COGNITIVE SCIENCE Vol 23 (3) 1999, pp. 305–336 ISSN 0364-0213
Copyright © 1999 Cognitive Science Society, Inc. All rights of reproduction in any form reserved.

Goals and Learning in Microworlds

CRAIG S. MILLER, JILL FAIN LEHMAN, KENNETH R. KOEDINGER

Carnegie Mellon University

We explored the consequences for learning through interaction with an educational microworld called Electric Field Hockey (EFH). Like many microworlds, EFH is intended to help students develop a qualitative understanding of the target domain, in this case, the physics of electrical interactions. Through the development and use of a computer model that learns to play EFH, we analyzed the knowledge the model acquired as it applied the game-oriented strategies we observed physics students using. Through learning-by-doing on the standard sequence of tasks, the model substantially improved its EFH playing ability; however, it did so without acquiring any new qualitative physics knowledge. This surprising result led to an experiment that compared students' use of EFH with *standard-goal* tasks against two alternative instructional conditions, *specific-path* and *no-goal*, each justified from a different learning theory. Students in the *standard-goal* condition learned less qualitative physics than did those in the two alternative conditions, which was consistent with the model. The implication for instructional practice is that careful selection and analysis of the tasks that frame microworld use is essential if these programs are to lead to the learning outcomes imagined for them. Theoretically, these results suggest a new interpretation for numerous empirical findings on the effectiveness of *no-goal* instructional tasks. The standing "reduced cognitive load" interpretation is contradicted by the success of the *specific-path* condition, and we offer an alternative knowledge-dependent interpretation.

I. INTRODUCTION

In the field of computer-aided instruction, highly interactive microworlds have gained importance as educational tools aimed at supporting learning through experience (Lawler, 1987; Schank & Farrel, 1988; Schauble, Glaser, Raghavan, Reiner, 1991b). In contrast to more traditional educational strategies that try to teach the target knowledge to the student directly, learning by exploration focuses on stimulating the student's initiative in gaining knowledge about the domain. Because microworlds both support exploration and behave according to the laws and constraints of the subject-matter domain, educators believe that student's activities in the microworld produce or foster education about the domain.

Direct all correspondence to: Craig Miller, School of CTI, DePaul University, 243 S. Wabash Ave, Chicago, IL 60604; E-Mail: cmiller@cs.depaul.edu.

Although such beliefs may be true in many cases, the learning outcomes achieved through microworld interaction depend largely on the surrounding instructional activities that structure the way students use and interact with microworlds. It should be no surprise that the activities students are prompted to perform in a microworld, and the goals they choose to pursue in light of these prompts, have a large impact on what is learned. However, all too often microworlds are promoted as being educational solutions in their own right, without qualification about the kinds of activities that should structure their use or direction about what activities will lead to which learning outcomes. Papert (1980) said that by using Logo children will have ‘‘mindstorms’’ and acquire ‘‘powerful ideas’’. That was the dream, but not the reality. Students do not learn powerful ideas from Logo (Pea & Kurland, 1984), unless the activity context is well engineered and targeted at well-defined learning objectives (Clements, 1986 and 1990; Klahr & Carver, 1988; Lehrer, Littlefield, Wottreng, 1991; Lehrer, Randle, Sancilio, 1989). The issue of activity context and the relation to educational objectives is even greater in educational games where there is the potential for students to be distracted by the game goals and, thus, not achieve the learning goals. On the other hand, it may be that the game motivates the student to do more or do better, and thus he or she learns more. This level of vague argumentation is common in the evaluation of educational software. What we proposed and demonstrated in this paper is the possibility of using a cognitive theory and computational learning model to derive, rather than vaguely argue, the probable learning outcomes of a complex interactive experience.

It is important to emphasize that this interactive experience is fundamentally a combination of the microworld interaction itself and the goals students choose to pursue as a result of the activities they are prompted to perform. The kinds of activities that can surround microworld use vary. Indeed, suggested activities can range from free exploration, where the only tasks required of the student are self-defined, to goal-driven exploration in which a set of tasks defined by the instructor (and often directly relevant to the test) delimits the students’ interactions. Several empirical studies suggest that training, in connection with a microworld-specific goal, can detract from the microworld’s pedagogically targeted objective. For example, Vollmeyer Burns, and Holyoak (1994) report that students who freely explored the effects of environmental parameters in an aquarium simulation acquired a better understanding of the simulation’s underlying properties than subjects who were given specific objectives. More generally, Sweller et al. reported empirical results indicating superior transfer from free exploration when compared to goal-specific problem solving across diverse tasks, including maze-tracing and number problems (Sweller & Levine, 1982), geometry problems (Sweller et al., 1983; Sweller, 1988) and kinematic problems (Sweller et al., 1983). It has been hypothesized that, theoretically, goal-driven problem solving has an inherent quality of impeding progress towards expert understanding (Sweller, 1988). This hypothesis thus stands in opposition to previous work that advocated the use of task-oriented interaction, touted for its apparent ability to keep the student focussed on the relevant aspects of the microworld (White, 1984).

Although the opposing viewpoints offer profoundly different prescriptions for microworld interaction, empirically they do not easily offer opposing predictions on which they can be evaluated. Indeed, just as an appropriately defined goal could require the student to focus on pedagogically relevant aspects, an inappropriate goal may require the student to focus on anything but what is relevant. Furthermore, such distraction may be particularly likely with a highly specialized goal, which allows for less overlap between the knowledge useful for the specialized goal and the pedagogically targeted knowledge.

It thus appears that progress towards reconciling these opposing viewpoints can come only by a detailed understanding of the microworld's goals and the knowledge necessary for achieving them. Although protocol analysis of students successfully completing a microworld goal can shed some light on the required knowledge, it does not ensure a complete, operational theory. For this reason, our methodology relies on the construction and analysis of a computer model of student learning and performance. We believe that the use of a computational model to explore the question of pedagogical efficacy has a number of important methodological strengths. First, it demands the clear specification of assumptions about what knowledge students have before the interaction and what strategies they use in their performance. Second, it makes explicit the assumed mechanisms of learning and permits observation of the process itself, making clear how and which prior knowledge contributes to the newly acquired knowledge. Finally, it can allow us to explore in a principled manner how changes to initial knowledge, strategy knowledge, or microworld features will effect what is learned.

The focus of our effort is Electric Field Hockey (EFH), an interactive microworld of electrical interaction. In particular, the underlying laws and constraints in the EFH environment simulate the motion of a free-floating, electrically charged particle under the influence of additional charged particles in fixed locations. The microworld is presented as a game in which students are given the goal of propelling the free-floating particle (the ''puck'') around obstacles and into a hockey net by fixing the position of additional charged particles on the playing field (see Figure 1). Thus, EFH falls somewhere in the middle of the continuum from free-exploration to test-driven interaction; the game does, by its very nature, define a specific goal for the student to achieve, but that goal is not, itself, the subject matter to be tested. Through the aid of a process model of student interaction and further empirical inquiry, our aim was to understand why the pursuit of a game-specific goal may impede learning; to ascertain whether learning is necessarily impeded with a specific goal; and if not, to qualify better the consequence of specific goals for microworld education.

Our presentation begins by describing Electric Field Hockey (EFH) and its pedagogical objective. Then, with the presentation of a model that interacts with EFH, we examine the knowledge needed for successfully pursuing EFH's game-like goals and, in particular, which microworld relationships need to be focussed on. Our analysis of the goal-specific knowledge in light of the pedagogical objective suggests that goal-oriented interaction will miss pedagogically targeted relationships. From this observation, we then formed hypotheses of how to change the interaction so that the student is more likely to focus on and acquire pedagogically relevant concepts. Finally, we present experimental results that

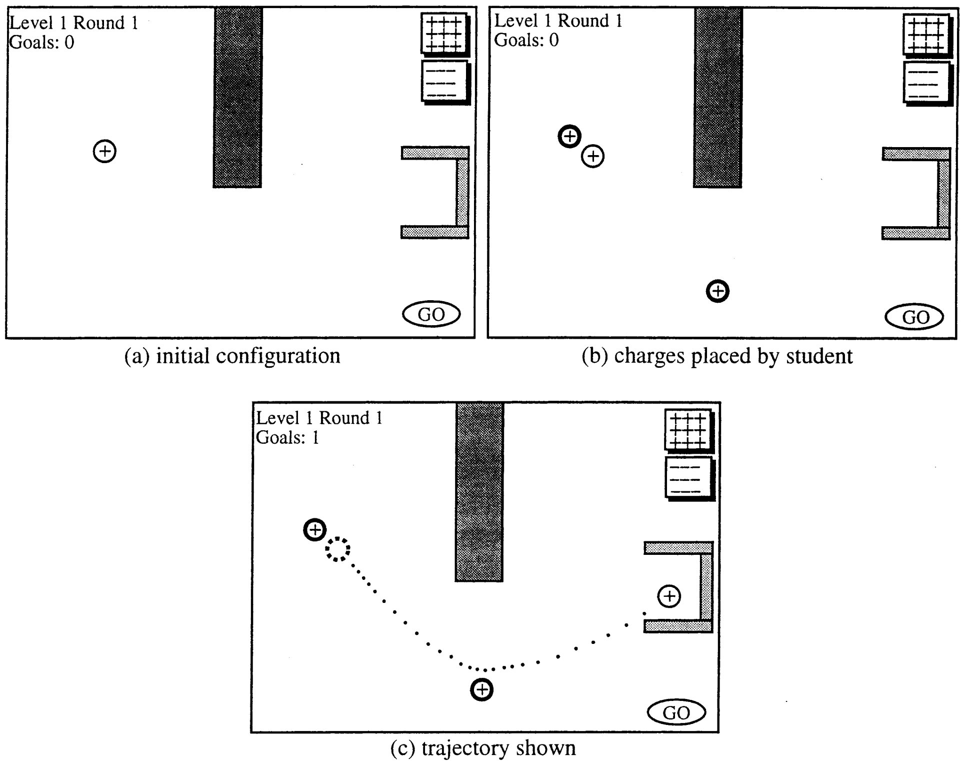


Figure 1. An example of simple interactions with Electric Field Hockey.

both support our hypotheses and qualify the generality of reducing goal specificity to improve learning.

II. THE INTERACTIVE MICROWORLD: ELECTRIC FIELD HOCKEY

Electric Field Hockey is an interactive, computer microworld whose pedagogical aim is to give physics students an intuitive feel for the qualitative interactions of electrically charged particles (Chabay & Sherwood, 1989, Sherwood & Chabay, 1991). The EFH microworld is capable of revealing the following pedagogically significant relationships: 1) the relationship between force and acceleration, 2) the relationship between an electrical charge's distance and its effective force (inverse square relationship), and 3) the relationship between the locations of multiple charges and their net effect (superposition).

As shown in Figure 1a, the student is presented a scenario with one charged particle, the puck, that becomes free floating as soon as the GO button is pushed. His or her goal is to propel and deflect the puck around obstacles and into the net located at the far right side of the screen. Before pushing the GO button, the student fixes additional charges, either positive or negative, to propel and deflect the free-floating charge along the desired trajectory (Figure 1b). Upon clicking the GO button, the original charge becomes free-

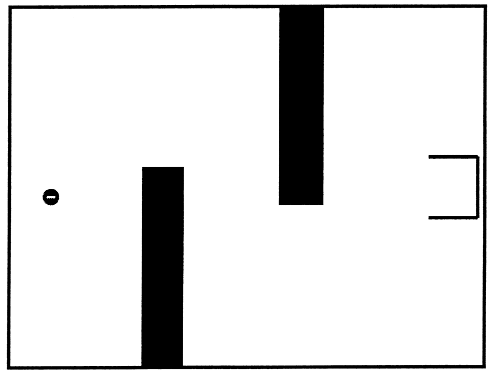


Figure 2. Example game situation at Level 5.

floating and follows a path in accordance with the forces exerted by the fixed charges. The microworld further illustrates the resulting path by placing a series of dots where the charge has traveled. Because each dot represents one unit of time, it is possible to infer relative velocities at various segments in the path; that is, a segment with dots spread apart indicates a faster velocity than a segment with dots closer together.

The student is free to add, remove, or adjust charges until the game's objective is achieved (Figure 1c). By staging more complex obstacles and sometimes an additional, unmovable charge, EFH offers six levels of increasingly difficult play. Figure 2 provides an example of a game situation at Level 5, whose solution requires the use of at least three charges.[1]

III. ASSUMPTIONS ABOUT KNOWLEDGE AND LEARNING IN THE MODEL

Our model, called EFH-Soar, is based on the observation of eight physics students playing EFH as part of the undergraduate electricity and magnetism course at Carnegie Mellon University. (All figures in this paper are snapshots automatically recorded during student interactions.) In constructing the model, we focussed on implementing playing strategies that are frequent, effective, and common to all the students. Note that our goal is not a comprehensive model of any particular student playing EFH. Instead, we seek an understanding of the playing strategies needed to meet the game's objective and the extent to which these strategies depend on pedagogically relevant relationships. Toward this end, we make a number of minimalist assumptions about the students' initial knowledge and the general learning process.

In delineating the knowledge required for the model, we assumed that the student sets him or herself the task of fulfilling the objective immediate to the game, namely, placing and adjusting charges so that the puck is propelled into the net. This is the minimal level of engagement required for working through the microworld, and the game-like environment encourages this goal-specific interaction (indeed, it may be why students find the microworld engaging and enjoyable to play). Clearly, it is possible for students to use strategies, such as self-explanation (Chi, Bassock, Lewis, Reimann, & Glaser, 1989) and deliberate hypothesis testing (Klahr & Dunbar, 1988; Lewis & Anderson, 1985; Simon & Lea, 1974), to engage in deliberate learning above and beyond the game's objective (although we see relatively little evidence for it in our protocols). However, such strategies would occur independent of the game's goal. Because our desire was to isolate the impact on learning of the microworld's goal per se, our modeling effort does not assume such strategies, but instead focuses on our notion of passive learning (described further below), which we believe is intrinsic to the problem-solving process.

Our model is implemented in Soar (Newell, 1990; Lehman, Laird, & Rosenbloom, 1996), and thus our theory rests on the learning assumptions of the Soar architecture, which sets the context for when learning occurs passively. In particular, our theory rests on the architectural assumptions that learning routinely occurs as a result of resolving subgoals during problem solving and that the acquired knowledge structure is indexed by the contextual cues used for that problem solving (Laird, Newell, & Rosenbloom, 1986). By assuming a non-deliberate learning approach, any episodic learning occurs only as a side-effect of achieving the immediate goal, which is placing and adjusting charges so that the puck is propelled into the net. From this assumption, we advance a theory for how and when successful and unsuccessful outcome episodes are stored.

The approach we have taken here is a revision of previous research and modeling work within the EFH domain (Conati & Lehman, 1993a and 1993b), which described a deliberate learning mechanism for explaining changes in student performance. Our current modeling effort departs from this previous work in two respects that are relevant to the conclusions we draw. First, our approach isolates the qualitative heuristics pertaining to the naive knowledge students typically have when they encounter the EFH microworld. This level of analysis enables us to specify the qualitative relationships that students use in pursuing the microworld goals. Second, our model limits learning to that derived from stated information needed in pursuing the microworld goal. Although our approach does not account for learning as a result of any deliberate hypothesis testing, it provides a principled understanding of what relationships are most likely to be acquired.

Because it is embedded in the Soar theory of cognition, EFH-Soar's passive learning assumptions contrast with Sweller's assertion (Sweller, 1988) that learning must generally compete for cognitive resources. Learning in Soar is a process that automatically occurs in the context of goal-oriented problem solving. In contrast, Sweller asserts that goal-oriented problem solving detracts from learning. Note, however, that the knowledge that Soar passively acquires as a consequence of problem solving is not guaranteed to be comprised of pedagogically relevant concepts, and, in fact, will not be if the goal knowledge does not depend upon such concepts.

IV. DESCRIPTION OF KNOWLEDGE, PROBLEM SOLVING, AND LEARNING IN THE MODEL

We now motivate the model's initial knowledge, show how it applies that knowledge to EFH's goal, and describe what new knowledge the model acquires in the process. By way

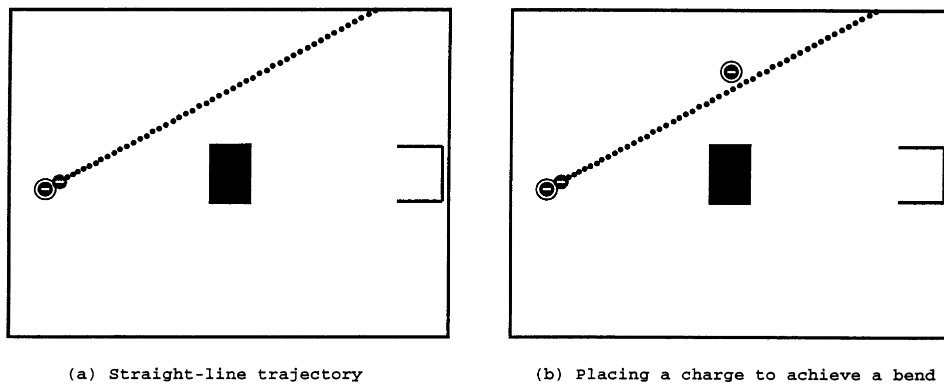


Figure 3. Typical placement strategies.

of example, we focus on a critical subtask that is needed throughout all levels of play, namely, the task of maneuvering the puck around an obstacle.

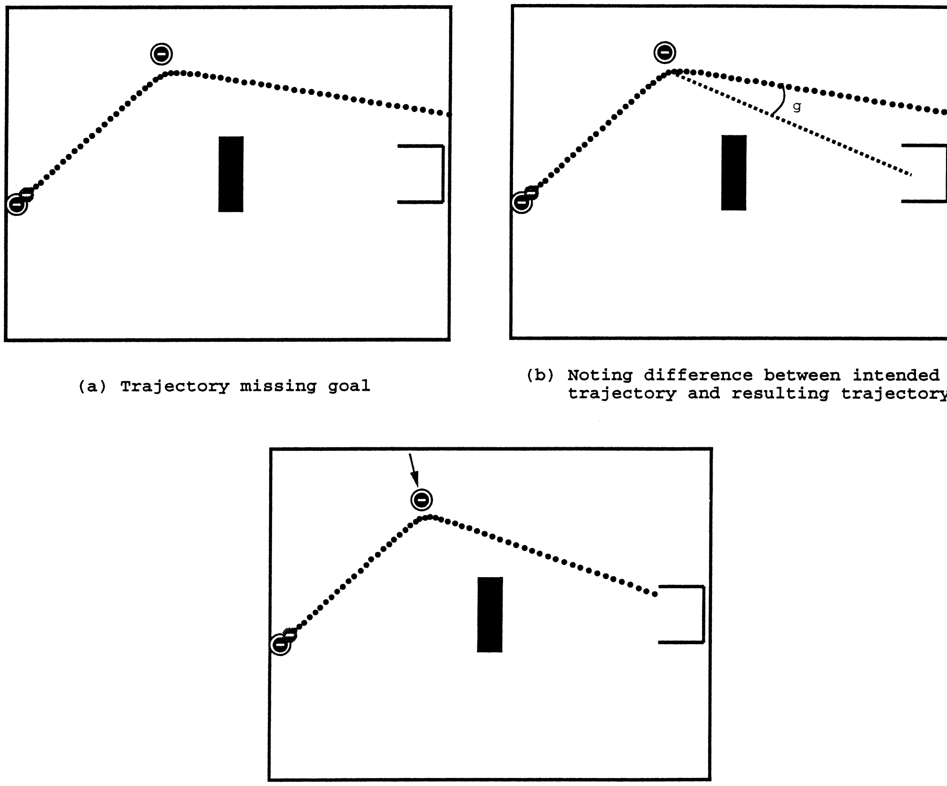
Our implementation initially has the following knowledge, which consists of a combination of some simple physics concepts of polarity and some commonplace, naive assumptions:

- A fixed particle with the same charge as the puck gives an initial push along the line that connects the charges, but in the opposite direction.
- The path of the free-floating charge will be deflected by like charges.
- Closer charges have more effect.

diSessa (1993) suggests that students often bring in a naive set of knowledge pieces, which he calls phenomenological primitives (p-prim). We believe that EFH-Soar's initial set of knowledge corresponds to a concrete set of p-prims for this domain. Our emphasis was to understand what the learning consequences are of applying these pieces of knowledge in the service of problem solving.

The goal of maneuvering around an obstacle can be broken down into two subgoals: achieving initial motion of the puck and bending the trajectory. Even upon first encountering EFH, our students seem to possess some basic knowledge about the qualitative properties of charges, for example, that like charges repel and opposite charges attract. Typically this knowledge is sufficient for achieving a straight puck trajectory by simply placing a like charge behind the non-moving puck (Figure 3a). All the observed students applied this strategy starting at the lowest level, as does our model.

EFH-Soar achieves a bending trajectory by placing a second like charge opposite of the intended bend's bisector (Figure 3b). Among the observed students, this was the most common strategy for effecting a bend in the trajectory, and every student used it at least once. A possible rationale, albeit naive, behind this relational charge placement could emanate from everyday experiences, such as observing a bouncing billiard ball or the



(a) Trajectory missing goal

(b) Noting difference between intended trajectory and resulting trajectory

(c) Moving charge closer for more bend

Figure 4. Game success through charge adjustment.

angle of objects reflected in a mirror. An alternate strategy, in which the student simply places the second charge opposite of the intended direction can be seen in Figure 4a. Once the free-moving particle is above the obstacle, the student anticipates that the propelling charge will suddenly ''push'' the particle towards the net. This strategy's underlying knowledge seems consonant with what diSessa (1993) calls the force-as-mover p-prim. Although it does not precisely deliver the intended effect (as shown in Figure 4a), appropriate follow-up adjustments can still lead to the goal, as demonstrated by both our student protocols and a variant of the model that used this placement strategy.

Although everyday schemas may provide a bias for placing the second charge relationally, they cannot provide the appropriate scaled distance. In terms of the example in Figure 3a, a student cannot know how far away to place the second charge from the envisioned trajectory, so, initially he or she must guess. It is only from interacting with EFH that he or she can induce the quantitative scaling knowledge required to play EFH efficiently because this scaling knowledge is a function of the game's parameters.

How does the student come to know just how far away to place the second charge? Presumably, he or she must internalize how placement differences correspond to outcomes that end in success or failure, and then access this knowledge, in one form or another, when placing charges in a new scenario. Under what conditions are these outcomes internalized? How general are they? What are the necessary cues for their future retrieval? And, most important, do they retain those features of the situation that will enable the student to induce pedagogically targeted concepts?

Before we address these questions, let us consider what the student could potentially learn from a placement and adjustment situation. First the student could learn some approximate distance knowledge; that is, how far away to place the second charge from the intended trajectory. The distance could even be an absolute figure, assuming that the incoming velocity and the desired angle of deflection do not significantly vary from trial to trial. In terms of pedagogically relevant knowledge, the student could notice that the greatest change in direction (acceleration) occurs at the point closest to the charge where the force is greatest. Even if the student does not encode this phenomenon in terms of the abstract concepts of force and acceleration, having noticed the event may serve in future understanding. With subsequent adjustments or by comparing the effect of adjustments over several trials, the student, by noticing that adjustments at great distances from the trajectory have less impact than equal adjustments at close distances to the trajectory, may also learn that the force of a charge rapidly decreases with distance.

By describing how EFH-Soar processes the microworld’s feedback and performs the appropriate adjustments, we hypothesized what a student actually notices and learns. We assumed that EFH-Soar has already placed two charges with the goal of propelling the puck over the obstacle and then back down into the net. The microworld produces the resulting trajectory (Figure 4a). How a student would perceive and process the microworld’s feedback is our immediate concern. With the feedback, the objective is now to adjust the current charges to reduce the difference between the current situation and the goal. Thus, EFH-Soar perceives the microworld’s feedback in terms of reducing this difference. In this case, by noting the direction of the angle formed by the intended trajectory and the actual trajectory (Angle g in Figure 4b), the model detects that the trajectory did not bend enough.

With a new situation, the model must engage in additional problem solving, which can encompass many strategies. The student is likely to draw upon his or her physics knowledge as well as knowledge of everyday phenomena. The model, by recasting the current goal in terms of achieving a greater bend in the trajectory, looks to see what will achieve a greater effect. Within this representation, the model possesses the knowledge that moving the charge closer to the trajectory will achieve this effect, a belief that is not only consistent with physics knowledge but also with a student’s everyday experience; that is, objects with a close proximity have more impact than objects farther away. Applying this knowledge achieves the result proposed by the action, moving the charge closer to the trajectory, as indicated by the arrow in Figure 4c.

For EFH-Soar, the proposal to move the charge closer completes a subgoal during problem solving. With the completion of the subgoal, the Soar architecture automatically creates a new knowledge structure in the form of a rule, called a chunk. As this new knowledge is a direct consequence of goal-oriented problem solving, the acquisition of the

```
IF charge is 5 units away from path vertex

   AND puck trajectory did not bend enough (wrt Angle g)

THEN move charge 1 unit closer to path
```

Figure 5. A chunk created from charge adjustment.

chunk is a passive process—no additional processing is required other than what is routinely supported by the architecture. Figure 5 presents the resulting chunk, simplified for explanatory purposes.[2]

The chunk's conditions contain two essential elements, namely, the distance of the charge from the planned path and the goal-oriented characterization of the resulting trajectory (i.e., the direction of the angle formed by the intended trajectory and the actual trajectory). These are elements that the model's subgoal knowledge depends on to infer the appropriate action. The Soar architecture automatically determines these conditions through a dependency analysis of the knowledge applied during problem solving in the subgoal (Laird et al., 1986). The chunk's action is the proposal to move the charge closer.

Acquiring the chunk facilitates future performance in two ways. First, should the situation described by the conditions arise again, the chunk immediately fires, thereby avoiding the time required for additional problem solving in one or more subgoals. In other words, performance of the task will show straightforward speedup from practice. Second, the chunk is a memory structure whose conditions encode an episodic result. EFH-Soar exploits this episodic knowledge by retrieving it when placing charges in future situations. The retrieval takes place by simulating, during problem solving, possible outcomes that may have occurred before. In this case, upon recreating the situation where a charge is placed 5 units away with a result of an overly shallow bend, EFH-Soar can detect that this event has occurred before through the successful match of the previously learned chunk. With this knowledge, the model can compensate by placing the charge closer in its initial placement, potentially eliminating the need for later adjustments altogether.

Thus, through simple problem solving, EFH-Soar acquires some distance knowledge that is useful for future goal-oriented performance. This improvement in performance corresponds to students' increasing ability to approximate a reasonable distance at which to place a charge with respect to the intended trajectory. In the next section, we discuss the knowledge structure's content in terms of pedagogically important relationships.

V. CLAIMS OF THE MODEL

To the extent that student learning is limited to the episodic knowledge acquired during problem solving, our model makes useful claims for the frequency and content of this knowledge. EFH-Soar acquires new episodic knowledge when adjusting charges. Adjust-

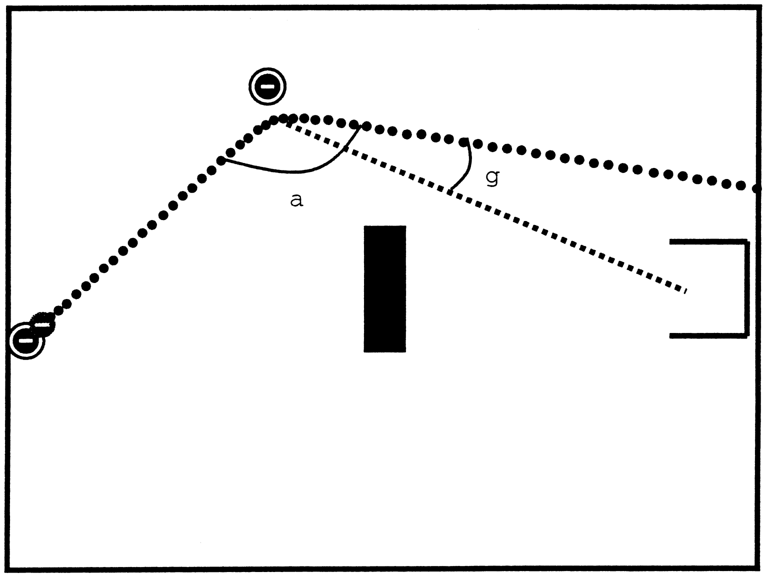


Figure 6. The goal-oriented angle (g) and the angle showing acceleration (a).

ing charges forces the model to attend to the microworld's feedback as well as the position of at least some of the charges and, as a product of Soar's learning mechanism, creates a new knowledge structure that encodes this episodic information. In its extreme form, the non-deliberate model suggests that no learning occurs when no adjustment is required, that is, when the goal is achieved on the first try.

By positing charge adjustment strategies, the model makes additional claims for the content of the episodic knowledge. The model's most useful strategy of moving the charge closer to achieve more bend requires the perception of the trajectory and the original position of the charge. These elements are included in the episodic memory structure shown in Figure 5. In particular, the trajectory is perceived in terms of Angle g, the difference between the intended trajectory and the resulting trajectory, as shown in Figure 6.

By comparing Angle g to the angle indicating the change of direction (Angle a), we see a discrepancy between the game-oriented and pedagogically relevant perspectives. The game-oriented perspective is embodied in the chunk shown in Figure 5, whereas the pedagogically relevant perspective considers the information present in Angle a independent of the goal of the game. From the latter perspective a student could potentially draw inferences between the position of the charge and the change in direction, encoding both 1) the extent of the change of direction in relation to the distance of the fixed charge from the trajectory, and 2) that the trajectory changes direction the greatest at the point closest to the fixed charge. Encoding these changes across similar contexts may ultimately reveal to the student that the force of the charge rapidly decreases with distance.

However, even with the pedagogical perspective, a comparison among similar contexts is not assured because the learned structure lacks additional relevant contextual informa-

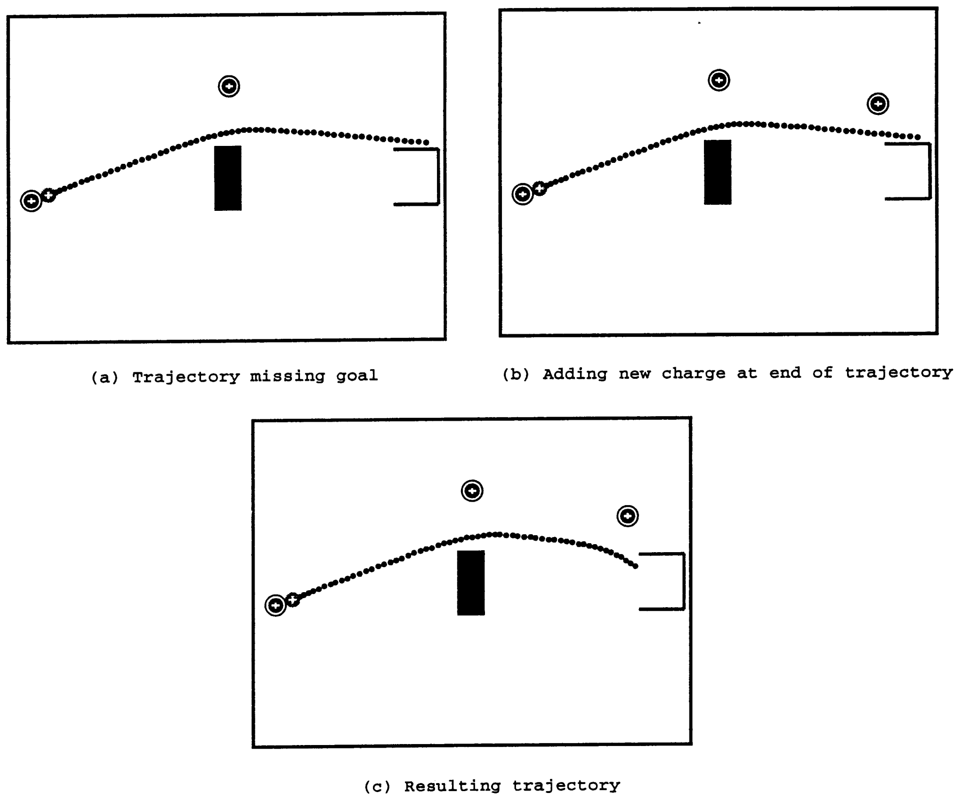


Figure 7. Game success by placing a charge near the end of the trajectory.

tion (e.g., velocity), and if attempted without this information may lead to erroneous conclusions. Moreover, the absence of velocity from the memory structure precludes the possibility of inducing any relationships between force and velocity.

Another adjustment strategy frequently practiced by students when the puck misses the net involves the placement of a new repelling charge near the end of the trajectory, as shown in Figure 7. Placing this charge requires only the perception of the small distance between the goal and the end of the trajectory. Because this strategy does not require any perception of previously placed charges, the model suggests that the microworld's feedback is not remembered within any useful context, in turn suggesting that the student learns no useful episodic knowledge when applying this strategy.

Note, also, that placing an additional repelling charge, even near the end of the trajectory, slightly alters the entire trajectory. Here the microworld potentially reveals the net effect of multiple charges (superposition). Yet, because the earlier portions of the trajectory are not involved in choosing this strategy, our model suggests that this difference will go unnoticed and unrecorded in episodic memory. A goal may still result, thus requiring no adjustment and no further opportunity for noticing the effect. Or, even if a goal does not result, adjusting the last charge according to the difference between the goal and the end of the trajectory does not require noticing changes that occurred earlier in the trajectory. Finally, a subtle change in the early stage of the trajectory, caused by a distantly placed charge, can cause a radical effect at a later stage. If the student fails to notice the early subtle change, he or she may mistakenly view the distant charge as the immediate cause of the radical change and, thus, perceive the change as a contradiction to the inverse square relation between distance and force.

Selecting adjustment strategies that reduce the difference between the resulting trajectory and the planned trajectory is, in effect, the application of means-ends analysis (MEA). Sweller (1988) has proposed a process account of how the application of MEA in pursuit of a specific goal detracts from learning the subject matter. In his problem-solving model MEA requires a greater cognitive load than simple exploration because it must maintain a stack of subgoals and their statuses. By assuming that both learning and subgoal management consume cognitive resources from a limited pool, he claims that MEA, in effect, inhibits learning.

EFH-Soar offers a contrasting theory. For Soar, learning occurs routinely and automatically during problem solving, during MEA, and otherwise. However, our analysis of the knowledge underlying the application of specific MEA operators (i.e., charge adjustment strategies) reveals that the state information necessary for understanding the domain can be ignored without jeopardizing game-level success. As a consequence, the knowledge structures that EFH-Soar acquires lack the episodic information for inducing pedagogically targeted relationships.

VI. PREDICTIONS OF THE MODEL

Based on our analysis in the previous section, we suggest that pursuing EFH's goal does not require it to notice or to use a number of pedagogically relevant relationships. If this claim is correct, then our model should suggest ways to modify interaction with EFH such that a student must re-attend the relationship between charge and trajectory and thus increase the likelihood of acquiring pedagogically relevant concepts. On the other hand, if our model has not adequately approximated student behavior and if EFH problem solving strongly supports relevant pedagogical learning, then we would fail to see significant improvements in learning with the prescribed alternatives. We now propose two alternative ways of interacting with EFH for which our analysis indicates increased likelihood of learning.

One alternate possibility for interaction is a *no-goal* situation; that is, a version of EFH without obstacles, net, and any specific task. Our model suggests that if students notice and acquire the charge and trajectory relationships that we have described, it comes from activity extrinsic to the achievement of the EFH goal. After successfully adjusting a charge, for example, a student could look back and see how the trajectory was affected, but the act is not required for achieving the game's goal. Further, the student could engage in additional placements and adjustments to directly test the microworld's properties. Time spent on achieving EFH's game-like goal could be better spent engaging in these activities; that is, EFH's goal may well distract students from pursuing useful interaction.

Consequently, we hypothesized that microworld interaction without any specified goal could provide more support for learning than interaction with the standard EFH goal. Our prediction is consistent with Sweller's cognitive load theory (Sweller, 1988), although ours is predicated on differences in knowledge acquired rather than cognitive processing load.

We will refer to interaction without any specified task as the *no-goal* condition. Here the term, ''no goal,'' is meant in the narrow, task-specific sense. By using this term, we are not suggesting that a student would not develop any self-imposed, internal goals. We use the term *standard-goal* for referring to the standard game interaction, where the student tries to position charges so that the ''puck'' avoids obstacles and goes into the hockey net.

Alternatively, our analysis allows for the possibility of changing the goal to require focussing on relevant relationships. In particular, we propose to specialize the goal further by illustrating a particular path (around EFH obstacles) and asking the student to arrange charged particles so that the moving charge will follow the illustrated path as closely as possible. This prediction is in direct contrast to Sweller's model, which argues that such specificity will detract from performance. We suggest, instead, that specialization of the goal will have two useful consequences:

1. Require the student to re-attend to the relationship between the charged particle and the trajectory.
2. Provide an additional reference point to which distance comparisons can be made.

Figure 8 illustrates the *specific-path* condition. Figure 8a shows the trajectory the student should try to achieve. Figure 8b shows a student's possible placement of charges and the trajectory that the student anticipates. Figure 8c includes the trajectory that resulted from the charge placement and the implemented modification to EFH that highlights the difference between the specified path and the actual path with connecting lines between alternating points.

The trajectory comparison highlights not only the post-bend difference, which students must notice to achieve the standard goal, but also the subtle difference in the trajectory before the bend. Without a trajectory-matching goal, students have little incentive to notice what happens to the trajectory before the bend. With the trajectory comparison, we suggest that students will attend to the entire change in direction, especially if they have the goal of matching the specified trajectory. This provides a necessary step to encoding the actual change in direction, a prerequisite for inducing both the distance and force relationship and the force and acceleration relationship.

The specified trajectory may also provide an additional reference point for noticing subtle changes early in the trajectory that are caused by a distantly placed charge. As subtle changes occurring early in the trajectory often produce radical differences in later stages, a student who notices the early, subtle difference should be less likely to be confused by how a distant charge can cause such a large effect. With this understanding, a student should be less likely to view the radical difference as a contradiction to the inverse square relationship between distance and force.

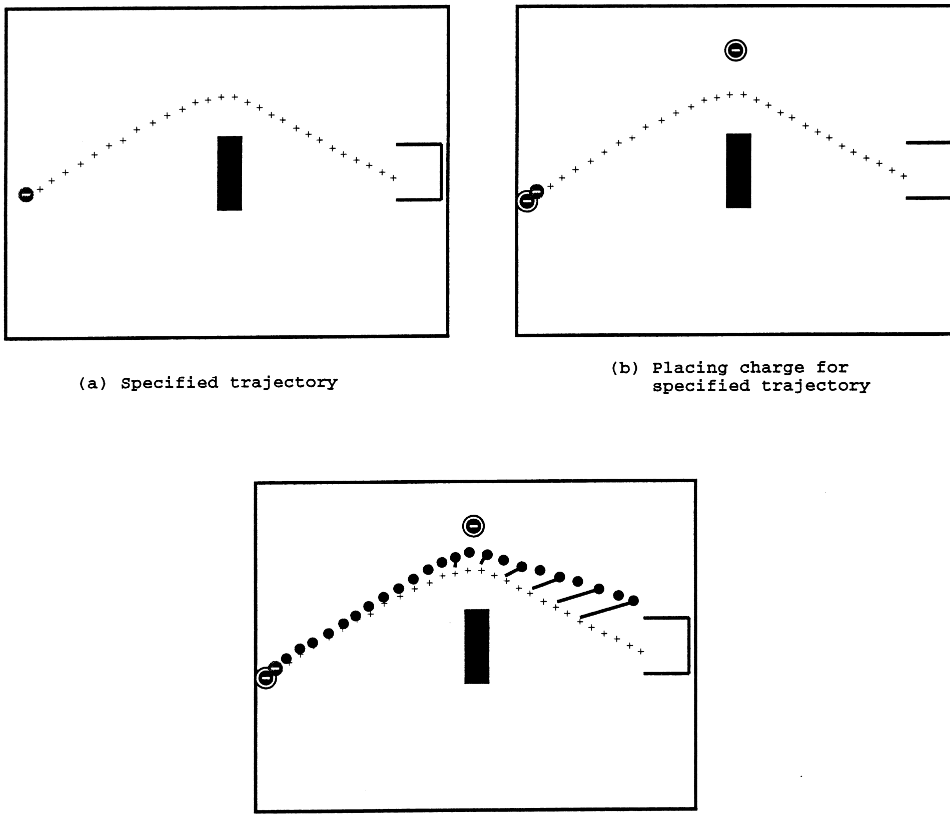


Figure 8. *Specific-path* condition: Specifying the intended path.

Figure 9 summarizes how the two contrasting theories would predict performance in the two modified conditions relative to the standard EFH environment. The first theory relies on our goal-dependent analysis, which indicates the extent to which the goal-based knowledge coincides with the pedagogical objective. We expected both of the proposed conditions to produce superior learning results—in terms of the targeted pedagogical objectives—compared to the standard EFH mode of interaction. This prediction contrasts with the hypothesis derived from Sweller's theory of goal-directed cognitive load (Sweller, 1988), which predicts that only the *no goal* situation will produce superior learning results. Under this view, the *specific-path* condition should produce inferior results. The *specific-path* situation requires a more complex set of constraints with both temporal and spatial interactions, which increase the use of MEA subgoals and, accordingly, add to the task's cognitive load. In the next section, we describe a study designed to test these predictions.

Predictions based on goal–directed transfer (Miller et al)

No goal > **Standard EFH Goal** < **Specific Path Goal**

No-goal condition is not "distracted" by having to focus on irrelevant goal-specific relationships of standard EFH play.

Specific path requires re-attending relationship of charge adjustment and trajectory change.

Predictions based on goal–directed cognitive load (Sweller et al)

No goal > **Standard EFH Goal** > **Specific Path Goal**

No-goal condition entails less cognitive load than standard condition which entails less load than the specific path condition because of increasing number of subgoals to be satisfied.

Figure 9. Predicted performance under two theories relative to standard EFH.

VII. TESTING THE PREDICTIONS: EXPERIMENTAL DESIGN AND RESULTS

Subjects

Subjects were 24 male and female undergraduate students, with diverse majors, from Carnegie Mellon University. They responded to an announcement posted on an electronic bulletin board. None had taken the electricity & magnetism (E&M) class in which EFH is used, nor had they previously interacted with EFH. They were paid $10 for their participation. Eight subjects were randomly assigned to each condition.

Materials

After a brief introduction, subjects interacted with one variant (described in more detail below) of the EFH program operating on a Unix workstation. For all students, the EFH program provided an introduction to the game, interactive instructions, and some simple example solutions. The program was instrumented to automatically record the subjects' moves.

A post-test, consisting of 17 problems and questions, was created to assess the students' acquisition of the microworld's pedagogically relevant properties. Most questions were taken from quizzes used in the E&M course; all were framed in terms of the EFH microworld. Questions were selected predominantly to test the subjects' understanding of the effect of distance, the effect of multiple charges (superposition), or the relationship between force and acceleration. Many questions presented an EFH scenario with a diagnostic arrangement of charges and asked the subject to indicate in which direction the free-floating charge would start to move. Other questions required the subject to supply a missing charge for a given trajectory. The post-test also included four true or false questions explicitly asking about the microworld's properties. The entire post-test is provided in the appendix.

Procedure

All Subjects. The computer program controlled the procedure and timing of the subjects' interaction. All subjects received the same introductory material, followed by additional instructions specific to their condition. For 30 min, subjects interacted with the version of EFH specific to their condition. After 30 min, subjects were required to stop and take the post-test. After completing the post-test, all subjects were given a final problem with obstacles and a net, for which they were given at most 15 min to find a solution.

No-goal Condition. For the 30-min time period, subjects in this condition interacted with an EFH microworld without obstacles or a net. After receiving the introductory instructions common to all conditions (which included some examples with an obstacle and goal), they were given the following instruction:

> Later, you will be given a situation with obstacles and a net.
>
> For the next 30 min, however, you are asked to ''experiment'' in a situation without any obstacles or a net. Your objective is to learn to understand the game's properties in any way you see fit.
>
> 30 min may seem like a long time, but you should try to continue experimenting as much as possible during this time.

Standard-goal Condition. For the 30-min time period, subjects in this condition played the standard version of EFH, which starts at Level 1—a situation with one small obstacle between the starting position and the goal—and becomes progressively more difficult with the use of more complicated obstacles and an occasional immovable charge. The interaction could proceed to Level 6. If a subject completed Level 6 before 30 min had elapsed, he or she was asked to continue playing new games at this level.

Specific-path *Condition.* For the 30-min time period, subjects in this condition had the same task as the *standard-goal* group, except that they were also given the desired solution trajectory, as shown and described in the previous section. They were given the following additional instructions accompanied by a working example:

> You will also be shown the trajectory path of a possible solution.
>
> You should try to arrange your charges so that the dots of the path you create match up with the dots of the shown solution.
>
> You will see lines connecting dots from the solution trajectory to your trajectory. This will give an idea of how well you have matched the solution trajectory.

Results

We found that students' performance on scientific tasks such as this one are often well predicted by their math SAT (mSAT) scores. Thus, we performed an analysis of covari-

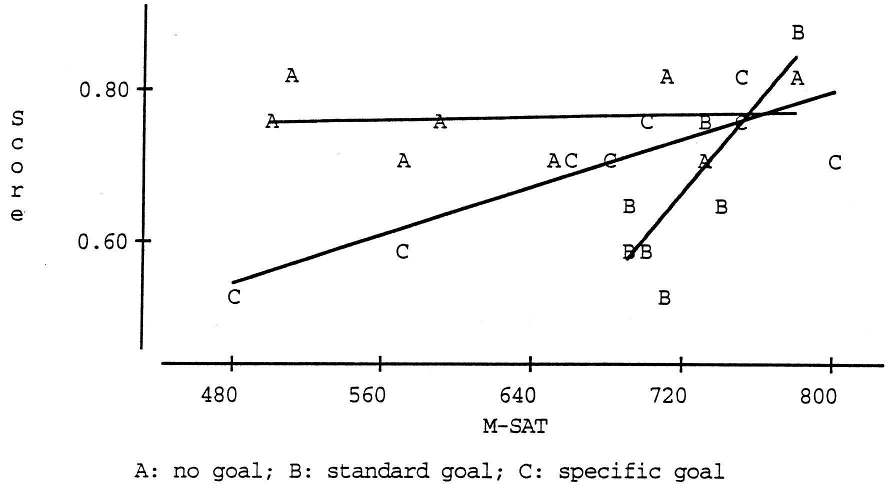


Figure 10. Score versus mSAT by goal condition.

ance with condition as a between-subjects factor and math SAT as the covariate. (One of the subjects had not taken the SAT and so was dropped from the analysis.) We found a significant effect of mSAT on student performance [$F(1, 17) = 22.39$; $p = .0002$], justifying its use as a covariate. Furthermore, there was a significant effect of condition [$F(2, 17) = 12.90$; $p = .0004$] and an aptitude-treatment interaction between the treatment conditions and mSAT [$F(2, 17) = 8.58$; $p = .0026$].

To better illustrate the aptitude-treatment interaction, Figure 10 shows a scatter plot of the post-test score versus the mSAT scores, broken down by the three goal conditions. Also depicted are the regression lines used in the full interaction model. The regression line for the *standard-goal* condition was particularly short because low mSAT scores were absent from this group. As shown, the aptitude-treatment interaction was largely a consequence of subjects in the *no-goal* condition doing uniformly well no matter what their mSAT, whereas in the other conditions higher mSAT subjects were more successful than lower mSAT subjects. In other words, the subgroup of students with the highest mSAT scores performed well independent of treatment, whereas those with lower mSAT scores showed significant differences as a result of experimental condition.

Using the full interaction model, we obtained estimated mean post-test scores at the average mSAT score (mSAT = 673). The *no-goal* condition produced the best (77%), the *specific-path* goal condition was next best (70%), and the *standard-goal* condition was considerably worse (52%).

For pair-wise comparisons at the average mSAT score, the Scheffe multiple comparison procedure indicated that both the *no-goal* and the *specific-path* conditions produced significantly higher scores than did the *standard-goal* condition. The difference between the *no-goal* and the *standard-goal* scores was 24.3%, with a 95% confidence interval for this difference ranging from 14.5% to 34.1%. The difference between the *specific-path* and the *standard-goal* scores was 17.495%, with a confidence interval ranging from 0.5%

TABLE 1
Pairwise Comparisons at the Average mSAT Score

Comparison	Difference (%)	Confidence interval (%)
No-goal > Standard-goal	24.3	$14.5 < \delta < 34.1$
Specific-path > Standard-goal	17.4	$0.5 < \delta < 34.3$
No-goal > Specific-path	6.9	$-7.2 < \delta < 21.0$

to 34.3%. The Scheffe procedure indicates no significant difference between the *no-goal* and the *specific-path* conditions (the 95% confidence interval was from −7.2% to 21.0%). Table 1 summarizes these comparisons.

DISCUSSION

EFH is intended to help students develop a qualitative understanding of the physics of electrical interactions. Earlier in the paper we described a computer model that learns to play EFH and presented an analysis of the knowledge the model acquired as it applied the game-oriented strategies we observed physics students using. The model assumes a non-deliberate learning approach based on the Soar unified theory of cognition, where episodic learning occurs only as a side effect of achieving the immediate goal. When given a sequence of standard tasks like those typically given to students, the model acquired new game-oriented skills and substantially improved its EFH playing ability. However, it did so without acquiring any new knowledge structures relevant to the targeted domain.

Using the model as a guide, we re-designed the standard tasks with a more specific goal that was hypothesized to lead to greater physics learning. In an experimental study we compared the *standard-goal* and *specific-path* approaches with a third *no-goal* condition. Students in the *standard-goal* condition generally learned less qualitative physics than those in the two alternative conditions, which was consistent with the model. The difference was particularly striking for students with moderate and lower aptitude as measured by their mSAT scores. Ostensibly, our assumptions of non-deliberate learning best apply to these cases. Important theoretical and instructional implications of our focus on task goals are discussed below.

Implications for Learning Theory: Knowledge Dependencies

Our model, embedded in the context of Soar's problem solving and learning mechanisms, argues for a knowledge-dependent analysis of microworld interaction. The knowledge-dependent interpretation suggests that what is important is not the presence or absence of externally defined goals, per se, but the relationship between the pedagogically targeted concepts and the knowledge required to interact successfully with the microworld, in whatever way that success is defined. Thus, goal-based problem solving will transfer to pedagogically relevant material exactly when the goal-dependent relationships coincide with pedagogically relevant relationships. This conclusion is supported by the *specific-*

path condition's predicted superior transfer to overall post-test performance relative to the *standard-goal* condition, and is consistent with a previous finding that reported the successful transfer of problem-dependent knowledge to related tasks (Morris, Bransford, & Franks 1977). The knowledge-dependent interpretation is also consistent with both the Soar and ACT (Anderson, 1993) unified theories of cognition, which have accumulated much other empirical support, as well as some other learning theories with pedagogical applications (VanLehn, Ohlsson, S., & Nason, 1994; Mertz, 1993).

The success of the *specific-path* condition relative to the standard condition stands in clear contradiction to Sweller's ''goals-hurt-learning'' interpretation (Sweller, 1988). On the other hand, the success of the *no-goal* condition is consistent with previous empirical findings by Sweller that have been interpreted as evidence against non-deliberate learning theories such as Soar. It is important, then, to understand how the knowledge-dependent view explains the success of the *no-goal* group, especially in light of the significant difference between *no-goal* and *specific-path* performance. Our original prediction that *no-goal* interaction would be superior to *standard-goal* interaction was based on the idea that students in the former condition would not be distracted by having to focus on irrelevant, goal-specific relationships inherent in standard EFH play; the group was more likely to spend its time on task focussing than on relevant relationships. Thus, although the prediction of the knowledge-dependent hypothesis does not differ from the prediction of a theory based on reduced cognitive load, the knowledge-dependent interpretation entails an additional test for confirmation: if the knowledge-dependent interpretation is warranted, then superior performance of the *no-goal* group should be differentially evident on those items of the assessment that share appropriate characteristics with the tasks students set for themselves.

Nine of the problems on the post-test asked students to make predictions about the interactions of a small number of charges, four were declarative true or false questions, and four were similar to EFH tasks (though designed particularly to assess qualitative physics principles). The charge interaction problems (questions 1 through 9) were similar to the kinds of tasks students in the *no-goal* condition were most likely to pose for themselves. It follows, then, from the knowledge-dependent hypothesis, that these questions should show the largest effect of the *no-goal* instruction. To test this hypothesis, we divided the post-test into two subscores, charge-interaction questions versus declarative+EFH questions, and performed another analysis of covariance, again with condition as a between-subjects factor and mSAT as a covariate, but this time adding question type as a within-subjects factor. As predicted, we found a significant interaction between condition and question type [$F(2, 22) = 3.76$; $p < .05$], whereby the difference between the *no-goal* condition and the other conditions was large for the charge interaction questions and small for the declarative+EFH questions. Table 2 shows the estimated means for the average mSAT score for the three conditions on the two subscores. *No-goal* students did 73% better than *standard-goal* students on the charge interaction questions, but only 22% better on the declarative+EFH questions.

We find further evidence of goal-specific transfer by examining more closely how subjects from the *standard-goal* and *specific-path* conditions did on particular post-test

TABLE 2
Post-Test Subscores Estimated at Average mSAT

	Charge interaction questions (n = 9)	Declarative + EFH questions (n = 8)
No-goal	.839	.685
Standard-goal	.484	.560
Specific-path	.721	.671

questions. We have suggested that the *specific-path* condition better supports the acquisition of the inverse square relation of force and distance. We also thought that the subjects in this condition were more likely to notice that the greatest change in direction (acceleration) occurs at the point nearest to the trajectory (acceleration-force relation). How well did subjects from the *standard-goal* and *specific-path* conditions do on test questions pertaining to these concepts? And how does this compare to other types of questions?

We targeted post-test questions for testing three previously discussed concepts. In particular, seven questions (numbers 3, 4, 6, 7, 8, 10a, and 10d) tested distance effects, two questions (numbers 11 and 13) tested the acceleration and force relation, and four questions (numbers 5, 9, 12, and 14) tested the net effect of charges (superposition). Table 3 presents the mean post-test score, broken down by condition and conceptual category. The table's results are estimated means at the average mSAT score. They suggest for which concepts the *specific-path* situation provided the best relative support. In particular, the results show that students in the *specific-path* condition fared, on average, approximately 18% better on the distance-related questions than did those in the standard condition. This is consistent with our prediction that the *specific-path* condition requires students to notice small, subtle changes when repositioning charges from afar. With fewer questions and higher variances, scores for the two remaining concepts were inconclusive. We did not anticipate seeing a difference in scores relating to superposition; yet a difference was revealed here. When we consider that the *specific-path* condition increases the likelihood that students notice subtle changes when repositioning the charges, they may notice equally well subtle changes caused by the additive effect of a new, distantly placed charge. If true, this would account for the higher superposition scores in the *specific-path* condition. We had anticipated a difference in the force and acceleration questions that favored the *specific-path* condition, which was marginally revealed here.

In contrast to the standard EFH model's lack of acquisition of pedagogically relevant concepts, our model does predict the acquisition of some game-relevant knowledge,

TABLE 3
Estimated Post-Test Means by Concept

Condition	Count	Distance (mean)	Acc*/Force (mean)	Superposition (mean)
Standard-goal	7	.439	.579	.515
Specific-path	8	.614	.593	.719

Note: Acc = Accelleration.

TABLE 4
Number Completing Each Level and Average Fit to Prototype Solution

	Standard-goal			Specific-path		
Level	*n*	Fit	*SD*	*n*	Fit	*SD*
1	8	13.7	4.8	8	15.2	5.5
2	8	18.1	5.8	7	15.6	9.0
3	7	27.1	12.4	7	18.7	11.1
4	7	20.0	13.3	6	13.5	5.1
5	6	52.9	17.1	4	28.6	19.1
6	4	50.1	4.7	2	33.4	24.7

namely, some absolute distance knowledge for roughly determining how far to place charges from an intended bend in the trajectory. Simulations in which the model played EFH demonstrate that, with some practice, it performs better at Level 5, in terms of requiring fewer adjustments to achieve the goal. Turning to the student data, both the *standard-goal* and *specific-path* conditions produced faster mean solution times (212 s and 290 s, respectively) than did the *no-goal* condition (326 s) for the Level 5 transfer task that followed the post-test. Although the differences were not significant, because of a large variance in the data, the differences are consistent with the claim that the *standard-goal* condition supports some game-relevant learning.

Unlike our goal-directed analysis, which predicts content-specific transfer, goal specificity makes broad learning claims. Indeed, the theory claims that goal-specific problem solving competes for cognitive resources to the detriment of learning, independent of the learning content. Our results suggest a limitation to using goal specificity to predict the extent of learning because the goal-specific condition produced post-test results superior to the *standard-goal* condition, and the *standard-goal* condition produced game-specific transfer superior to the *no-goal* condition.

Because our *specific-path* condition is ostensibly a specialization of the *standard-goal* condition, we have assumed that the *specific-path* task does, in fact, consume more cognitive resources than the *standard-goal* task. In providing additional support for this position, we took a further look at student performance in the *standard-goal* and *specific-path* conditions. The subjects in the *specific-path* group progressed through the microworld levels more slowly, on average, than the subjects in the *standard-goal* group (Table 4 indicates how many subjects of both conditions successfully completed each level). The *specific-path* goal, thus, seemed to be a more difficult goal, which we interpret as taking a larger cognitive load. Anecdotally, one subject noted after having played under the *specific-path* condition and then a *standard-goal* situation, ‘‘I didn’t really need dotted path lines. They were hard to follow anyway.’’

We were curious as to what extent the *specific-path* subjects followed the specified path, and whether their effort correlates to their post-test score. We predicted that subjects who followed the specified-path the most closely would do the best on the post-test. Because all results were automatically recorded by the computer, we were able to compare their actual goal-scoring trajectory with the specified trajectory. We used a comparison algorithm, which greedily searched for a minimal point-to-point mapping of the compared trajectories, and then averaged the distances between the paired points. This rendered a quantitative measure of how much the subject's worked-out solution fit the pre-specified prototype solution.

Table 4 lists the averaged measured fits for both conditions for each level. A smaller number represents a better fit. The table reveals that, for all levels except Level 1, the *specific-path* group produced better fits to the prototype solution than the *standard-goal* group, confirming that the *specific-path* group followed, to some extent, the instructions they were given. Both groups followed similar patterns by level. To some degree, the *standard-goal* group was also likely to follow the prototype path, as it is a minimal path that roundly clears the obstacles. For Levels 1–4, the fits were relatively good for both conditions. These levels are either easy to control with few charges (Levels 1 and 2) or the obstacles tightly constrain the path (Level 3 and 4). In contrast, Levels 5 and 6, which allow for more path freedom and require more charges, had more variance and worse fits on average for both conditions. In checking whether the variance could account for post-test performance, we found the Level 5 correlations between fit distance and post-test performance to be −0.83 and −0.85 for the *standard-goal* and *specific-path* conditions, respectively. Although the causality is speculative, these correlations are consistent with the position that subjects who try to follow a specific, minimal path are more likely to observe the pedagogically targeted relationships.

What seems unequivocal, at least, is that different goals lead to different learning outcomes. In the knowledge-dependent account of why this should be so we have focused on the detailed features of the microworld implicated in performance. To take a broader view of the issue, we note that Schauble, Klopfer, and Raghavan (1991) distinguish between students who used an engineering model of experimentation and those who used a science model. The behavior of the engineering group was characterized by manipulation of variables to produce a desired outcome, whereas the science group was characterized by broader exploration and more selectiveness in interpreting evidence, especially disconfirming evidence. In other words, the *no-goal* condition predisposed students to scientific modeling, whereas the *standard-goal* condition predisposed students to an engineering approach. We saw superior performance by the *no-goal* group on the scientifically-oriented post-test, which was consistent with Schauble et al.'s (1991b) findings. However, when it came to playing the game, the engineering group had the advantage.

Implications for Instruction: Appropriate Framing of Microworld Use

Perhaps the most practical result of this work is that careful selection and analysis of the tasks that frame microworld use is essential if such environments are to lead to the learning outcomes imagined for them. A simple self directed versus goal directed rule of thumb is not a panacea. In planning microworld use for the classroom, teachers need to ask how the strategies students employ in the learning activity are like or unlike those that fulfill the intended instructional objectives. If microworlds are to be effectively used, it is incumbent on microworld developers to make clear 1) what learning outcomes are to be expected from microworld use, ideally in the form of sample assessment items; and 2) what is the context for microworld use that will lead to such outcomes, ideally in the form of sample tasks or learning activities.

One common and faulty reasoning pattern in microworld use and, more generally in educational reform, is to reject careful consideration of learning outcomes and assessment tasks because ‘‘the traditional tests just don’t measure what we’re after.’’ One can easily see that, in the EFH domain, traditional tests of quantitative electrical relationships (e.g., find acceleration of the particle given it is . . .) are not appropriate. However, to go further and assume that because a microworld does not address quantitative objectives it must lead to qualitative ones, is unwarranted. As we have shown, it is quite possible, even for college students, to get caught up in the gaming aspects of a microworld and acquire strategies that are peculiar to the game, but do not result in qualitative knowledge applicable outside the microworld. Certainly the entertainment aspects of microworlds can play an important motivational role in their educational impact, but this must be supported with careful selection of tasks for microworld use; teacher awareness of the potential for non-productive, game-oriented strategies; and careful monitoring and assessment of student strategies and learning outcomes. The justification for the prediction that task A leads to learning outcome B can be found in a process model that establishes how the knowledge required to do A is reorganized by the actual experience of doing A to produce B. In other words, the point is not to ‘‘teach to the test’’ but, rather to teach to the process.

Although the *no-goal* condition fared the best among our three conditions, this does not imply that this condition provides the best learning support among all possible conditions. Our choice of conditions was partly motivated by the desire to test the effect of goal-directed cognitive load and we thus forewent testing alternate conditions whose goal specificity was ambiguous relative to the other conditions. We suspect a more structured interaction would provide a more efficient vehicle for supporting learning. We reach this conclusion by noting that a large percentage of the *no-goal* interaction was ostensibly wasted through the construction of complex charge configurations, too complex to discern the abstract properties of the microworld. By encouraging simpler configurations and explicitly emphasizing relevant microworld relationships, a more productive use of time would likely result.

Finally, we would like to emphasize that our results qualify rather than deny the usefulness of goal specificity. Using goal specificity has its merits in that it is generally simple to apply. An extensive analysis of the domain is not required. It has correctly predicted the learning merits of certain goal conditions in previous work, and for two pairs of conditions here (*no-goal* versus *standard-goal* and *no-goal* versus *specific-path*), has correctly predicted which condition better supports learning. However, our conclusion suggests that goal specificity is limited in that it does not consider the relationship between knowledge used in service of the goal and the pedagogical objective, and thus does not account for learning that occurs non-deliberately during problem solving. Because an analysis of the goal-directed knowledge does account for this learning, we can use it to determine a specific goal for microworld activity that will support the learning of its pedagogically targeted concepts.

Although goal specificity may serve as a useful heuristic for estimating learning potential within a given microworld, its theoretical underpinning as determined by cognitive load could be called into question. Its theory is not only inconsistent with the results presented here, but also with previous results reporting the successful transfer of problem-dependent knowledge to related tasks. Furthermore, it lies in direct opposition with leading unified theories of cognition, where learning and problem solving are tightly integrated. Despite the inconsistencies, the phenomena supporting cognitive load theory cannot be ignored and thus require an account of how problem solving may impede learning in the context of these integrated learning theories. The theoretical work provided in this paper takes a step in this direction by describing how learning can be ubiquitously present during problem solving, yet fail to have the appropriate form and content for pedagogically relevant transfer.

Acknowledgments: We thank the administrators of Carnegie Mellon University's User Studies Lab for the use of their facilities acquired through the National Science Foundation Equipment Grant #9022511. We thank Clayton Lewis, Wayne Gray, Mark Guzdial, and one anonymous reviewer for their comments on earlier drafts. Finally, we thank Allan Rossman for his considerable assistance in preparing and reviewing our statistical results. This research was sponsored in part by the Markle Foundation. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Markle Foundation.

NOTES

1. For this and subsequent game figures, we will omit some of the game's details, such as the GO button and the charge boxes.
2. The unit for the distance measure is arbitrary. Actual numbers were kept to illustrate that the distances in the representation are constants.
3. Level 6 has the same obstacle configuration as Level 5 but demands the use of fewer charges.

REFERENCES

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chabay, R., & Sherwood, B. (1989). *The electricity project and the cT programming language*. (Technical Report 89–10). Pittsburgh, PA: Carnegie Mellon University, Center for Design of Educational Computing.

Chi, M. T. H., Bassock, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanation: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145–182.

Clements, D. H. (1986). Effects of Logo and CAI environments on cognition and creativity. *Journal of Educational Psychology, 78,* 309–318.

Clements, D. H. (1990). Metacomponential development in a Logo programming environment. *Journal of Educational Psychology, 82,* 141–149.

Conati, C., & Lehman, J. F. (1993a). EFH-Soar: Modeling education in highly interactive environments. In P. Torasso (Ed.), *Advances in artificial intelligence, lecture notes in artificial intelligence, LNCS 728* (pp. 47–58). Berlin: Springer-Verlag.

Conati, C., & Lehman, J. F. (1993b). Toward a model of student education in microworlds. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 353–358). Hillsdale, NJ: Lawrence Erlbaum Associates.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction, 10,* 105–225.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12,* 1–55.

Klahr, D., & Carver, S. M. (1988). Cognitive objectives in a Logo debugging curriculum: Instruction, learning, and transfer. *Cognitive Psychology, 20,* 362–404.

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1986). Chunking in soar: The anatomy of a general learning mechanism. *Machine Learning, 1,* 11–46.

Lawler, R. (1982). Designing computer based microworlds. *Byte, 7,* 138–146.

Lehman, J. F., Laird, J. E., & Rosenbloom, P. S. (1996). A gentle introduction to soar, an architecture for human cognition. In D. Scarborough & S. Sternberg (Eds.), *Invitation to cognitive science* (Vol. 4., pp. 211–253). Cambridge, MA: MIT Press.

Lehrer, R., Littlefield, J., & Wottreng, B. (1991). *Seeding Mindstorms with LogoWriter: Using LOGO in the elementary classroom.* Fontana, WI: Interactive Education Technologies.

Lehrer, R., Randle, L., & Sancilio, L. (1989). Learning preproof geometry with LOGO. *Cognition and Instruction, 6,* 159–184.

Lewis, M., & Anderson, J. (1985). Discrimination of operator schemata in problem solving: Learning from examples. *Cognitive Psychology, 17,* 26–65.

Mertz, J. S. (1993). Using a simulated student for instructional design. In J. Greer (Ed.), *Proceedings of the 7th World Conference on Artificial Intelligence in Education* (pp. 453–460). Charlottesville, VA: Association for the Advancement of Computing in Education.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519–533.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas.* New York: Basic Books.

Pea, R. R., & Kurland, D. M. (1984). On the cognitive effects of learning computer programming. *New Ideas in Psychology, 2*(2), 137–168.

Schank, R. C., & Farrel, R. (1988). Creativity in education: A standard for computer-based teaching. *Machine-Mediated Learning, 2,* 175–194.

Schauble, L., Glaser, R., Raghavan, K., & Reiner, M. (1991a). Causal models and experimentation strategies in scientific reasoning. *The Journal of the Learning Sciences, 2,* 201–238.

Schauble, L., Klopfer, L. E., & Raghavan, K. (1991b). Students' transition from an engineering model to a science model of experimentation. *Journal of Research in Science Teaching, 28*(9), 859–882.

Sherwood, B., & Chabay, R. (1991). Electrical interactions and the atomic structure of matter: Adding qualitative reasoning to a calculus-based electricity and magnetism course. In M. Caillot (Ed.), *Proceedings of the NATO Advanced Research Workshop on Learning Electricity or Electronics with Advanced Educational Technology* (pp. 23–35). Berlin: Springer-Verlag.

Simon, H. A., & Lea, G. (1974). Problem solving and rule induction: A unified view. In L. W. Gregg (Ed.), *Knowledge and cognition* (pp.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Simon, T. (1987). Claims for Logo—What should we believe and why? In J. C. Rutkowska & C. Crook (Eds.), *Computers, cognition, and development* (pp. 115–133). New York: John Wiley and Sons.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12,* 257–285.

Sweller, J., & Levine, M. (1982). Effects of goal specificity on means-ends analysis and learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 463–474.

Sweller, J., Mawer, R., & Ward, M. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General, 112,* 639–661.

VanLehn, K., Ohlsson, S., & Nason, R. (1994). Applications of simulated students: An exploration. *Journal of Artificial Intelligence in Education, 5,* 135–175.

Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1994). The impact of goal specificity and systematicity of strategies on the acquisition of problem structure. In A. Ram & K. Eiselt (Eds.), *Proceedings of the 16th Annual Meeting of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.

White, B. Y. (1984). Designing computer games to help physics students understand Newton's laws of motion. *Cognition and Instruction, 1,* 69–108.

APPENDIX

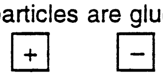
The following figures are the pages of the post-test that were administered to students upon their completion of one of the three goal conditions.

Electric Field Hockey Problems

The following pages have problems for you to solve. Based upon your experience of interacting with Electric Field Hockey, please provide, as best as you can, your answer to each question.

For all problems:

Squared charged particles are glued down, and cannot move:



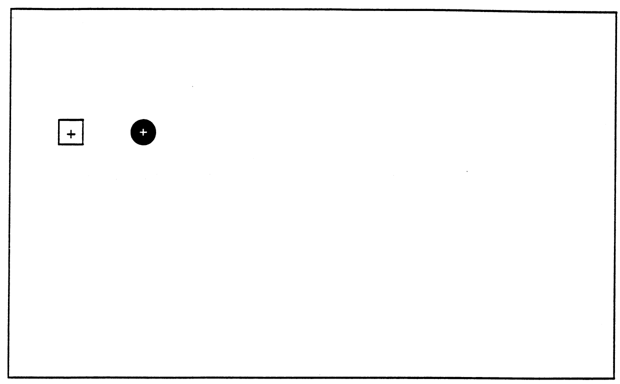
Round charged particles move freely without friction. They always start from rest:

Charges with two signs (plus or minus) are equivalent to two charges in the same place:

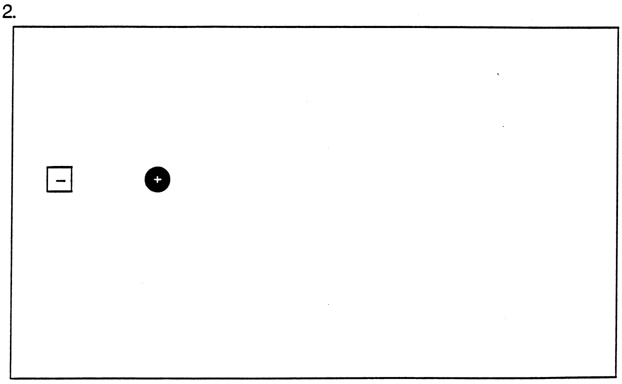
Some problems show a scaled ruler in the diagram, which will inform you of the distance between charges. The distances between the scale's segments are always equal:

For problems 1-6, draw an arrow showing the direction in which the round particle will begin to move. If the round particle won't move, write "won't move."

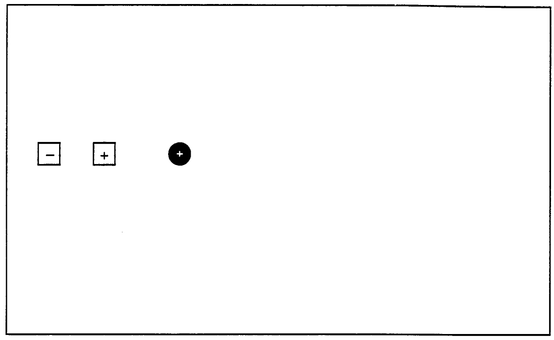
1.



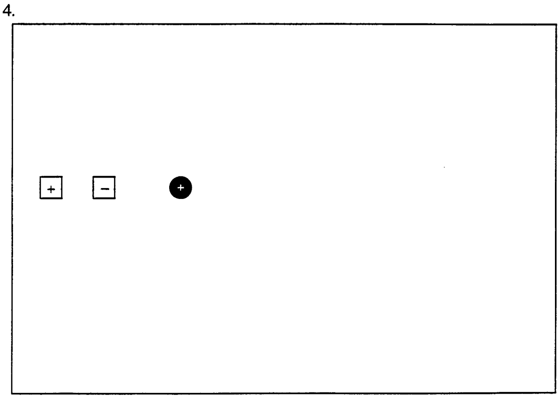
2.



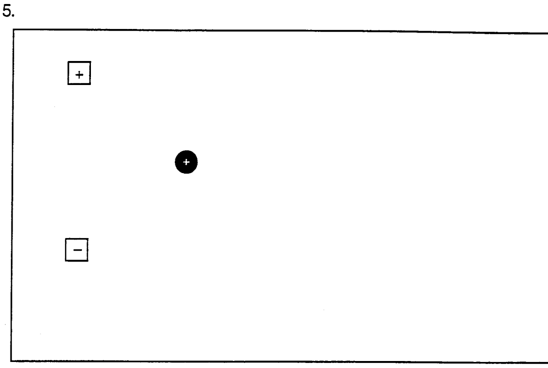
3.



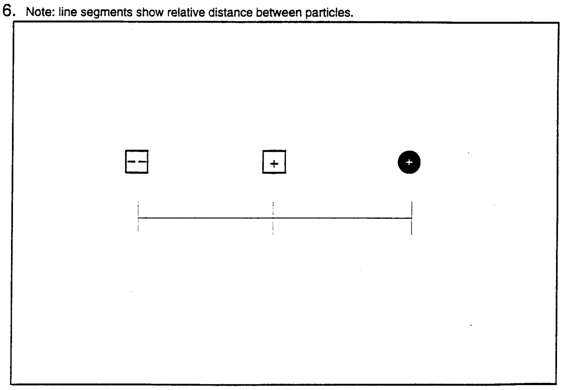
4.



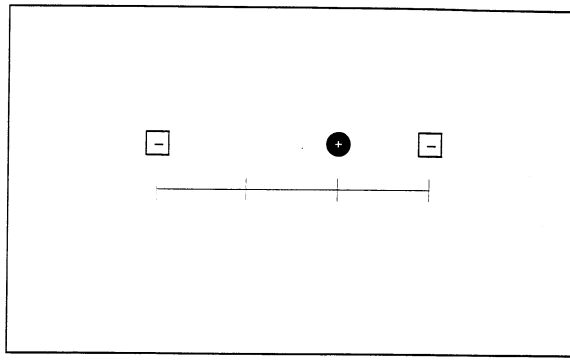
5.



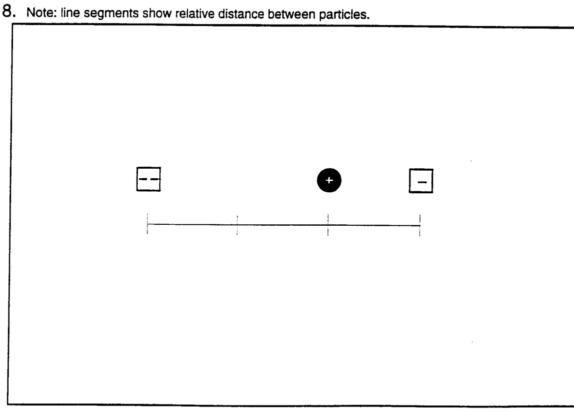
6. Note: line segments show relative distance between particles.



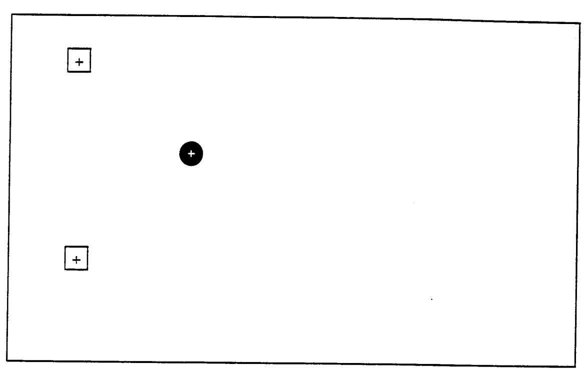
7. Note: line segments show relative distance between particles.



8. Note: line segments show relative distance between particles.



9.



10. Circle TRUE or FALSE

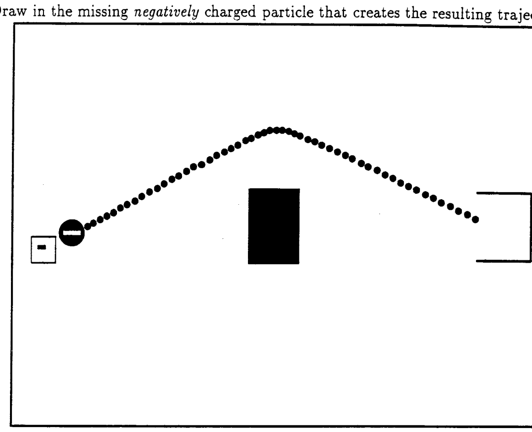
a) TRUE FALSE If the distance between two charged objects is doubled, the force on each object becomes half as large.

b) TRUE FALSE Charged objects exert forces on each other only if there is nothing between them.

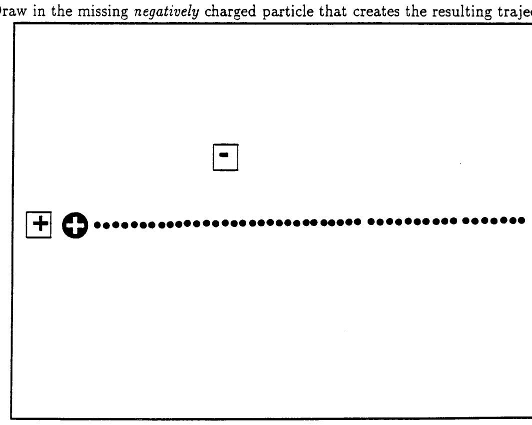
c) TRUE FALSE Like charges repel.

d) TRUE FALSE If two charged objects are sufficiently far apart, the force between them is exactly zero.

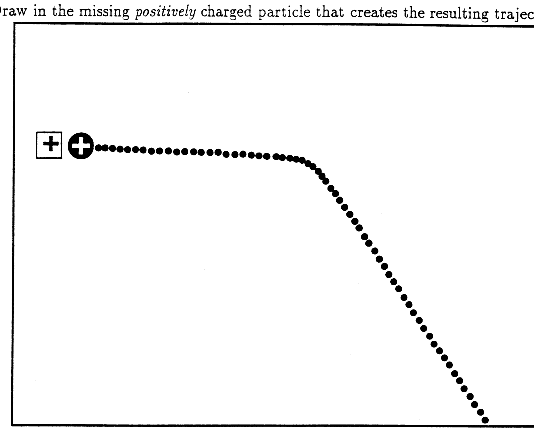
11. Draw in the missing *negatively* charged particle that creates the resulting trajectory.



12. Draw in the missing *negatively* charged particle that creates the resulting trajectory.



13. Draw in the missing *positively* charged particle that creates the resulting trajectory.



14. Draw in the missing *positively* charged particle that creates the resulting trajectory.