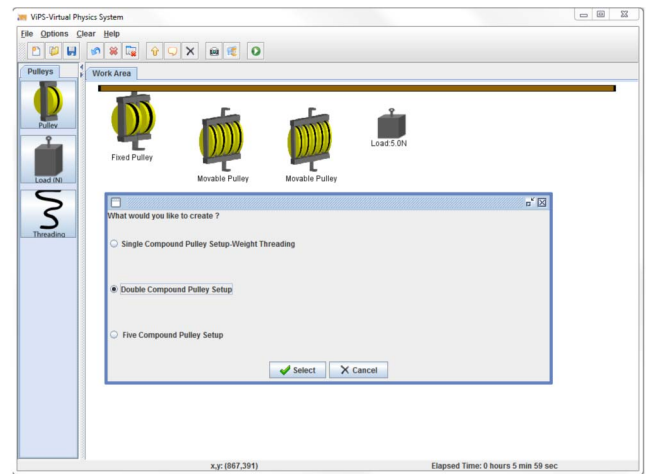


Fig. 11. Possible setups generated using the components in the work area.

in the work area will lead to impossible or invalid pulley setups, it generates a feedback known as setup feedback.

2. The student creates a set of valid components in the work area, but has no idea of what to do next, i.e., how to thread a string through the pulleys to complete the setup construction. In this case, the ViPS delivers feedback about the next moves the student has to make. This is known as threading hint feedback.
3. After creating and simulating one or more setups, the student submits his/her problem solutions. The system evaluates this and generates messages known as problem feedback.
4. The ViPS can coach students when needed during the process of problem solving, and this is known as problem hint feedback.

3.6 Student Module

The student model stores information that is specific to each student and enables the system to identify different students. It includes information about each individual student’s interactions with the system, pre- and post-knowledge levels and misconceptions (as identified from the pre- and post-tests), and the problem solving behavior of the student, which enables the system to understand the student’s current knowledge level. Modeling student’s knowledge within an ITS involves a good deal of inherent uncertainty [3]. Over the past decade, many researchers have proposed various approaches to address its difficulty, including fuzzy logic [10], rule-based systems [3], and Bayesian networks [4], [11]. In the ViPS, we used Bayesian networks to model student knowledge.

4 EVALUATION OF THE ViPS

We conducted evaluation studies of the ViPS at one university with 12 engineering majors enrolled in their first physics course, and at another university with 210 preservice elementary teachers enrolled in a physics course. Though ultimately the ViPS is intended for middle school use, our iterative design approach to the ViPS involves the following stages:

1. initial design;
2. usability test of the initial design with the target middle school population;
3. redesign;
4. evaluation with more advanced (i.e., college) students regarding the usefulness and usability of the system;
5. redesign; and
6. deployment in middle schools for further evaluation of usability and usefulness.

Myneni [14] provides details of the initial design (stage 1) and usability testing with middle school students (stage 2), which showed that the interface was usable, but also revealed problem areas that were then corrected in redesign (stage 3). Below we describe the evaluation studies of the ViPS that we conducted, focusing on answering the following research questions:

1. Do students learn from the ViPS?
2. Can ViPS remediate the misconceptions students might have?
3. Do building and experimenting with a physical pulley system followed by working with the ViPS or vice versa change the learning trajectory of students through the two experimentations?

4.1 Experimental Procedure

A total of 220 students, 12 engineering majors from the first university and 208 preservice elementary teachers from the second university, took part in the evaluation studies. For the first experiment (only 5), 12 participants from the first university and 50 from the second were assigned to the virtual-only condition in which participants constructed pulley systems using the ViPS and solved problems. However, data from two participants at the first and three participants at the second university could not be used for analysis because of gaps in collected data. For the second experiment, 158 participants from the second university were randomly assigned to two experimental conditions: 1) the physical-virtual (PV) condition in which participants worked in groups of two, first with physical pulleys and next with the ViPS to solve problems related to one misconception, and (2) the virtual-physical (VP) condition in which participants worked in groups of two, first with the ViPS and then with physical pulleys to solve problems related to the same misconception. All students answered a usability questionnaire at the end of their sessions, which was used to assess user satisfaction with the system.

Study procedure for virtual-only condition:

- *Pretest.* In a pre-test, the participants were asked to answer 18 questions related to pulley systems individually in the ViPS to measure their knowledge.
- *Problem solving.* The participants individually solved problems related to the misconceptions that were identified, and underwent tutoring, as needed using the ViPS.
- *Post-test.* In a post-test, the participants were asked to answer 18 questions (same questions as pre-test, but displayed in different order) related to pulley systems individually in ViPS to measure their

knowledge. The ViPS used this test to detect any remaining misconceptions.

- *Usability survey.* All participants were asked to fill out a usability survey individually to measure their overall satisfaction in using the ViPS.

Study procedure for physical-virtual and virtual-physical conditions:

- *Pretest.* In a pre-test, the participants were asked to answer 18 questions related to pulley systems individually on paper to measure their knowledge.
- *Group assignment.* Participants were paired and pairs were randomly assigned to either the PV group or the VP group.
- *Problem solving.* Each group solved problems related to one misconception (“the more pulleys there are in a setup, the easier it is to pull to lift a load”) using either actual pulleys or the ViPS depending on their assignment to the PV or VP condition.
- *Mid-test.* In a mid-test, the participants were asked to answer 18 questions related to pulley systems (the same questions as pre-test, but presented in different order) individually on paper to measure their knowledge after solving problems using either actual pulleys or the ViPS.
- *Problem solving.* Each group then solved problems related to the same misconception using either actual ViPS or pulleys depending on their assignment to the PV or VP condition.
- *Post-test.* In a post-test, all the participants were asked to answer 18 questions related to pulley systems (the same questions as pre-test and mid-test, but presented in a different order) individually on paper to measure their knowledge after solving problems using actual pulleys and then the simulation or vice versa.
- *Usability survey.* All participants were asked to fill out a usability survey individually to measure their overall satisfaction in using the ViPS.

4.2 Materials

Variations of a questionnaire with 18 questions similar to the one illustrated in Fig. 3 was used to evaluate participants’ initial, mid and final knowledge about pulleys and corresponding physics concepts such as force, energy, and work. These tests were also used to identify whether the participants exhibited any of the six misconceptions. Note that in the actual ViPS usage, it is the system that poses and evaluates the pre- and post-tests as a part of its tutoring strategy. The virtual-only study evaluated this. However, taking this approach in the PV and VP conditions would have meant that each participant might interact differently with the ViPS, because the ViPS adapts its tutorial strategy in response to the student performance in the pre-test. Therefore, to keep student interactions with the ViPS consistent across PV and VP conditions, in the second evaluation study these tests were presented on a paper and scored by the experimenter, and all participants solved problems related to the same misconception using actual pulleys and the ViPS in both conditions.

TABLE 3
Correlation between the Pretest Score and Learning Gain

	N	R	P	Std Beta
Virtual Only	57	0.66	0.001	-0.664
Virtual-Physical (VP)	80	0.66	0.001	-0.663
Physical-Virtual (PV)	78	0.62	0.001	-0.620

TABLE 4
Descriptive Statistics for Paired-Sample T-Test of Research Question 1 (Maximum Possible Score is 18)

	N	Mean	SD	SEM
Pretest Scores	57	4.57	3.21	0.42
Posttest Scores	57	13.71	2.98	0.39
Difference		9.14		

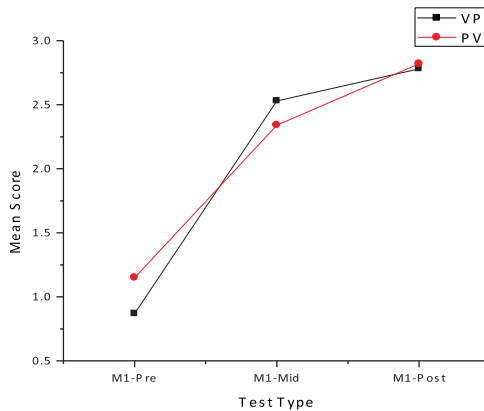


Fig. 12. Mean pre-, mid-, and post-test scores of VP and PV groups.

4.3 Data Analysis and Results

Research Question 1: Do students learn from the ViPS?

Students in all groups learned from their respective activities. Furthermore, lower prior knowledge students learned more. Linear regression found a significant negative correlation (see Table 3) between the pre-test score and learning gain in virtual-only, virtual (VP), and physical (PV) groups. It is not surprising that these correlations are strong as many of the students have low pre-test scores.

A paired-sample t-test was performed on the pre-to-post test scores of the students in the virtual-only group ($n = 57$) to evaluate their learning gain (see Table 4 for mean and SD values) after solving problems using the ViPS. There was an increase in pre-to-post test scores with statistical significance ($t(56) = -17.66, p = 0.001$). Scores went up by 300 percent from an average score of 4.57 in pre-test to 13.71 in post-test.

A repeated measures mixed analysis of variance (ANOVA) test was performed on pre-test to mid-test scores of the VP group and the PV group (158 students or 79 pairs in both groups solved problems related to the same misconception, but the first group used the ViPS, whereas the second used actual pulleys) to compare their learning gains. Results showed that the learning gain was higher (see Fig. 12 and Table 5 for test means) for the VP group that used the ViPS to solve problems first, with a

TABLE 5
Descriptive Statistics for Paired Samples T-Test of Research Question 2

	N	Mean	SD	SEM
Pretest Misconceptions	55	5.10	1.24	0.16
Posttest Misconceptions	55	1.76	1.13	0.15
Difference		-3.34		

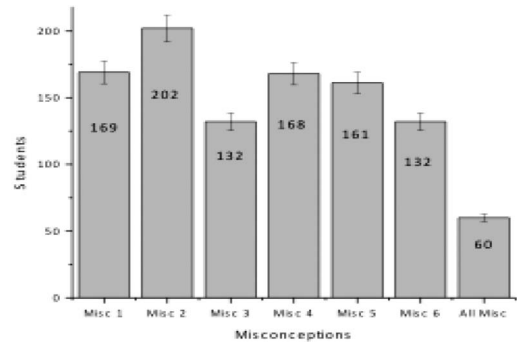


Fig. 13. Frequency of misconceptions.

statistically significant p value ($F(1, 156) = 4.54, p = 0.035, \eta^2 = 0.28$, and power = 0.563). This shows that students who used the ViPS to solve problems gained more knowledge than the students who used physical pulleys. These results answer the first research question in the affirmative.

Research Question 2: Can the ViPS remediate the misconceptions students might have?

A paired-sample t-test was conducted to compare the number of misconceptions identified in the pre-test to post-test (see Table 5 for pre- and post-misconceptions) in the virtual-only group. There was a significant reduction in the number of misconceptions from pre-test to post-test with statistical significance ($t(54) = 16.6, p = 0.001$). On average, each student exhibited five misconceptions after pre-test and two misconceptions after post-test. The number of misconceptions decreased significantly after working with the ViPS. These results indicate that the ViPS helped students learn about pulley systems and also remediated 60 percent of their misconceptions, and therefore answers our second research question in the affirmative.

Fig. 13 shows the detected frequency of each misconception. The most common misconception among all the students who participated in the evaluation experiments is Misconception 2 (see Table 1) followed by Misconception 1 and Misconception 4. Out of all the students, 60 exhibited all the six misconceptions. That misconceptions persist in college students is an interesting finding, given that these misconceptions point to a naïve understanding of physics concepts such as work, force, and energy, and that science curricula in middle through high school cover topics such as simple machines to help students better understand these concepts.

Research Question 3: Do building and experimenting with a physical pulley system followed by working with the ViPS or vice versa change the learning trajectory of students through the two experimentations?

TABLE 6
Average Pre, Mid and Post-Test Scores of VP
and PV Groups (Maximum Possible Score is 3)

	Virtual-Physical(VP)	Physical-Virtual(PV)
Pretest	0.87	1.15
Midtest	2.53	2.34
Posttest	2.78	2.82

A repeated measures mixed ANOVA test was performed on pre-test to mid-test scores of both the PV and VP groups to compare their learning gains.

Both groups solved problems related to only one misconception to maintain consistency across the two groups. Results showed that the learning gain (see Fig. 12) was higher in the VP group that used the ViPS to solve the problems, with a statistically significant p value ($F(1,156) = 4.54, p = 0.035, \eta^2 = 0.28$, and power = 0.563). Similarly, a repeated measures mixed ANOVA test was performed on mid-test to post-test scores, and we found that the students in the PV group who used the ViPS to solve the problems had a higher learning gain than those who used physical pulleys (see Fig. 12), but the gain was not statistically significant ($F(1,156) = 2.24, p = 0.137, \eta^2 = 0.014$, and power = 0.319).

Repeated measures mixed ANOVA was also performed on pre-test to post-test scores of both the groups to see the effect of VP or PV sequencing of experimental conditions on learning gain or conceptual understanding of students. We found that the pre-test to post-test increase in scores was significant in both the groups (see Table 6 for test means; note that only scores on three questions in the pre-, mid-, and post-tests related to the particular misconception addressed by the problems students solved in the two conditions are used in this analysis), but there was no statistically significant difference in learning gains between the groups. This shows that the sequence of experimental conditions (VP or PV) had no effect on the final level of students' conceptual understanding, i.e., both sequences improved their understanding by similar levels.

4.4 Analysis of the ViPS Usage

Interactions between the ViPS and the students were logged by the system. Results of analyzing these log files are presented below.

4.4.1 Number of Problems Solved

Linear regression found a significant positive correlation ($N = 57, R = 0.756, R^2 = 0.571, p = 0.03$, Standardized Beta = 0.792) between learning gain and number of problems solved in the virtual-only group. On average, each student solved eight problems while working with the ViPS tutor. The other two groups (PV and VP) were excluded from this analysis as they solved problems related to only one misconception (three problems).

4.4.2 Number of Simulations Created

Linear regression found a positive correlation between learning gain and number of simulations created, but the

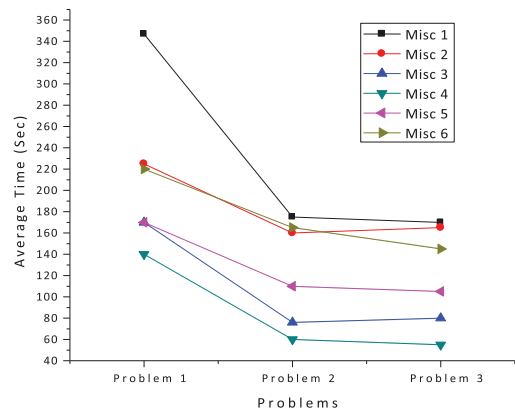


Fig. 14. Average time taken to solve three problems in each misconception category.

value of p is not statistically significant ($N = 57, R = 0.039, R^2 = 0.002, p = 0.830$). On average, each student created 14 simulations.

4.4.3 Problem Solving Efficiency

Fig. 14 shows the average time taken to solve the three problems in each misconception category using the ViPS (see Table 1). These data were obtained from system logs. A repeated measures ANOVA revealed an overall significant difference in the average time taken to solve the three problems while working with the ViPS ($F(1,140) = 9.1, p < 0.02$). The time required to solve a problem decreased significantly as students solved subsequent problems in the same misconception category. Though students took more time to solve the first problem in each misconception category, they became faster at solving subsequent problems related to that misconception.

5 CONCLUSION AND FUTURE WORK

In this paper, we presented an intelligent simulation and tutoring system called the ViPS for learning physics concepts through exploring a class of simple machines. The ViPS is innovative in several ways. First, the ViPS employs the coached problem solving approach [24] to detect and effectively tutor for common student misconceptions regarding physics concepts exemplified in the pulley systems. The ViPS is able to dynamically infer valid pulley setups from the components that a student selects and places on the workspace, and to adaptively generate hints based on student actions. Second, the ViPS is a new tool for virtually experimenting with—creating, exploring, and simulating—pulley setups, which are hard to build and manipulate in the physical world. Third, the graphical interface of the ViPS is designed according to the cognitive theory of multimedia learning [12] to help students connect abstract and difficult concepts of physics with representations at a more tangible level. Fourth, the ViPS brings together the concepts of virtual experimentation and intelligent tutoring in one platform. An evaluation of the ViPS was conducted at two institutions of higher education, where over 200 students worked with the system in multiple studies. Results from pre- and post-tests showed that the ViPS was effective in helping students learn and in

remediating their misconceptions. The system was shown to be easy and satisfying to work with, usable and useful, and more beneficial than working with real pulley setups.

An analysis performed on the log files generated during the ViPS evaluations supported these conclusions:

1. the less prior knowledge a student has, the more he or she learns from the ViPS;
2. amount of learning is directly related to the number of problems a student solves and the number of simulations he or she runs;
3. the more a student works with the ViPS, the faster he or she is able to solve problems; and
4. the ViPS is able to reduce the number of misconceptions students commonly exhibit.

It is interesting to note that we identified an average of five misconceptions in college student study participants, even though middle school curricula in physics are generally expected to address and remedy such misconceptions. Though the ViPS was successful in remedying many of these misconceptions in college students, this finding needs to be re-evaluated in middle schools.

We now discuss limitations of this research and some future directions. The system can be gamed by students so inclined; however, it has been acknowledged in the literature that this is a problem common to learning environments [2]. The system is not capable of diagnosing the full range of student difficulties with pulley systems. As currently designed, it diagnoses six common misconceptions with a pre-test with problems tailored to reveal specific student difficulties. The system is yet to be tested with the target population of middle school students. Though an earlier pilot study (not reported here) showed that middle school students found the system to be usable, large scale testing of the system in middle schools is to be done in the future. We also plan to extend the tutor's knowledge base and problem repertoire to cover more misconceptions. Another avenue is to enhance the system prompts to ask students to explain their actions and to process their free text responses to provide more nuanced guidance.

REFERENCES

- [1] J.R. Anderson and E. Skwarecki, "The Automated Tutoring of Introductory Computer Programming," *Comm. ACM*, vol. 29, pp. 842-849, 1986.
- [2] R. Baker, J. Walonoski, N. Heffernan, I. Roll, A. Corbett, and K. Koedinger, "Why Students Engage in 'Gaming the System' Behavior in Interactive Learning Environments," *J. Interactive Learning Research*, vol. 19, no. 2, pp. 185-224, 2008.
- [3] B.G. Buchanan and E.H. Shortliffe, *Rule-Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley, 1985.
- [4] C. Conati, A. Gertner, K. VanLehn, and M. Druzdzel, "On-Line Student Modeling for Coached Problem Solving Using Bayesian Networks," *Proc. Sixth Int'l Conf. User Modeling*, pp. 231-242, 1997.
- [5] T. De Jong and W.R. Van Joolingen, "Scientific Discovery Learning with Computer Simulations of Conceptual Domains," *Rev. Educational Research*, vol. 68, pp. 179-202, 1998.
- [6] H. Ebbinghaus, *Classics in Psychology, 1885: Memory*, R.H. Wozniak, ed., vol. 20, Thoemmes, 1998.
- [7] N.D. Finkelstein, W.K. Adams, C.J. Keller, P.B. Kohl, K.K. Perkins, and N.S. Podolefsky, "When Learning about the Real World is Better Done Virtually: A Study of Substituting Computer Simulations for Laboratory Equipment," *Physical Rev. Special Topics - Physics Education Research*, vol. 1, no. 1, p. 010103, 2005.
- [8] D. Fossati, B.D. Eugenio, C. Brown, S. Ohlsson, D. Cosejo, and L. Chen, "Supporting Computer Science Curriculum: Exploring and Learning Linked Lists with iList," *IEEE Trans. Learning Technologies*, vol. 2, no. 2, pp. 107-120, Apr.-June 2009.
- [9] A. Gertner and K. VanLehn, "Andes: A Coached Problem Solving Environment for Physics," *Proc. Fifth Int'l Conf. Intelligent Tutoring Systems*, pp. 133-142, 2000.
- [10] G.J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice-Hall, 1995.
- [11] R.J. Mislevy and D.H. Gitomer, "The Role of Probability-Based Inference in an Intelligent Tutoring System," *User-Modeling and User Adapted Interaction*, vol. 5, pp. 253-282, 1996.
- [12] R.E. Mayer, "Multimedia Learning: Are We Asking the Right Questions?" *Educational Psychologist*, vol. 32, pp. 1-19, 1997.
- [13] R.E. Mayer, *Multimedia Learning*. Cambridge Univ., 2009.
- [14] B.M. McLaren, K.E. DeLeeuw, and R.E. Mayer, "Polite Web-Based Intelligent Tutors: Can They Improve Learning in Classrooms?" *Computers and Education*, vol. 56, no. 3, pp. 574-584, 2011.
- [15] L.S. Myneni, "An Intelligent and Interactive Simulation and Tutoring Environment for Exploring and Learning Simple Machines," doctoral dissertation, Auburn Univ., 2011, unpublished.
- [16] J. Minstrell, "Student Thinking and Related Assessment: Creating a Facet-Based Learning Environment Chapter 3 in: Grading the Nation's Report Card: Research from the Evaluation of NAEP," *Nat'l Research Council, The Nat'l Academies*, pp. 44-73, 2000.
- [17] C. Murray and K. VanLehn, "A Comparison of Decision-Theoretic, Fixed-Policy and Random Tutorial Action Selection," *Proc. Eighth Int'l Conf. Intelligent Tutoring Systems*, pp. 116-123, 2006.
- [18] A. Papadimitriou, M. Grigoriadou, and G. Gyftodimos, "Interactive Problem Solving Support in the Adaptive Educational Hypermedia System MATHEMA," *IEEE Trans. Learning Technologies*, vol. 2, no. 2, pp. 93-106, Apr.-June 2009.
- [19] B.J. Reiser, J.R. Anderson, and R.G. Farrell, "Dynamic Student Modeling in an Intelligent Tutor for LISP Programming," *Proc. Int'l Joint Conf. Artificial Intelligence*, pp. 8-14, 1985.
- [20] S. Ritter, J. Anderson, K. Koedinger, and A. Corbett, "The Cognitive Tutor: Applied Research in Mathematics Education," *Psychonomics Bull. and Rev.*, vol. 14, no. 2, pp. 249-255, 2007.
- [21] J. Tan and G. Biswas, "Simulation-Based Game Learning Environments: Building and Sustaining a Fish Tank," *Proc. First IEEE Int'l Workshop Digital Game and Intelligent Toy Enhanced Learning*, pp. 73-80, 2007.
- [22] D. Tsovaltzi, N. Rummel, B.M. McLaren, N. Pinkwart, O. Scheuer, A. Harrer, and I. Braun, "Extending a Virtual Chemistry Laboratory with a Collaboration Script to Promote Conceptual Learning," *Int'l J. Technology Enhanced Learning*, vol. 2, pp. 91-110, 2010.
- [23] L.M. Triona, D. Klahr, and C. Williams, "Point or Click or Building by Hand: Comparing the Effects of Physical versus Virtual Materials on Middle School Students' Ability to Optimize an Engineering Design," *Proc. 27th Ann. Conf. Cognitive Science Soc.*, pp. 2202-2205, 2005.
- [24] K. VanLehn, "Conceptual and Meta Learning during Coached Problem Solving," *Proc. Third Int'l Conf. Intelligent Tutoring Systems*, pp. 29-47, 1996.
- [25] K. Van Lehn, C. Lynch, K. Schulze, J.A. Shapiro, R.H. Shelby, L. Taylor, D.J. Treacy, A. Weinstein, and M.C. Wintersgill, "The Andes Physics Tutoring System: Five Years of Evaluations," *Proc. Artificial Intelligence in Education Conf.*, 2005.
- [26] L.S. Vygotsky, *Mind and Society: The development of Higher Mental Processes*. Harvard Univ., 1978.
- [27] Z. Zacharia and O.R. Anderson, "The Effects of an Interactive Computer-Based Simulation Prior to Performing a Laboratory Inquiry-Based Experiment on Students' Conceptual Understanding of Physics," *Am. J. Physics*, vol. 71, no. 6, pp. 618-629, 2003.
- [28] Z.C. Zacharia, G. Olympiou, and M. Papaevripidou, "Effects of Experimenting with Physical and Virtual Manipulatives on Students' Conceptual Understanding in Heat and Temperature," *J. Research in Science Teaching*, vol. 45, pp. 1021-1035, 2008.



Lakshman S. Myneni received the MS and PhD degrees in computer science and software engineering from Auburn University in 2008 and 2011, respectively. He is currently with the Interactive and Intelligent System Research Laboratory, Department of Computer Science and Software Engineering, Auburn University. His research interests include learning technologies, intelligent tutoring systems, and user experience.



N. Hari Narayanan received the BE degree in electrical engineering from the Birla Institute of Technology and Science, the ME degree in automation from the Indian Institute of Science, the MS degree in computer science from the University of Rochester, and the PhD degree in computer and information science from Ohio State University. He is the John H. and Gail Watson Endowed chair professor of computer science and software engineering at Auburn

University, and directs Auburn's Intelligent and Interactive Systems Laboratory. His research interests include learning science and technology, cognitive science, human-computer interaction, and artificial intelligence.



N. Sanjay Rebello is currently an associate professor in the Department of Physics at Kansas State University. He has established a national reputation in physics education research. In addition to numerous publications and presentations in national and international conferences, he has received a number of grants for research in the teaching and learning of physics. In 2004, he was awarded the Presidential Early Career Award for Scientists

and Engineers. His research interests include teaching and learning of physics including the use of visualizations and problem solving.



Amy Rouinfar received the BS degree in physics from Florida State University in 2010. She is currently working toward the PhD degree in the area of physics education and is a graduate research assistant in the Physics Department at Kansas State University. Her research interests include visual attention and attentional cueing in physics problem solving.



Sadhana Puntambekar is currently a professor of learning sciences at the University of Wisconsin-Madison. Her expertise is in helping students learn from text and scaffolding students' learning of science.