

## **Prospective Adaptation in the Use of External Representations**

In press, Cognition and Instruction

Lee Martin

University of California, Davis

Daniel L. Schwartz

Stanford University

Running Head: Prospective Adaptation

Corresponding Author:  
Lee Martin  
School of Education  
1 Shields Avenue  
Davis, CA 95616

## Abstract

An important element of adaptive expertise involves stepping away from a routine to retool one's knowledge or environment. The current study investigated two forms of this adaptive pattern: fault-driven adaptations, which are reactions to a difficulty, and prospective adaptations, which are proactive reformulations. Graduate and undergraduate students with no medical training engaged in a medical diagnosis task that involved complex information management. The graduate students, who were relative experts in information management and data analysis, uniformly made prospective adaptations by taking the time to create external representations of the available information before they diagnosed a single patient. In contrast, the undergraduate students only made representations reactively, when experimental manipulations made their default behaviors impractical. Graduate students tolerated the time lost creating representations in favor of future benefits – well-structured representations led to more optimal diagnostic choices. Overall, the results indicate that long-term educational experiences are correlated with prospective adaptation, even in a novel task domain that is not a specific part of those educational experiences. This research provides new metrics for evaluating educational interventions designed to move students along a trajectory toward adaptive expertise.

The ability to adapt is at a high premium for jobs and futures that involve changing circumstances (Augustine, 2005). Even so, there has been relatively limited work on the characteristics of adaptive behaviors or the conditions that lead to them. In part, this may be due to American psychology's emphasis on immediately efficient behaviors (Schwartz, Bransford, & Sears, 2005). Problem solving performance, which would seem to be at the heart of adaptation, is typically evaluated by speed and accuracy. These are measures of efficiency, whereas adaptation may involve letting go of short-term efficiencies to take the time to learn or develop new ways to accomplish an activity. In this paper, we demonstrate one class of adaptive behaviors: namely, the stepping away from case-by-case problem solving to set the stage for handling a class of problems.

Prior to the empirical demonstration, we open the discussion with a review of Hatano's seminal construct of adaptive expertise (e.g., Hatano & Inagaki, 1986). We distinguish two types of adaptive behaviors: fault driven adaptations, which occur in response to an impediment and seek to address an immediate need, and prospective adaptations, which occur in the absence of immediate external pressures and attempt to anticipate future activity. Our particular focus is on people's spontaneous introduction of visual representations to manage complex information.

### Routine and Adaptive Expertise

Hatano and colleagues differentiated routine and adaptive experts (Hatano & Inagaki, 1986; Hatano, Miyake, & Binks, 1977; Hatano & Osawa, 1983). Routine experts possess a high degree of procedural efficiency. Their primary example involved abacus masters who, through years of practice, had developed an internal simulation of the abacus. Using their mental abacus, the masters could perform amazing feats of mental arithmetic such as summing ten 10-digit numbers with a mere two second delay between each. The abacus masters were clearly experts,

yet at the same time, their understanding was narrow and inflexible. Their competence was restricted to a small set of arithmetic tasks, and they did not seek new contexts in which to apply or extend their skills.

Hatano contrasted the routine expertise of the abacus masters with adaptive expertise. Adaptive experts can (1) verbalize the principles underlying their skills, (2) judge conventional and non-conventional versions of the skills as appropriate, and (3) modify or invent skills according to local constraints. When considering Hatano's distinction, it is useful to note that adaptive and routine expertise are tied to specific activities rather than types of people. For instance, a person who exhibits adaptive expertise for reading literature may employ routine expertise for word decoding.<sup>1</sup>

Hatano's distinction suggests an important and educationally relevant question: how do "novices become adaptive experts – performing procedural skills efficiently, but also understanding the meaning and nature of their objects" (Hatano & Inagaki, 1986). One critical aspect of adaptive expertise is the ability to let go of existing routines to try something new. In the absence of a specific motivation to adapt or be innovative, people tend to get by if they have a set of reasonably effective skills. Luchins and Luchins' (1959) classic studies of einstellung, or rigidity of behavior, illustrate this tendency. People were taught a method for measuring water using several jars in sequence. Once they had mastered the complex, multi-jar method, they continued to use it for simple cases that could be solved with only one or two jars.

People often approach novel problems or environments as if they were old, familiar ones. In the early days of computers, typists treated word processors like typewriters. They easily

---

<sup>1</sup> It is also worth noting that the distinction between routine and adaptive forms of expertise does not imply that any particular group of previously studied experts, such as the medical and chess experts so well studied in the expertise literature, are routine experts. On the contrary, these experts are most likely adaptive experts, as their respective areas of expertise require extensive and on-going learning (see Ericsson, 2006 for reviews).

learned functions already available on a typewriter, such as entering text and repositioning the cursor, but many did not explore to learn more advanced functions, such as cut and paste, unless they received tasks or instructions that explicitly demanded those functions (Lewis & Mack, 1982; Sander & Richard, 1997). There are good reasons for this. Routines are effective: once established, they are relatively low cost and low risk. As such, people often use what they already know, especially when they are unaware of the possibilities, or when learning interferes with engaging the activity itself, as in the case of reading instructions before playing a new videogame.

A challenge for taking a trajectory towards adaptive expertise is that letting go of good-enough routines can create a temporary “implementation dip” that comes with change (Fullan, 1993; Strauss, 1982). It can be necessary to sacrifice short-term efficiency to retool one’s knowledge or context for the prospect of long-term gain. Figure 1 depicts two hypothesized trajectories for the development of proficiency in a task.<sup>2</sup> The dots in the figure represent points in time where people accomplish instances or components of a recurrent task. In the routine pattern, people begin work on a task immediately and employ familiar methods, moving steadily through the task towards the accepted goal. Over time, they become more efficient as they directly engage the task, most likely following the ubiquitous power law of skill learning (e.g., VanLehn, 1996). In the adaptive pattern, people take an initial period to explore or adapt their ideas, practices, and/or environment. They are slower to start, but they can make up the lost time if they make an appropriate adaptation.

---

<sup>2</sup> People may have multiple goals when working on a task – e.g., they may wish to take a walk for exercise while also wanting to “explore the neighborhood” – and these goals may be complementary or opposing. For simplicity, we assume in this discussion that progress on a task is measured by movement toward a single goal, as is common in the problem solving literature.

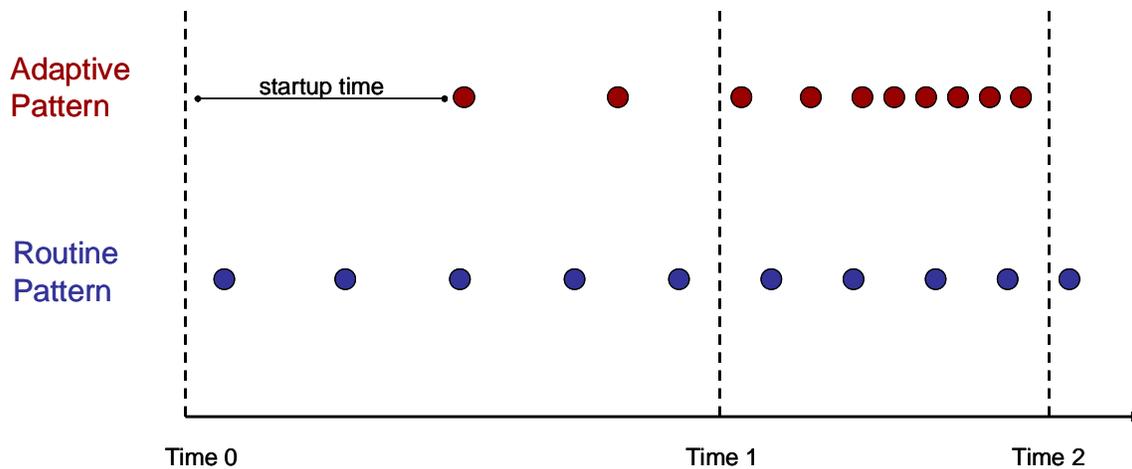


Figure 1. Two approaches to a recurring task.

Not all adaptations will be appropriate. Some will be effective, but others will be mal-adaptations. In honoring Hatano’s term “adaptive,” we do not mean to equate adaptive and effective. Instead, we note that without some form of adaptation, people will remain in the routine pattern. Both the routine and adaptive patterns have their relative merits, and it is important to maintain a balance of the two. In problem solving, for example, one must balance time demands for problem planning and exploration with actual solution execution (Schoenfeld, 1985). Similarly, in design processes, it is important to set aside time for brainstorming, fact-finding, and prototyping. At the same time, it is important to know when to stop adapting the plan, so there is still enough time to meet production deadlines (Eger, Eckert, & Clarkson, 2005).

The relative value of these two patterns will often be a function of the performance and evaluation horizon. Consider the two evaluation points, Time 1 and Time 2, shown in the figure. If a task has a Time 1 deadline and it is not recurrent, then the routine approach can be more effective. In computer programming, a “hack” can be just as effective as an elegant solution, if the program is small and the code will not be reused. However, if the horizon extends to Time 2,

then the adaptive pattern may be more effective. For example, if the computer program someday needs to handle new types of data and procedures, an elegant solution will make it easier to modify.

The evaluation horizon has important implications for the assessment and design of instruction. If student tests occur in the short-term, there will be a pull for immediate efficiency and the routine pattern will prevail, because this pattern achieves short-term results. However, a test at Time 1 can be misleading, if the goal of instruction is to prepare students for a longer trajectory of learning (Bransford & Schwartz, 1999). This is one reason that metrics of adaptation may be useful. They may better predict the effectiveness of instruction for a distant Time 2.

#### Catalysts to Adaptive and Routine Patterns

Given the natural tendency to stick with the tried and true, when and why do people engage in the adaptive pattern? Hatano and Inagaki (1986) proposed three factors that support a trajectory towards adaptive expertise. One factor involves the degree to which risk is associated with performance: “when a procedural skill is performed primarily to obtain rewards, people are reluctant to risk varying the skills, since they believe safety lies in relying on the ‘conventional’ version” (p. 269). A second factor is whether a situation has sufficient variability to warrant adaptation. For instance, formal instruction often attempts to reduce all ambient variability so students can focus on a procedural skill. This can have the unintended consequence of preventing students from considering how to vary procedures in response to new situations. The third factor involves the degree to which the local culture emphasizes understanding. “A culture, where understanding the system is the goal, encourages individuals in it to engage in active

experimentation. That is, they are invited to try new versions of the procedural skill, even at the cost of efficiency” (p.270).

To this list, we would like to add two potential catalysts to the adaptive pattern. One catalyst is reactive: when a behavior becomes too burdensome or flawed, and it is hard to get by with usual routines, people will try something new. As Stevens et al. (2006) note, adaptations often occur in response to a specific problem, whether a new crisis or a chronic snag. We will call these fault-driven adaptations, when the situation “forces” adaptation (assuming an individual or group does not quit the situation altogether). Fault-driven adaptations fit well with the common definition of adaptation as an “effective change in response to an altered situation” (White et al., 2002, p. 2).

The second catalyst is proactive in nature and leads to prospective adaptations. In this case, people engage the adaptive pattern before they confront any specific snags. Extensive planning is one good example of a prospective adaptation. Schoenfeld (1992) described how one mathematics faculty member solved a difficult problem. Most notably, he spent the majority of his time working to understand the problem and to choose an approach to solving it. In contrast, high school and college students tended to “read, make a decision quickly, and pursue that direction come hell or high water” (p. 61). The differences between expert and novice planning that Schoenfeld describes map well on to the distinction between adaptive and routine patterns.

It is important to note that people may engage the adaptive pattern as a matter of course, particularly when they have experienced its value for specific classes of activity. Experts, who have engaged in the adaptive pattern many times, passing through many implementation dips, may have learned that long-term benefits often outweigh short-term costs in certain areas of their practice. For example, Ericsson, Krampe, and Tesch-Römer (1993) noted that many chess

masters improve primarily by playing chess. They follow the routine pattern. In contrast, grand masters devote hours a day to deliberate practice – they study and attempt to predict moves in published games from great chess matches. They step back from playing chess to develop a deeper understanding that can set the stage for greater future success. Although they exhibit the adaptive pattern by habit, it is nevertheless adaptive because it does not follow the more typical routine of just playing chess to get better.

### Expertise in Representational Tools

The adaptive pattern can comprise different activities, as in the cases of deliberate practice or planning ahead. In the current work, we are particularly interested in adaptations that involve the introduction of a tool that changes the conditions of problem solving. Not only are experts able to perceive and make use of structures in the environment that are invisible to novices (Chase & Simon, 1973), they can also create new structures to adapt the environment. Kirsh (1996) argues that “introducing a tool is one of the easiest ways to change an agent’s action repertoire, for now it is possible to do things previously unattainable, or unattainable in a single step” (p. 438). New tools can simply amplify a person’s existing capabilities, as pliers enhance grip strength, or they may completely transform the nature of a task, in the way that a spreadsheet changes financial planning or the printing press changed information distribution (Eisenstein, 1979; Pea, 1985).

Representational tools are especially relevant to adapting cognitive tasks. A good representation can greatly improve problem solving (Zhang, 1997), and introducing a representation can help with a class of related instances, even if they are not fully known beforehand. A list representation, for example, helps identify duplicates and missing items,

which would be especially useful when aggregating information ahead of a series of relevant problems (Collins & Ferguson, 1993).

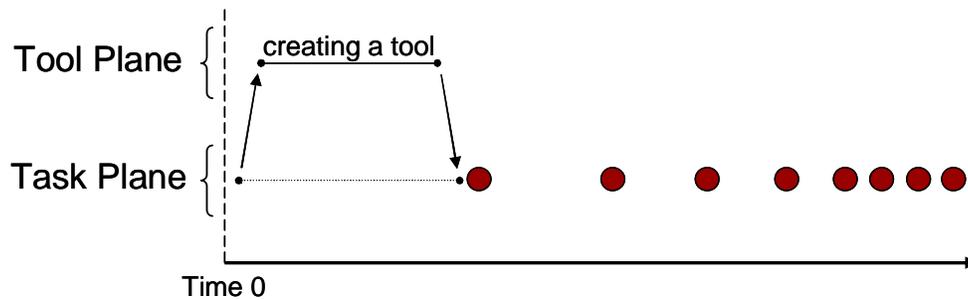


Figure 2. Startup time can be spent creating a tool.

Figure 2 presents the hypothesis that one important instance of the adaptive pattern involves the creation of a representational tool to help with a class of problems, much like an algebraic representation solves for a class of problems whereas arithmetic solves for a single instance. The figure distinguishes two planes of activity. The task plane involves solving the received problems. The tool plane involves stepping away from the immediate task to fashion a representational tool designed to help when returning to the task plane. Of course, people can shuttle between these two planes of activity throughout the history of a task; the figure oversimplifies for explanatory ease.

There were several reasons to believe that we might find instances of representational adaptive expertise, where people move to the tool plane to create an organizing representation. One reason is that representational tools, such as lists and matrices, are relatively easy to make and modify with just a pencil and paper, at least compared to making physical tools like hammers or steam engines. Thus, the experienced cost of the adaptive pattern will be low.

A second reason is that people are relatively good at making visual representations and they readily appreciate their value for understanding and learning, if given appropriate experiences (Enyedy, 2005; Petrosino, Lehrer, and Schauble, 2003). diSessa, Hammer, Sherin, and Kolpakowski (1991), for example, studied a small group of 6<sup>th</sup>-grade students as they worked to invent representations of motion. They found that, with proper guidance, the students were able to create representations of motion that closely approximated canonical versions. In addition, the students developed a good understanding of the benefits and drawbacks of various representational conventions. This work led them to put forth the concept of meta-representational competence, which diSessa and Sherin (2000) define as “the full range of capabilities that students (and others) have concerning the construction and use of external representations” (p. 386).

A third reason is that a simple repertoire of representations (e.g., trees, matrices, Venn diagrams) can be applied broadly with only moderate adaptation. Therefore, representational adaptation is a type of adaptive expertise that one might expect to transfer to novel settings (cf. Novick & Hurley, 2001). For instance, in one study, middle-school students received experiences inventing and working with a variety of visual representations (Schwartz, 1993). Several weeks before and after the intervention, the students received in-class assignments that were amenable to visualization and served as surreptitious pre- and posttests. At pretest, no students constructed visualizations. At posttest, nearly half of the students tried to invent a visualization, even though they had not learned a visualization specific to the structure of the posttest problems.

To probe for prospective adaptations with representations, the following research employed a medical diagnosis task. The relevant information for making a series of diagnoses was distributed across many sheets of paper. Would the participants move to the tool plane to

organize this information to help in subsequent diagnoses, or would they simply shuffle through the papers to solve each new case? We varied both the task demands and the level of participant experience to see whether these would influence the use of fault-driven and prospective adaptations.

### Overview of the Experiment

The identification of expertise generally takes one of two broad approaches. In the practice-centered approach, an expert is associated with a stable practice, role, or domain of knowledge. For example, chess experts can be identified by the fact that they win in organized competitions; history experts can be identified by the fact that they have helped to define the historical domain of study; and chick-sexing experts can be identified by their tenure on the job (Beiderman & Shiffrar, 1987). The practice-centered approach has led to a number of important questions. Is a community's identification of an expert related to social prestige (Agnew et al., 1997)? Does identification as an expert lead to expert-like behaviors? Do experts have personality traits that differentiate them from non-experts with equal years of experience (White, et al., 2005)? Can experts transfer their abilities from one domain to another, for example, through interdisciplinary work?

The second approach to identifying expertise is skill-centered. In this approach, which we adopt, an expert is identified by performance or experience with a specific skill. For example, Novick (1988) sorted undergraduates into novice and expert categories based on their SAT mathematics scores. The skill approach is modular in character, because the skill does not comprise the totality of any readily identified role, and the skill can often recur in many different activity contexts.

To investigate routine and adaptive patterns within a problem-solving context, the study examined two groups as they worked through a complex medical diagnosis task. One group comprised graduate students and recent PhDs from a variety of non-medical fields, including engineering, physical and social sciences. Although they were not full-fledged disciplinary experts, they had spent years working with and trying to understand complex data within their discipline, and thus were relative experts at handling complex information. The other group comprised undergraduates, who knew how to make representations but were relative novices compared to the graduate students.

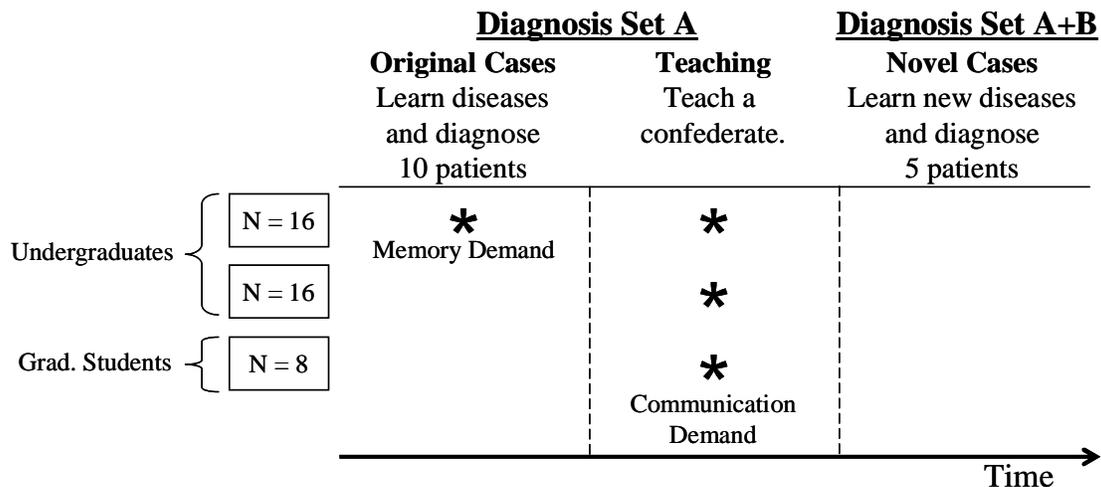


Figure 3. Design of the Experiment. The asterisks indicate places where we attempted to induce fault-driven adaptation. The graduate students test whether sustained educational experiences lead to prospective adaptation. The final diagnosis session at the right of the figure tests whether students who have experienced fault-driven adaptation will use prospective adaptation when the fault has been removed.

Figure 3 provides an overview of the total research design. There are several components to the design, so we will describe the logic in some detail working from left to right in the figure. As mentioned above, we included undergraduates as relative novices and graduate students as relative experts. The medical diagnosis task was novel for both. Many expertise studies have

used problems that are familiar to experts, who can readily solve them in routine ways. Using a relatively novel problem can help probe for adaptive expertise (e.g., Wineburg, 1998).

In the first phase, participants received a set of reference cases, one per sheet of paper. Each reference case included a patient, the test results, and the diagnosis. These provided the information that participants would need to diagnose new patients presented on the computer one at a time. For diagnosing new patients, participants ordered tests on the computer to gather diagnostic information, and they were told to minimize the number of test requests to “keep costs down.” When ready, they submitted their diagnosis on the computer.

Participants received the reference cases and the computer diagnosis task at the same time, so they could begin diagnosing the new patients on the computer at any point. Blank paper for note taking was available to participants at all times. Pilot data showed that it was possible to diagnose the new patients successfully based on the reference cases with or without taking notes. Therefore, it was possible to accomplish the task without shifting to the tool plane.

Participants completed one of three conditions. The first two conditions contrasted graduate and undergraduate students when they had continuous access to the reference cases throughout the diagnosis task. For example, when diagnosing a patient on the computer, they could consult the reference cases if they wanted. Our hypothesis was that the graduate students would exhibit prospective adaptation – they would move to the tool plane and take the time to create a visual representation of the reference cases before starting their diagnoses on the computer. In contrast, the undergraduates would follow a more routine pattern by beginning their diagnoses of the new cases quickly and shuffling through the reference cases while working on each new diagnosis. If so, to our knowledge, this would be a rare documentation of a correlation

between long-term educational experiences and subsequent adaptiveness in a novel task domain (cf., Miyake & Pea, 2007).

The third condition, run with undergraduates only, was designed to demonstrate fault-driven adaptation. These students only had intermittent access to the reference cases – they could not look at the reference cases when they were diagnosing a new patient on the computer. They could look at the reference cases at any other time (e.g., between patients) but not when they were actually in the process of ordering tests and diagnosing a patient. We assumed that students in this condition would find they could not remember the reference cases well enough to do each diagnosis. This would cause a fault in problem solving, and the students would be driven to adapt their problem solving by creating notes to help them remember the reference cases. We did not create a parallel, intermittent access condition for the graduate students, because pilot data indicated they would make prospective adaptations, so the fault condition would not have affected them.

In the second phase, we asked the students to teach another person how to solve the diagnosis problems. Diagnosis with sprawling facts is a difficult process to describe. We hypothesized that this would create pressure to make representations for communicative purposes. The introduction of a teaching demand was specifically aimed at those who failed to create a representation in the first phase of problem solving, and allowed us to ask several questions. Would they exhibit prospective adaptation and create representations for communication even if they had not done so for problem solving? If so, it would show that they were capable of creating representations, making their failure to do so earlier more notable. It would also set the stage for another question: would they capitalize on their new, communicative representations upon returning to problem solving in the final phase of the study?

In the final phase, students returned to the diagnosis task. There were two changes in this phase. First, we added a novel set of reference cases that covered two new diseases. The students now had to diagnose new patients on the computer that could have any of the expanded set of diseases. Adding new diseases made the students' earlier problem solving representations out of date. Second, we removed all fault-driven pressures to the adaptive pattern; everyone had continuous access to all the reference cases.

A critical question was, how would students' prior experiences in the study affect their creation and use of representations in this final problem solving phase? For those who had created representations in response to a memory burden (undergraduates in the intermittent condition), would they still bother to make a representation once the burden was removed? Stated another way, would they move from their fault driven adaptation to a prospective adaptation? For those who created representations for teaching (phase 2) but not for their own use (phase 1), would they build upon their phase 2 representations, or would they not appreciate the functionality of the representations for both communication and problem solving? Finally, for those who made prospective adaptations from the beginning, would they show superior learning for the new diseases in phase 3. That is, did the adaptive pattern in the first phase prepare students for future learning in the third phase? Ideally, the answers to these questions will have instructional implications for putting people on a trajectory to adaptive expertise.

## Method

### Participants

Participants included 32 undergraduate students (21 women, 11 men) recruited from a paid subject pool. We also recruited 8 people who were pursuing or had recently received a PhD

in the sciences or engineering (5 women, 3 men). None of the participants had any medical training. All of the participants were from a highly selective university.

### Materials

The physical materials used in this study included two sets of reference cases (A & B), a computer program for performing diagnoses with new patients, an answer sheet, and pens and paper for note taking.

Each reference case resembled a simple medical chart, printed on a single side of a sheet of paper. A patient's name, their doctor's name, and their ID number appeared at the top of each page. Nine medical tests, along with associated results, appeared below. Each reference case included the patient's diagnosis at the bottom of the sheet. Each disease appeared as the diagnosis for exactly two reference cases. Twelve reference cases were used for Set A (covering six diseases), and four reference cases were used for Set B (covering two new diseases). Appendix A includes a sample reference case and a table showing the relationship between diseases and tests.

A computer program provided an interface for diagnosing new patients. Figure 4 shows that each diagnosis screen gave a patient's name and showed nine possible medical tests. Next to the name of each medical test was a light blue rectangle which, when clicked, revealed the result of that test. Once participants believed they had enough information to make a diagnosis, they wrote down their diagnosis on a diagnosis sheet of paper (with one line for each patient diagnosis). They could then click the "Next" button to reveal the correct diagnosis.

Patient name: E. Moore

Standard diagnostics:  
 Blood Pressure [text box]  
 Temperature  High  Normal  
 Heart rate [text box]

Patient complains of:  
 [text box]

Standard Blood screen:  
 Anemia  Positive  Negative  
 Elevated white blood cell count [text box]

Jacobson Test-Battery:  
 Reflex test [text box]  
 Visual acuity test  Abnormal  Normal  
 Treadmill test [text box]

Next

Figure 4. Screenshot from the computer interface for diagnosing new patients. Several tests have been ordered (clicked) to reveal their results.

### Procedure and Design

The study had three phases: original cases, teaching, novel cases. Only the first phase, original cases, included a between-subjects manipulation.

#### Original Cases.

Participants heard they would be working on a medical diagnosis task, and that they would be relying on a set of reference cases to diagnose new patients. They were shown how to use the computer interface for a practice diagnosis. Participants were told to order the fewest number of medical tests possible for each patient (to keep costs down). They then received the set of 12 reference cases, covering 6 diseases (Set A). The researcher explicitly indicated the pen, blank paper, and answer sheet for the diagnoses.

Undergraduate participants were randomly assigned to the continuous access or intermittent access conditions, with 16 in each condition. All the graduate students were assigned

to the continuous access condition. The intermittent condition imposed a memory burden – the participants were not allowed to look at the reference cases while making a diagnosis on the computer (they could look at the cases between diagnoses). The continuous condition allowed participants to consult the reference cases at any time. Participants in all conditions could take notes at any point, and they were allowed to look at their notes at any point. Participants worked at their own pace until they had worked on 10 cases<sup>3</sup>. Everyone saw the 10 cases in the same order, and they could not go back. For each new patient, participants could order up to nine medical tests. We recorded their test order for subsequent analyses of diagnostic efficiency.

### Teaching.

At the conclusion of the original cases, the students heard that they would teach another person how to diagnose patients with the six possible diseases. They heard that the new person was familiar with the structure of the task, in terms of the reference cases and the computer interface, but not with the particular diseases and symptoms that they had been working with. They had five minutes to prepare to teach, during which the experimenter left the room. They had all the original cases, plus sheets of paper to help prepare. After five minutes, the experimenter returned with a confederate who was blind to condition and our hypotheses. The students had five minutes to explain their method for diagnosing patients. When the students finished with their explanation, or when five minutes had elapsed, the experimenter dismissed the confederate and began the next part of the study.<sup>4</sup>

---

<sup>3</sup> Many participants had external time constraints that limited the total amount of time they could spend in the study. In four instances, participants were moving too slowly through the original cases to finish the study within the time allowed. In these instances, the experimenter waited until 30 minutes had elapsed, allowed the participant to finish their current diagnosis, and then had them move on to the next phase of the study, even though all ten diagnoses had not been completed. All participants were able to complete the Teaching and Novel Cases phases in full. We discuss our procedures for handling these missing data below.

<sup>4</sup> A secondary manipulation in the teaching phase had no effect and will not be discussed in the Results. During the teaching part of the study, participants were randomly assigned to one of two conditions, which varied in whether or not we provided a set of practice cases they could use to help teach. Very few participants made use of this practice

### Novel Cases.

Everyone received a set of four new reference cases, which covered two new diseases (Set B). Their task was the same as before – they should diagnose new patients as accurately as possible, while minimizing the number of medical tests they ordered for each patient. All participants received full access to the reference cases (both Set A and B). Students in the intermittent condition were told that they no longer needed to set the reference cases aside at any point. The students worked at their own pace and completed five new diagnoses in the same order across participants. The new patients included diseases from both Set A and Set B.

### Coding

The study included a number of outcome measures. Standard measures included time and accuracy, described in the Results section. We also coded the students' representations and we measured the optimality of their test selection for diagnosing each patient. We describe these two measures here.

---

set, and neither the use of the practice set nor the condition to which participants were assigned had any effect in the subsequent analyses.

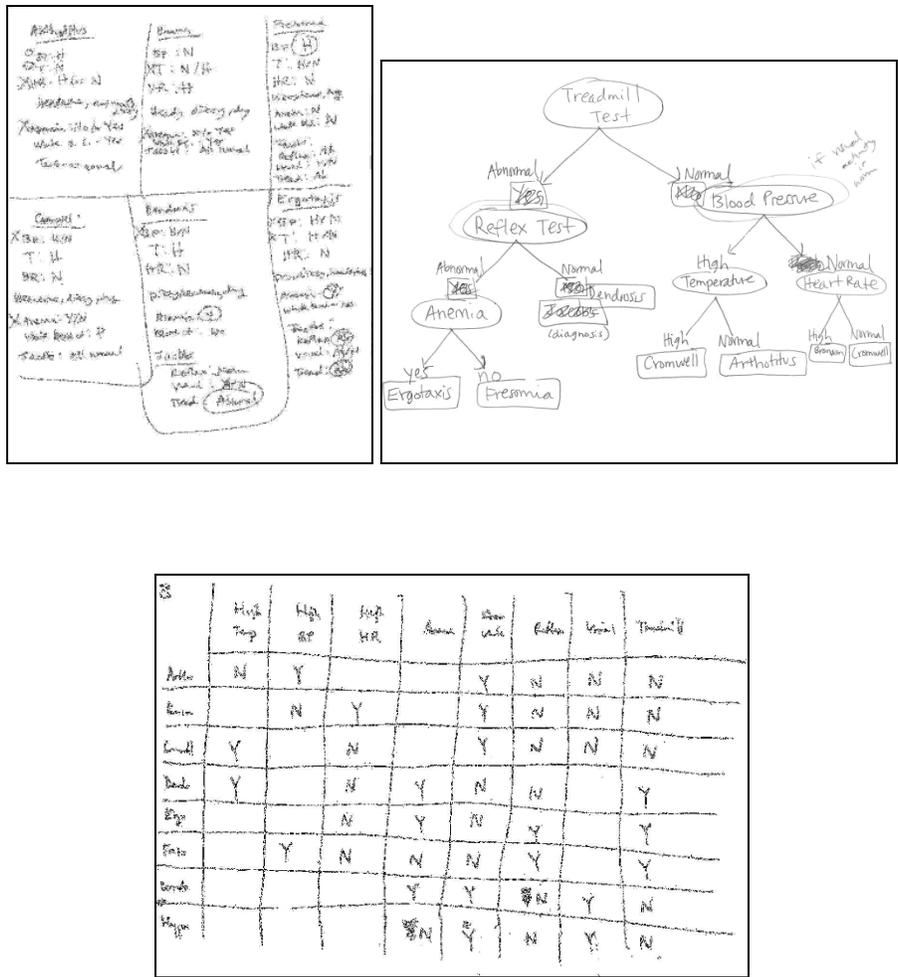


Figure 5. Examples of representations. The two upper representations were created by one participant. The upper-left corner shows a List representation created during the original cases. The upper-right corner shows a Tree representation created during the teaching phase. The lower representation was created by a different participant. It shows a Matrix representation created during the novel cases phase.

Participants created a variety of representations throughout the study, varying from no representation whatsoever to a representation that spanned seven pages. Figure 5 shows representations from two participants. We coded each representation for each phase of the study into one of three categories. When students did not make a representation, they received a code of No Representation for that phase. When students made a representation that organized information by symptom or disease, they received a code of List. The upper-left representation in

Figure 5 is an example of a List representation. A code of Matrix/Tree indicated that the representation contained either a decision tree structure, with nodes and branches, or a matrix structure with rows and columns for diseases and symptoms. The upper-right representation in Figure 5 contains tree structure elements, while the lower representation in Figure 5 has a matrix structure. Both were coded as a Matrix/Tree representations. While these categories collapse across some details, they were sufficient to code all of the representations without losing too much detail for subsequent analyses. For example, no subjects simply copied the information from the reference cases onto the paper; they all imposed some additional organization, for instance, by integrating the symptoms from two patients who had the same disease. Some students made both a List and a Matrix/Tree representation in the same phase. They were categorized according to their final representation. In all such cases, they made a List first, and then shifted to a Matrix/Tree.

To establish the reliability of the coding scheme, a second coder trained with a set of six sample representations. The secondary coder then evaluated representations from 20 of the participants, blind to their conditions. Overall agreement was 93%.

The diagnosis task was designed so that all participants could perform at high levels of accuracy. Accuracy was unlikely to be informative with respect to the benefits of the routine and adaptive patterns. Therefore, we created a measure of how optimally students moved through the search space of possible diagnostic tests (recall that students were told to order the fewest number of medical tests possible for each patient). The measure, called the Weighted Optimality Ratio (WOR), is a measure of how close each test sequence was to being perfectly optimal. Given the information available at a given moment, students could order tests that would

eliminate alternative diseases more or less effectively. Appendix B provides a detailed explanation of the WOR metric.

## Results

The study can be partitioned into two sets of outcomes. The first set involves whether participants engaged in the routine or adaptive patterns. We attempted to influence these outcomes through our choice of experimental conditions and participants. The second set of outcomes concerns how the routine and adaptive patterns are associated with performance on the diagnosis task. We present our results chronologically by phase of the experiment.

### Original Cases

#### Tool Plane Activity.

Table 1. Frequency of creation of a representation for the original cases, organized by condition and type of representation.

	Continuous	Intermittent	Continuous
	<u>Undergraduate</u>	<u>Undergraduate</u>	<u>Graduate</u>
Matrix/Tree	1	2	6
List	2	12	2
No Representation	13	2	0

We first consider how condition influenced the creation and use of representations. Table 1 shows the frequency with which participants in each of the conditions created representations of various types. The conditions led to large differences in the types of representations created,  $\chi^2(4, N = 40) = 34.6, p < .001$ . Notably, only 19% of undergraduates in the continuous access condition created a representation of any sort, in contrast to 88% of undergraduates in the intermittent access condition and 100% of graduate students in the continuous condition. These

differences indicate several things. First, given continuous access to the reference cases, undergraduates chose to “get by” without creating an external representation. Most students in this condition accomplished the task by shuffling papers, organizing them into piles as they went.

Second, graduate students with continuous access unanimously chose to create external representations. They presumably could have gotten by just like the undergraduates, but they instead chose to make a prospective adaptation and create a representation. A closer look at the representational activity indicates that 50% of the graduate students evolved their representations (shifted from List to Matrix/Tree), whereas only 12.5% of the undergraduates in the intermittent condition evolved their representations,  $\chi^2(1, N = 24) = 4.0, p < .05$ . Thus, the prior experiences of the graduate students led them to create representations when they did not have to, and often more than once.

Finally, although undergraduates in the intermittent access condition created representations at a high rate, they did so for a different reason than the graduate students. Because otherwise equivalent undergraduates who did not have the memory burden did not make representations, we can conclude that undergraduates in the intermittent access condition were making fault-driven adaptations not prospective ones. In addition, graduate students created different types of representation than undergraduates. Of all the undergraduates who created representations, across both continuous and intermittent access conditions, 21% created Matrix/Tree representations, while 75% of graduate students created Matrix/Tree representations.

The pattern seen above, where condition was a significant predictor of the creation of external representations, supports the hypothesis that expertise, in the graduate students, and a

memory burden, in the intermittent access condition, lead to an adaptive pattern with respect to the creation of representations. This pattern also holds true when we examine time data.

#### Task Plane Efficiency -- Time.

With respect to time, we have two main questions. First, how long did students spend preparing, before they began the first diagnosis item? We will call this the startup time. Second, once they began the first item, how long did they spend diagnosing the ten patients in the original cases? We will call this the diagnosis time. We conducted a multivariate analysis of variance, with condition as the between-subjects factor and startup time and diagnosis time as dependent measures<sup>5</sup>. There was a main effect of condition on startup time,  $F(2, 37) = 21.93, p < .001$ . Undergraduates in the continuous access condition were the fastest to begin the first item ( $M = 102$  s), followed by undergraduates in the intermittent access condition ( $M = 503$  s), and then the graduate students in the continuous access condition ( $M = 937$  s). Post hoc tests using Tukey's HSD showed that all pairwise differences between conditions were significant,  $p < .01$ . Participants differed by condition in how long it took them to begin diagnosing, but not in how long it took them to complete their diagnoses once they had begun. There was no effect of condition on time spent diagnosing once participants began,  $F(2, 33) = 0.22, p > .05$ .

The time differences between conditions were driven by the representational activity (see Table 1 for the association of condition and representation type). This relationship becomes evident in Figure 6, which shows the time data across the ten diagnosis items, organized by type of representation. Numbered dots indicate the mean time at which participants began each

---

<sup>5</sup> As mentioned above, four participants did not complete the diagnoses in the original cases. Omitting their data from analysis of time might bias the results, as these participants were slower than average, so we extrapolated from their existing data to fill in missing data. Participants sped up in an approximately linear fashion across the diagnoses. We computed the regression equation for the speed increase from one problem to the next (slope = 0.5, intercept = 28s,  $r = .6, p < .01$ ). We used these parameters to extrapolate the missing values given the last data point from each participant. To compensate for the computed values, we subtracted 4 degrees of freedom for the F-test on diagnosis times.

diagnosis. The first dot, labeled “1,” indicates when participants began their first diagnosis on the computer, and the “End” dot shows the mean ending time.

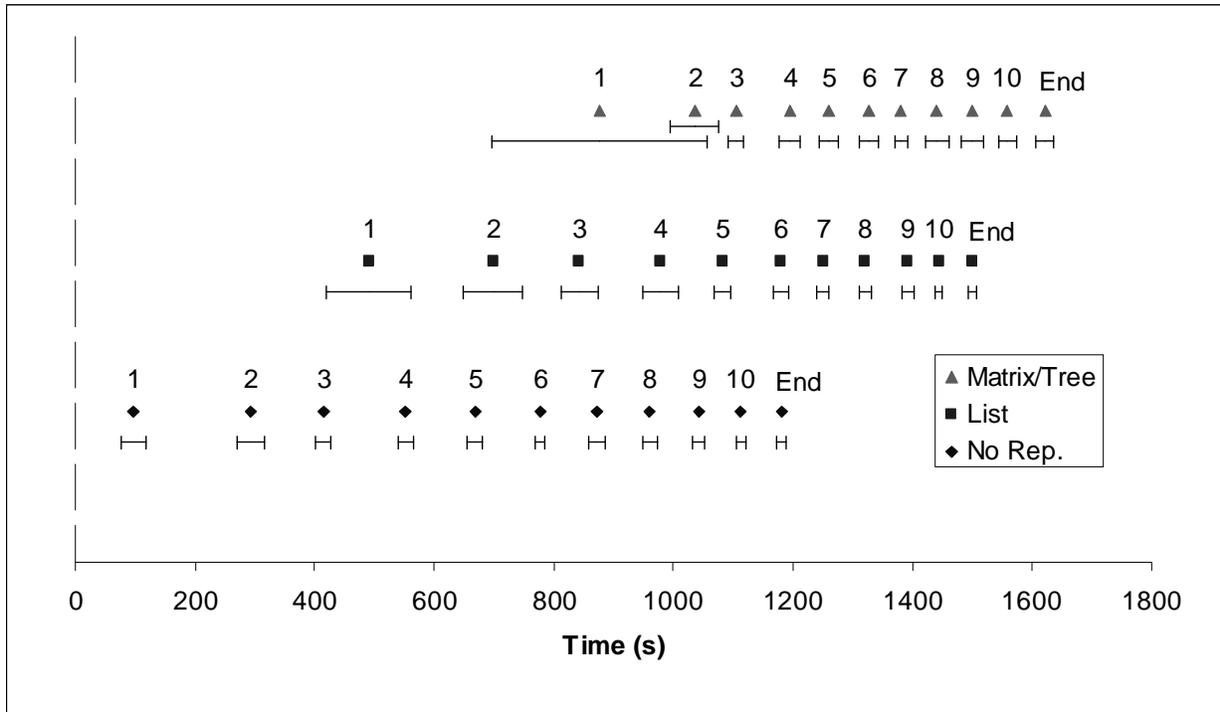


Figure 6. Speed of performance on the original cases. Error bars indicate +/- 1 standard error of mean time per problem.

We conducted a second multivariate analysis of variance, this time with type of representation as the between-subjects factor, and startup time and diagnosis time as dependent measures. There was a main effect of representation type on startup time,  $F(2, 37) = 18.03, p < .001$ . The No Representation group was the fastest to begin ( $M = 96$  s), followed by the List group ( $M = 491$  s), and then by the Matrix/Tree group ( $M = 876$  s). Tukey’s HSD showed that all pairwise differences between groups were significant,  $p < .05$ . There was no effect of representation type on time spent diagnosing,  $F(2, 33) = 2.38, p > .05$ . Creating a representation, especially a complex Matrix or Tree representation, was not time efficient. The benefits of the adaptive pattern, if any, lay elsewhere.

## Task Plane Efficiency -- Accuracy

By design, the diagnosis task could be completed accurately regardless of condition or type of representation employed, and neither condition nor type of representation were significant predictors of mean accuracy across the ten items in the original cases. (Mean accuracy by condition: continuous undergraduate 91%, intermittent undergraduate 89%, continuous graduate 94%). However, in a medical setting, each diagnostic test has an associated cost, and students were told to minimize the number of tests they ordered. Therefore, we also looked at whether the participants made an optimal selection of tests using the WOR statistic.

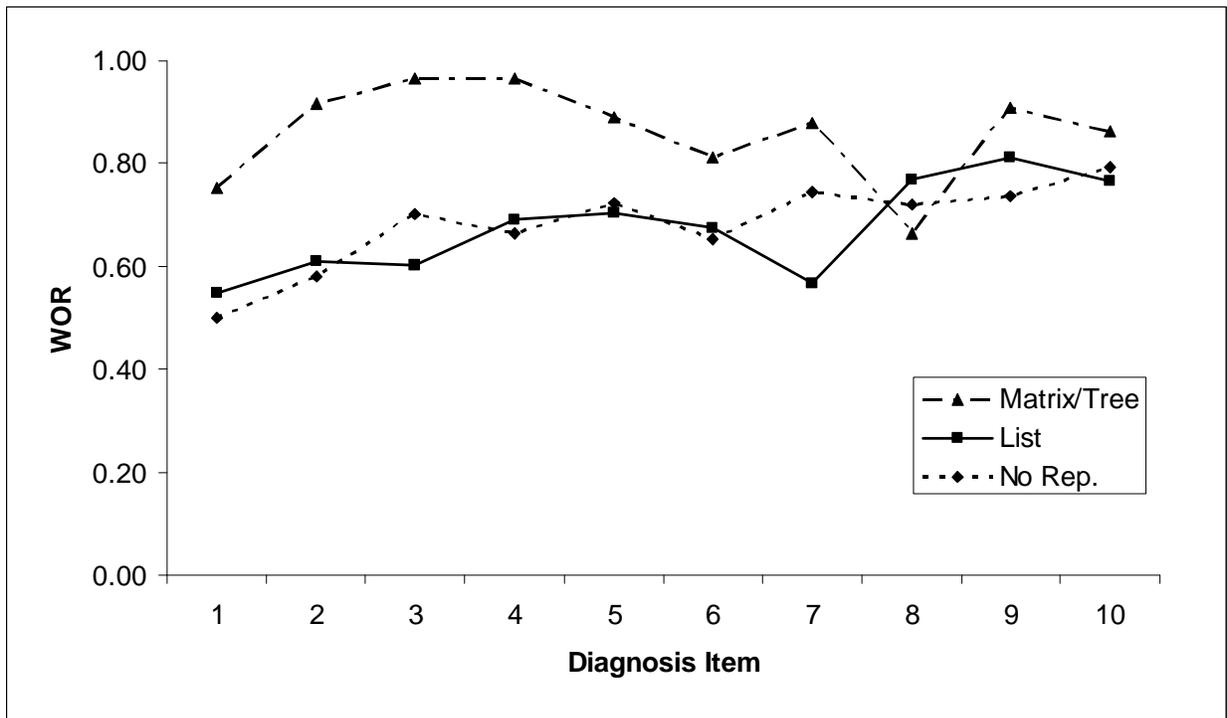


Figure 7. WOR across original cases, by type of representation.

As we will describe, type of representation, as shown in Figure 7, but not condition, predicted participants' performance in test-ordering efficiency. We conducted a 3x10 repeated measures analysis with condition as the between-subject factor, and WOR for each diagnosis

item (case) as the repeated measure<sup>6</sup>. There was a main effect of diagnosis item – participants' WOR improved over time,  $F(6.4, 211.4) = 3.54, p < .01$  (degrees of freedom adjusted for sphericity violation). There was no effect of condition,  $F(2, 33) = 2.01, p > .05$ , and there was no interaction between diagnosis item and condition,  $F(12.8, 211.4) = 0.81, p > .05$  (degrees of freedom adjusted).

A stronger picture emerges when we consider the types of representations that students created. We conducted a 3x10 repeated measures analysis with type of representation as the between-subject factor, and WOR for each diagnosis item as the repeated measure. We again found a main effect of diagnosis item, with participants improving over the course of the ten items,  $F(9, 297) = 3.84, p < .01$ . There was also an effect of type of representation,  $F(2, 33) = 5.48, p < .01$ . Tukey's HSD showed that those creating Matrix/Tree representations ( $M = .86$ ) significantly outperformed those creating Lists ( $M = .67$ ) or No Representation ( $M = .68$ ),  $p < .05$ . Those creating Lists and No Representation did not differ significantly. There was also a significant interaction between diagnosis item and type of representation,  $F(18, 297) = 1.96, p < .05$ . This effect is driven by the relatively early efficiency of the Matrix/Tree representations. Thus, although the prospective adaptation of creating Matrix/Trees led to slower and no more accurate diagnoses, it did create a situation where participants could better consider all the symptoms and diseases simultaneously which led to more cost-effective test requests. (We have no definitive explanation for the various peaks and troughs, such as the large drop in performance on item 8 for those using Matrix/Tree representations.)

### Teaching

---

<sup>6</sup> As mentioned above, four participants did not complete all ten diagnosis items in the original cases. Their data were omitted for all analyses of WOR in the original cases. Their mean WORs on the problems they did complete were no different from the overall mean WOR, thus there was no reason to believe that omitting or extrapolating their data would affect the results.

In the teaching phase, the measure of interest is what kind of representation, if any, participants created to aid them in teaching. Table 2 holds the representational activity for the teaching phase. The table only includes new representational work undertaken in the teaching phase, whether it was done for a new representation or as a modification to an existing representation.

Table 2. Frequency of creation of a representation specifically for teaching, organized by condition and type of representation.

	Continuous	Intermittent	Continuous
	<u>Undergraduate</u>	<u>Undergraduate</u>	<u>Graduate</u>
Matrix/Tree	2	6	5
List	8	6	0
No Representation	6	4	3

The undergraduates from the continuous access condition increased their rate of representational activity for the teaching phase from 19% to 63%,  $\chi^2(1, N = 16) = 6.35, p < .05$ . We interpret this to mean that communication demands, like memory demands, can lead adults to construct representations. The large majority of these undergraduates made their representations during the five minutes they had to prepare for teaching. Thus, in some sense, they were making prospective adaptations in anticipation of teaching. However, they received five minutes to prepare, so the activity of making the representations did not pull them away from the activity of teaching. The students' choice to make representations during the five-minute preparatory period suggests they had prior experiences that inculcated the value of making representations for communicating complex information. The two other conditions

exhibited non-significant drops in the percentage of people who created some sort of representation compared to the earlier problem-solving phase,  $\chi^2(1, N = 16) = .82, p > .05$ , and  $\chi^2(1, N = 8) = 3.69, p > .05$ , undergraduate and graduate students, respectively. This drop in representational activity was because several participants relied on the representations they had made for the original cases. These results are consistent with the hypothesis that a communication demand, in the form of teaching, would lead college students to create representations.

### Novel Cases

#### Tool Plane Activity.

We begin by considering the types of representations that participants created for the novel cases. Data from the novel cases provide an opportunity to consider how prior experiences within the study influenced participants' approach to handling new information and solving new diagnosis problems. Recall that for the novel cases, all the students had continuous access to the reference cases (Sets A & B), so fault-driven pressures to adaptation were no longer present. Table 3 shows the representational activity for the novel cases. As before, the table only includes new representational work undertaken in the novel cases phase, whether done for a new representation or as a modification of an existing representation.

Table 3. Frequency of creation of a representation for novel cases, organized by condition and type of representation.

	Continuous	Intermittent	Continuous
	<u>Undergraduate</u>	<u>Undergraduate</u>	<u>Graduate</u>
Matrix/Tree	1	4	6
List	2	8	1
No Representation	13	4	1

As expected, graduate students continued to make representations at a high rate. The more interesting contrast involves the two undergraduate conditions. Recall that 88% of undergraduates in the intermittent access condition made representations for the original cases. For the novel cases, they were now in a situation where the memory burden was gone. Table 3 shows that 75% made representations here as well. That is, they moved from fault-driven adaptation for the original cases to prospective adaptation as they undertook new representational work for the novel cases. This can be contrasted with the students in the continuous access condition. These students had made representations for the teaching phase, but they did not make use of these representations for problem solving. They simply set the teaching representations aside to continue with the routine pattern. The undergraduates who had been in the continuous condition created representations at a significantly lower rate than undergraduates who had been in the intermittent access condition,  $\chi^2(1, N = 32) = 10.16, p < 0.01$ , despite their relatively high rate of representation seen in the teaching phase.

Table 4. Creation and modification of representations in original and novel cases, across all conditions.

<u>ORIGINAL CASES</u>	<u>NOVEL CASES</u>	
	<u>No Representation</u>	<u>Representation</u>
No Representation	15	0
Representation	3	22

A second way to describe the data is to note that the creation of representations for diagnosis of the original cases was highly associated with the creation of representations for the

novel cases. This relation is shown in Table 4,  $\chi^2(1, N = 32) = 24.91, p < .001$ . Interestingly, the relation only held for representations that were created and used for diagnosis. Representations created for communicating, at least in the short run, had no effect on future use for diagnosis.

#### Task Plane Efficiency -- Time

The novel cases showed a similar time course as the original cases. As before, startup time refers to the time before working on the first diagnosis, and diagnosis time refers to the time taken to diagnose the five patients. We conducted a multivariate analysis of variance, with condition as the between-subjects factor and startup time and diagnosis time as dependent measures. There was a main effect of condition on startup time,  $F(2, 37) = 7.05, p < .01$ . Tukey's HSD showed that the graduate students ( $M = 303$  s) began the first problem significantly later than the undergraduate continuous access condition ( $M = 66$  s),  $p < .01$ . The undergraduate intermittent access condition ( $M = 153$  s) did not differ significantly from either of the other conditions. There was no effect of condition on diagnosis time once they began,  $F(2, 37) = 2.87, p > .05$ .

As with the original cases, the effect of condition on time was driven by representational activity. We conducted a multivariate analysis of variance, with type of representation for the novel cases as the between-subjects factor, and startup time and diagnosis time as the dependent measures. There was a main effect of type of representation,  $F(2, 37) = 23.28, p < .001$ . Tukey's HSD showed that the Matrix/Tree group ( $M = 339$  s) began the first problem significantly later than the No Representation group ( $M = 42$  s) and the List group ( $M = 132$  s),  $p < .01$ . The List and No Representation groups did not differ significantly. Unlike the original cases, there was an effect of type of representation on diagnosis time,  $F(2,37) = 3.50, p < .05$ . Although they got off to a slower start, those who created representations completed the diagnosis items more quickly..

Figure 8 shows participants' progression through the five diagnosis items in the novel cases. Dots indicate the mean start time for each problem, counted from the beginning of the novel cases phase. As is visible in Figure 8, those in the Matrix/Tree group did not have a chance to make up the time they spent initially. However, they were solving each problem faster, and given a sufficiently long time horizon, those using a Matrix/Tree representation may have surpassed those who began more quickly.

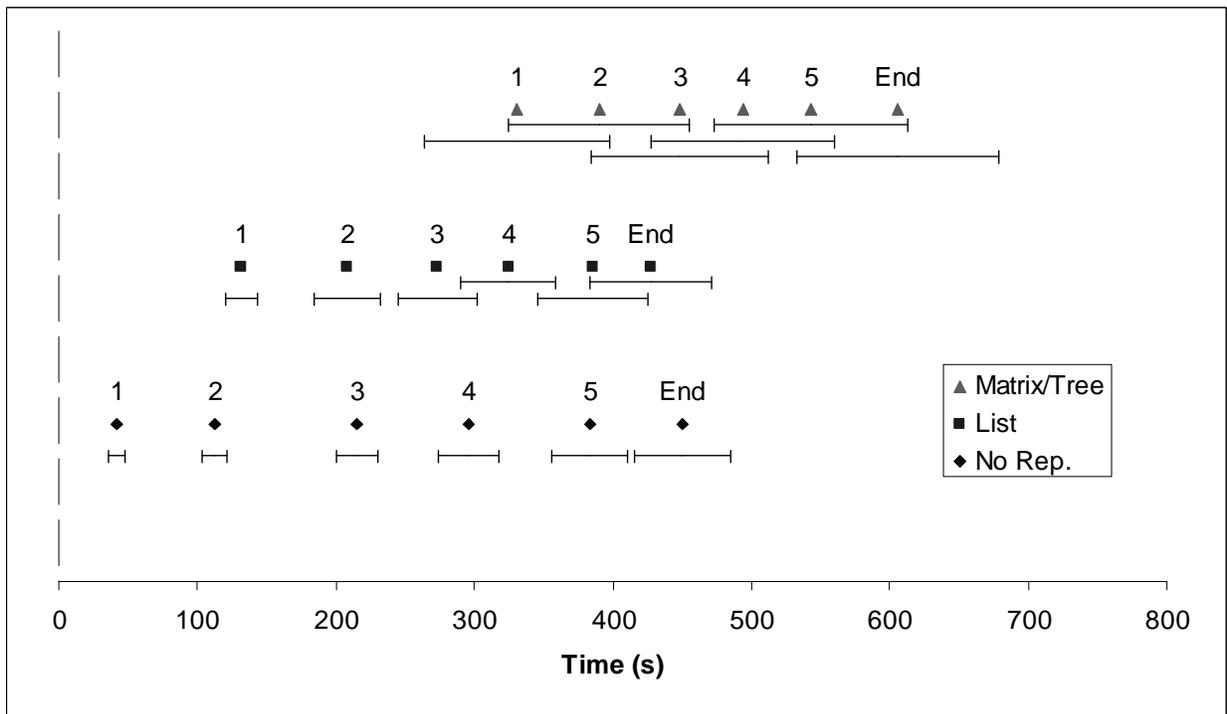


Figure 8. Speed of performance on the novel cases. Error bars indicate +/- 1 standard error of time per problem.

### Task Plane Efficiency -- Accuracy

As in the original cases, neither condition nor representation was a significant predictor of mean accuracy across the five diagnosis items in the novel cases. Instead, we focus on an analysis of the WOR. Again, type of representation, but not condition, predicted participants'

performance in diagnosis. We first conducted a 3x5 repeated measures analysis with condition as the between-subject factor, and WOR for each diagnosis item as the repeated measure. There were no significant effects: diagnosis item,  $F(2.9, 109.5) = 1.54, p > .05$  (degrees of freedom adjusted); condition,  $F(2, 37) = 1.17, p > .05$ ; diagnosis item by condition,  $F(5.9, 109.5) = 0.50, p > .05$  (degrees of freedom adjusted).

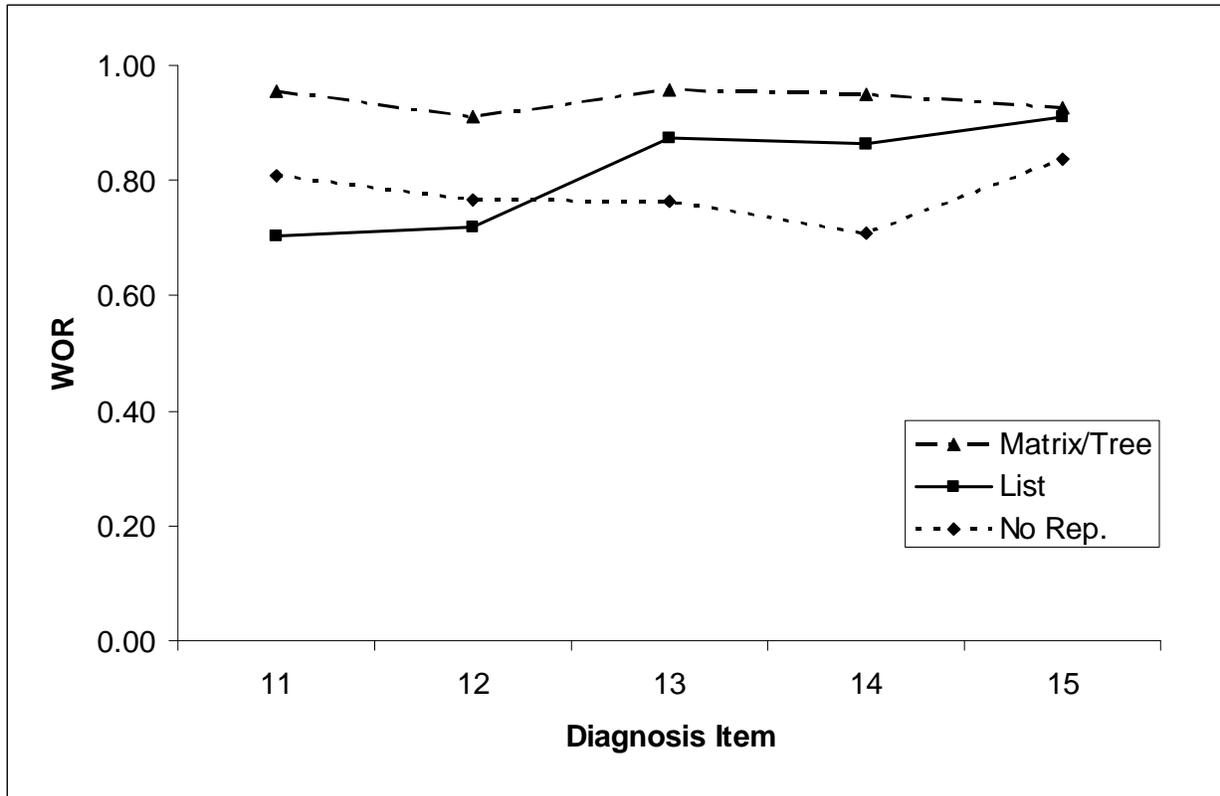


Figure 9. WOR across novel cases, by type of representation.

Type of representation tells a more interesting story. Figure 9 shows WOR by item across the novel cases, with separate lines for each type of representation. We conducted a 3x5 repeated measures analysis with type of representation created or modified for novel cases as the between-subjects factor and WOR for each diagnosis item as the repeated measure. There was a main effect of type of representation,  $F(2, 37) = 4.29, p < .05$ . Tukey's HSD showed that the

Matrix/Tree group ( $M = .94$ ) significantly outperformed the No Representation group ( $M = .78$ ),  $p < .05$ . The List group ( $M = .81$ ) did not differ significantly from either of the other groups. There was no main effect of diagnosis item,  $F(3.05, 112.97) = 2.66, p > .05$ . However, there was an interaction between diagnosis item and type of representation,  $F(6.10, 112.97) = 2.74, p < .05$ . The interaction was driven by the stable efficient performances of the Matrix/Tree representation and the gain of the List representation relative to No Representation.

### Discussion

The main finding involves the difference between the graduate and undergraduate students who had continuous access to the reference cases throughout the study. All of the graduate students in the continuous access condition made a prospective adaptation to create representations to aid them in diagnosing patients, while only one-fifth of undergraduates did so. Given that the undergraduates achieved equal levels of accuracy without representations, we presume the graduate students would have also been able to complete the task accurately without making representations. Nevertheless, the graduate students moved to the tool plane to create representations. This move cost them roughly fifteen minutes of startup time before diagnosing their first patient, compared to two minutes for the undergraduates. By the metrics of overall time and diagnostic accuracy, prospective adaptation was not a particularly wise thing to do. However, it did have some more subtle benefits.

When students made matrices or decision trees to organize the information in the reference cases, they made more cost-effective diagnoses. They ordered tests that optimally pruned the search space. Thus, by introducing a visual tool to organize the information space, they were prepared for all combinations of symptoms and diseases and could handle each instance well.

A second subtle benefit involved the amount of time needed to diagnose each patient. The time spent making representations was not recouped by the end of the original cases. However, by the end of the novel cases, participants who used the adaptive pattern had nearly caught up to those who had used the routine pattern. Projecting forward, they would have likely outstripped the routine pattern given more cases. By examining the students' behaviors on a subsequent task that required learning about novel cases, we begin to see the potential benefits of prospective adaptation – it can prepare people to learn more effectively when conditions change.

The differences between the graduate and undergraduate students cannot be attributed to the undergraduates not knowing how to make visual representations. In the intermittent condition, the undergraduates were not allowed to look at the reference cases when making a diagnosis. This created an excessive memory burden, and these students demonstrated a fault-driven adaptation. They created visualizations to get around the fault of limited memory capacity. They also created them for the demands of teaching; most students in the continuous access condition did not make representations for problem solving, but they did create them for the teaching phase. Thus, the undergraduates were capable of creating representations given the right demands – they made them for mnemonic and instructional purposes – but, for the original cases, they did not make prospective visualizations to optimize their search strategies.

For the final phase of the study, students had to diagnose patients given an expanded set of diseases. They had full access to the cases throughout diagnosis, so there was no memory burden to drive fault-driven adaptation, and most of them had completed some sort of representation of the original cases – some did so prospectively, some in response to a memory burden, and some in response to the task of teaching. Would students invest the time needed to build upon their representations to tackle the new problems (prospective adaptation), or would

they follow the routine pattern and work with the materials as given? The graduate students continued to follow the adaptive pattern, as expected. The undergraduates presented a more interesting case. Even though both groups of undergraduates had created representations that could have, in theory, been built upon, only those who made representations under fault conditions (intermittent access) switched to prospective adaptation. They did not simply deploy their already-produced representations. They modified those representations to account for the new information in the novel cases, which led to the signature start-up costs of prospective adaptation. The undergraduates who had only made representations for teaching (continuous access condition) did not build upon their representations, nor did they create new representations from scratch. They solved the problem with the materials as given.

One possible explanation for those undergraduates who switched to prospective adaptation is that they had experienced the value of the representations for helping with the procedural aspects of diagnoses. Even though their initial catalyst to visualization was the memory burden, they may have experienced the representations' value for problem solving, which led them to continue to create representations even after the fault was removed. If this interpretation holds up under further scrutiny, it suggests one way to help promote adaptive expertise. Allow students to experience the benefit of an adaptation, and they will be more likely to use it prospectively. This instructional prescription differs from a teaching approach that favors always giving students the routine solution during instruction, so they never experience the need to adapt. Of course, our interpretation and instructional implications are speculative, because these students may have created representations out of momentum – they just kept doing what they had done before. But, it is informative to note that the undergraduates who only made visualizations for teaching did not continue to make representations for the novel cases out of

momentum. They returned to the routine pattern. Perhaps these latter students only experienced the value of the representations for teaching, and this did not allow them to experience their functional value for problem solving.

## General Discussion

### Measuring Adaptive Behaviors

One purpose for the current research was to develop a metric of adaptive expertise that can be agnostic with respect to what a culture considers adaptive versus maladaptive. The expressions “adaptive pattern” and “routine pattern” should not imply that adaptive is good, whereas routine is bad. Sometimes the adaptive pattern is a bad idea – in high risk situations the implementation dip associated with the adaptive pattern can be dangerous (e.g., driving on the freeway or performing surgery). Likewise, the routine pattern can be entirely appropriate, as in the case of high automaticity for decoding of words.

A challenge with defining the word adaptive is that what one person or group considers adaptive, another may not. So, rather than creating measures of the adaptive pattern that might be normative and potentially contentious, we relied on measures that are used in the study of routine expertise (e.g., time to accurate completion). The advantage of this approach is that it may be useful to researchers from different traditions who are interested in adaptive expertise but have different explanatory frameworks and values (e.g., cognitive, cultural, economic, education). Using standard efficiency metrics, we found (a) an adaptive pattern, where people stepped away from the immediate task plane, and (b) a routine pattern, where people immediately engaged in the task. We further found that the time spent away from the task plane was replaced by time on a tool plan in an attempt to restructure the situation by creating and introducing organizing representations.

In addition to differences of opinion on the value of a particular adaptation, there is a second problem – even if there is agreement on adaptive value, a specific adaptation’s value will be a function of the evaluation horizon. Sometimes, the appropriate evaluation horizon and criteria for deciding if an adaptation is appropriate will be straightforward (e.g., swimming to safety). But, at other times, the appropriate evaluation horizon is impossible to determine ahead of time. At such times, it is especially beneficial to be able to describe people’s efforts toward adaptation, without needing to answer the impossible question of whether their adaptation will ever pay dividends in the long run. The adaptive and routine patterns, with their focus on people’s initial approach, allow us to categorize efforts toward adaptation, while remaining agnostic to the ultimate evaluation horizon.

Nevertheless, one would hope that the move to the adaptive pattern, though inefficient at first, can yield desirable behaviors in the long run. By moving to the adaptive pattern and away from the press of immediate performance, people can innovate new means to achieve their goals or even create new goals. Csikszentmihalyi and Getzels (1970) described a study where art students had to select from various objects; set them on a table; and then, paint them. Some students decided what their final painting would look like as they selected the objects. Other students selected objects and then let the composition emerge while painting. The paintings of the latter students were rated as more creative by art critics. The authors concluded that these latter students exhibited a greater “concern for discovery,” because they did not pre-figure the solution to their task, but instead, created conditions where new ideas and structures could emerge.

The routine pattern allows for improvement in performance, but only to a point. People following routines will eventually reach a performance plateau (e.g., Ericsson, Krampe, & Tesch-Römer, 1993). There are no guarantees that an adaptation will be successful, but the adaptive pattern offers the potential to go beyond what can be achieved by tuning routines (e.g., see Kirsh's 1993 discussion of super-optimizing strategies). For instance, students who used the adaptive pattern to make lists were no more cost-effective than their peers who followed the routine pattern for the original cases. From this slice of time, the adaptive pattern appears less useful than the routine pattern. However, for the novel cases, the list makers finally reaped the benefits of the adaptive pattern. They became increasingly cost-effective at ordering tests, whereas students who followed the routine pattern reached a sub-optimal plateau. This finding highlights the great challenge of moving to a trajectory towards adaptive expertise. On the one hand, the adaptive pattern is inefficient in the short run, and it offers no guarantees of long-term benefits. The routine pattern has a strong allure given this situation. On the other hand, without engaging the adaptive pattern, people have no chance of achieving greater gains on a longer evaluation horizon.

#### Prior Experiences and Prospective Adaptation

A second purpose of the research was to examine the viability of the distinction between fault-driven adaptation and prospective adaptation. Identifying prospective adaptation may be useful, because it reminds us that people can adapt proactively rather than only reacting. Different experiences may be necessary to develop prospective adaptation. Under identical conditions, the graduate students exhibited prospective adaptation, whereas the undergraduate students did not. The undergraduates stayed within the routine pattern, unless there were fault conditions that made their routine unworkable. This finding raises the question of what led the

graduate students to adapt prospectively. Presumably, the graduate students could have completed the diagnosis task just like the undergraduate students had – following the routine pattern of handling each case as it arose, instead of preparing for all the cases ahead of time.

One possible explanation for the graduate-undergraduate difference invokes some form of enduring personality trait that makes some people more adaptive than others. By this account, experiences in graduate school did not engender prospective adaptation. Rather, a property of the graduate students, which correlated with graduate school selection and prospective adaptation, was responsible. One interesting way to examine this trait-based explanation would be to follow the undergraduate students post-baccalaureate. Recall that 19% of the undergraduates made prospective adaptations compared to 100% of the graduate students. Thus, by the current data, the trait hypothesis should predict that those undergraduates who did not make prospective adaptations (81%) will also not enter a doctoral program that involves high levels of information analysis.

Our preferred explanation for the graduate students' prospective adaptation concerns the nature of their learning experiences in graduate school. Graduate students in the sciences and engineering, such as those in this study, spend many hours creating and using representational forms, often over long timescales, as they work to organize and understand data as part of their research. From this work they have had a good opportunity to gain knowledge of how to construct and use representations forms, as well develop the broader meta-representational competence that includes knowledge of varying functions of representation, representation appropriateness, and so forth (diSessa & Sherin, 2000). For example, in the current study, the graduate students more frequently created matrices and trees, which led to superior cost-effectiveness when ordering tests.

In addition to experience with representations, we suspect there is an additional set of experiences behind their prospective adaptations. The graduate students had probably experienced many fault-driven adaptations while working on data tasks. We suppose they frequently started on the task plane to analyze some of their research data case-by-case. Because their analyses were of direct consequence for their own research (Reeves & Bell, in press), and because they could take a relatively long evaluation horizon, they were frequently driven to the tool plane to fashion a representation that could help answer all the questions their data might hold. Eventually, these recurrent fault-driven adaptations led students to adapt prospectively. In the current study we found evidence of this shift on short timescale: the undergraduates who were induced into a fault-driven adaptation for the original cases eventually made prospective adaptations when the fault was removed for the novel cases. If the experience of fault-driven adaptation is what led the graduate students to rely on prospective adaptation, we should expect the graduate students to show prospective adaptation for other common elements of their graduate experience, and not just visualizing. For example, graduate students who collect their own data eventually learn that is worth the time to prepare a structured, master data file before starting to do analyses, as opposed to building different data files as different analyses come to mind.

Undergraduate students also have experience with representations, but their experience tends to be confined to classrooms, where the work is less independent and more constrained to shorter time periods. Moreover, undergraduate education typically tells students when to make a representation, so they do not have an opportunity to experience adaptation. Our hypothesis is that experiencing the value of adaptation for a category of activities is the progenitor of the

adaptive pattern for similar activities. Further research will need to test this hypothesis experimentally.

### The Time Course of Prospective Adaptation

Given our hypothesis about the importance of prior experiences, what were the likely cognitive processes during the prospective adaptation? One possibility is that the graduate students engaged in an explicit period of deliberation and decision making as they began the task. They may have considered a variety of possible problem solving approaches and compared the expected costs and benefits of each one. The estimation of utility is a plausible hypothesis for many cases of prospective adaptation. When people decide they need to fix their tennis swing, for example, they indubitably make an explicit decision that the future benefits will make up for the time and awkwardness of retooling their swing. However, the graduate students did not exhibit any evidence of explicitly estimating the relative utility of prospective adaptation versus the routine pattern. In follow-up research (Martin, \*), graduate students completed the diagnosis task while thinking aloud. Afterwards, they were explicitly asked why they made representations. As in the current study, the graduate students uniformly exhibited prospective adaptation. Importantly, neither the verbal protocol nor the exit interview contained any evidence that graduate students estimated the functional utility of making a representation. Instead, they stated that they did not really think about it – they just made representations as a matter of course.

Graduate students chose an initial approach quickly and with little deliberate thought, but nonetheless managed to end up with appropriate and effective representational tools. This pattern suggests a story involving two forms of transfer: similarity-based transfer and dynamic transfer (Schwartz, Varma, & Martin, 2008). In a similarity-based transfer, people recognize that a new

problem or situation is like an old one, and they map in the relevant schema, analog, or procedure to approach the problem. The graduate students recognized the diagnosis task as the type of situation that called for representations, despite the surface particulars and novelty of the task, and they made the (relatively) far transfer from their specific domain of expertise to the novel task of medical diagnosis. These students had enough experience with information management to habitually create representations prospectively when problem solving in situations of complex information.

The graduate students' adaptation did not end with the similarity-based transfer. The current diagnosis task differs from most transfer tasks examined in the literature, because students had an opportunity to complete diagnoses multiple times and received feedback on their diagnostic accuracy. Given these more interactive conditions, we found evidence of a dynamic transfer (Rebellow et al., 2005), where students slowly evolved their original representations in response to contextual conditions. About half of the graduate students significantly altered their representations midway through the first phase. Their initial, prospective adaptation placed them in a situation to better "see" the task such that they could adapt their representations to better match its structure. Thus, students who began with lists discovered that the situation had a many-to-many mapping between symptoms and diseases (Schwartz, 1993), and this led them to integrate new representational forms such as trees and matrices.

### Conclusion

Invention and innovation are forms of adaptation. The truism that "Necessity is the mother of invention" has a good deal of merit. As found in the current study, problematic situations compel people to fault-driven adaptations and new ways to solve problems. One of the contributions of this work, however, is to show that necessity is not the only parent of adaptation

and innovation. There was no necessity that drove the graduate students to produce and alter representations. Like the undergraduate students, they could have accomplished the diagnosis task without making visualizations. Nevertheless, they adapted the situation to create a representation that could handle all the cases instead of handling each one independently.

The prospective adaptation put the graduate students in an interesting situation. By habit, the graduate students put themselves into a situation where they could continue to adapt through a dynamic transfer. Because they had made the representations, they could continue to refine them as they learned more about the specifics of this new problem solving context. This gives one possible view of how some innovations arise. By habit or intent, people transform problem solving situations to be more amenable to thinking, learning, and further adaptation. In this transformed environment, processes of refinement, revision, and insight can lead to new creations that were not in play at the outset. This can be contrasted with the routine pattern, where people work with situations as they are and just try to get the job done. Unless the environment provides them with faults to overcome, there is little space for innovations to emerge.

If we return to the challenge of producing more adaptive learners in our schools, it will be necessary to create a set of experiences so that adaptation is valued alongside routine efficiencies. A trajectory to adaptive expertise will depend on a balance of routine, highly efficient skills coupled with a willingness to step away from some of those routines to adapt new ways of doing things. An important standing question is whether designed adaptive experiences within the confines of a classroom context can foster prospective adaptation beyond that classroom. Our current leading hypothesis is that experiences with the value of the adaptive pattern on a long evaluation horizon, for example as a side effect of experiencing successful

fault-driven adaptations,, leads to future prospective adaptations in related situations. Often times, in a moment of frustration, we beg children to just accept that there is a longer horizon; for example, “you’ll need to know algebra later, trust me.” Nevertheless, we assume that children appreciate proof more than proclamation. One way to create the proof is to let them experience the value of adaptation directly, for example, by creating opportunities to experience an adaptive pattern that pays off (Schwartz & Martin, 2004).

In the current study, we examined graduate students whose adaptive work probably included many hours outside of the classroom, for example, when wrangling their own data. Thus, based on this study, we cannot make strong claims about the effects of classroom adaptive experiences on developing a trajectory to adaptive expertise. The contribution of this paper is to suggest ways to measure the adaptive pattern of adaptive expertise, so it will be possible to find out in the future.

#### ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under grant SLC-0354453. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

- Agnew, N. M., Ford, K. M., & Hayes, P. J. (1997). Expertise in context: personally constructed, socially selected, and reality-relevant. In: P.J. Feltovich, K.M. Ford, & R.R. Hoffman (Eds) Expertise in context (pp. 219-244). Menlo Park, CA: AAAI Press.
- Augustine, N. R. (2005). Rising above the gathering storm: Energizing and employing America for a brighter economic future. Washington, DC: National Academy Press.
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. Psychological Bulletin, 128, 612-637.
- Barron, B. (2004). Learning ecologies for technological fluency: Gender and experience differences. Journal of Educational Computing Research, 31(1), 1-36.
- Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 640-645.
- Biswas, G., Schwartz, D. L., Leelawong, K., Vye, N., & TAG-V. (2005). Learning by teaching: A new agent paradigm for educational software. Applied Artificial Intelligence, 19, 363–392.
- Bransford, J. D. & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad and P. D. Pearson (Eds.), Review of Research in Education 24, 61-100. Washington, D.C.: American Educational Research Association.
- Carraher, D.W. & Schliemann, A.D. (2002). The transfer dilemma. Journal of the Learning Sciences, 11, 1-24.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.

- Csikszentmihalyi, M., & Getzels, J. W. (1970). Concern for discovery: An attitudinal component of creative production. *Journal of Personality*, 38, 91-105.
- diSessa, A. A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Meta-representational expertise in children. *Journal of Mathematical Behavior*, 10(2), 117-160.
- diSessa, A. A., & Sherin, B., (2000). Meta-representation: An introduction. *The Journal of Mathematical Behavior*, 19(4), 385-398.
- diSessa, A. A., & Wagner, J. F. (2005). What coordination has to say about transfer. In J. P. Mestre (Ed.) *Transfer of learning from a modern multidisciplinary perspective* (pp. 121-154). Greenwich, CT: Information Age Publishing.
- Dufresne, R., Mestre, J., Thaden-Koch, T., Gerace, W., Leonard, W. (2005). Knowledge representation and coordination in the transfer process (2005). In J. P. Mestre (Ed.) *Transfer of learning from a modern multidisciplinary perspective* (pp. 155 - 216). Greenwich, CT: Information Age Publishing.
- Eger, T., Eckert, C., & Clarkson, P. J. (2005). The role of design freeze in product development. *Proceedings of 15th International Conference on Engineering Design, Melbourne, Australia* (pp. 164-165).
- Eisenstein, E. L. (1979). *The printing press as an agent of change: Communications and cultural transformations in early modern Europe*. New York : Cambridge University Press.
- Enyedy, N. (2005). Inventing mapping: Creating cultural forms to solve collective problems. *Cognition and Instruction*, 23(4), 427-466.
- Ericsson, K. A. (2006). *The Cambridge handbook of expertise and expert performance*. New York : Cambridge University Press.

- Ericsson, K. A., Krampe, R. Th., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. Psychological Review, 100, 363-406.
- Fullan, M. (1993). Change forces: Probing the depths of educational reform. London: Falmer Press.
- Gentner, D., Brem, S., Ferguson, R. W., Markman, A. B., Levidow, B. B., Wolff, P., & Forbus, K. D. (1997). Analogical reasoning and conceptual change: A case study of Johannes Kepler. Journal of the Learning Sciences, 6(1), 3-40.
- Goldman, S., Martin, L., Pea, R., Booker, A., & Blair, K. P. (2006). Problem emergence, problem solving, and mathematics in family life [summary]. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), Proceedings of the 7th International Conference of the Learning Sciences, Bloomington, Indiana (pp. 1088-1090), Mahwah, NJ: Lawrence Erlbaum Associates.
- Greeno, J. G., Smith, D. R., Moore, J. L. (1993). Transfer of situated learning. In D. K. Detterman & R. J. Sternberg (Eds.) Transfer on trial: Intelligence, cognition, and instruction. Norwood, NJ: Ablex.
- Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. P. Mestre (Ed.) Transfer of learning from a modern multidisciplinary perspective (pp. 89-120). Greenwich, CT: Information Age Publishing.
- Hatano, G., & Inagaki, K. (1986). Two Courses of Expertise. In H. Stevenson, H. Assume, & K. Hakuta (Eds). Child development and education in Japan (pp. 262-272). New York: Freeman.

- Hatano, G., Miyake, Y., & Binks, M. G., (1977). Performance of expert abacus operators. Cognition, 5, 57-71.
- Hatano, G., & Osawa, K. (1983). Digit memory of grand experts in abacus-derived mental calculation. Cognition, 15, 95-110.
- Kirsh, D. (1996) Adapting the environment instead of oneself. Adaptive Behavior, 4, 415-452.
- Lewis, C, & Mack, R. (1982). Learning to use a text processing system: Evidence for thinking aloud protocols. Communications of the Association for Computing Machinery, 25, 387-392.
- Luchins, A. S., & Luchins, E. H. (1959). Rigidity of behavior: A variational approach to the effect of einstellung. Eugene, OR: University of Oregon Books.
- Mestre, J.P. (1994, February). Cognitive aspects of learning and teaching science. In S.J. Fitzsimmons & L.C. Kerpelman (Eds.), Teacher enhancement for elementary and secondary science and mathematics: Status, issues and problems (pp. 3-1 - 3-53). Washington, D.C.: National Science Foundation (NSF 94-80).
- Mischel, W., Shoda, Y., & Rodriguez, M. I. (1989). Delay of gratification in children. Science, 244, 933-938.
- Miyake, N., & Pea, R. (2007). Redefining learning goals of very long-term learning across many different fields of activity. Proceedings of the CSCL 2007: Computer Supported Collaborative Learning, New Brunswick, NJ.
- Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 510-520.

- Novick, L. R., & Hurley, S. M., (2001). To matrix, network, or hierarchy: That is the question. Cognitive Psychology, *42*, 158-216.
- Palinscar, A. S. & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. Cognition and Instruction, *1*(2), 117-175.
- Pea, R.D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), Distributed cognitions (pp. 47-87). New York: Cambridge University Press.
- Petrosino, A. J., Lehrer, R., & Schauble, L. (2003). Structuring error and experimental variation as distribution in the fourth grade. Mathematical Thinking and Learning, *5*(2&3), 131-156.
- Rebellow, N. S., Zollman, D. A., Allbaugh, A. R., Englehardt, P. V., Gray, K. E., Hrepic, Z., Itza-Ortiz, S. F. (2005). Dynamic transfer: A perspective from physics education research. In J. P. Mestre (Ed.) Transfer of learning from a modern multidisciplinary perspective (pp. 217- 250). Greenwich, CT: Information Age Publishing.
- Reeves, S. & Bell, P. (in press). Children's self-documentation and understanding of the concepts 'healthy' and 'unhealthy.' International Journal of Science Education.
- Sander, E., & Richard, J. (1997). Analogical transfer as guided by an abstraction process: The case of learning by doing in text editing. Journal of Experimental Psychology: Learning, Memory, and Cognition, *23*(6), 1459-1483.
- Schwartz, D. L. (1993). The construction and analogical transfer of symbolic visualizations. Journal of Research in Science Teaching, *30* (10), 1309-1325.
- Schwartz, D. L., & Bransford, J. D. (1998) A time for telling. Cognition & Instruction, *16*, 475-522.

- Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. In J. P. Mestre (Ed.) Transfer of learning from a modern multidisciplinary perspective (pp. 1-52). Greenwich, CT: Information Age Publishing.
- Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: the hidden efficiency of encouraging original student production in statistics instruction. Cognition and Instruction, 22(2), 129-84.
- Sherin, B. (2000). How students invent representations of motion: A genetic account. Journal of Mathematical Behavior, 19(4), 399-441.
- Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), Handbook for Research on Mathematics Teaching and Learning (pp. 334-370). New York: MacMillan.
- Stevens, R., Mertl, V., Levias, S., & McCarthy, L. (2006). Money matters: The social and material organization of consequential financial practices in families [summary]. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), Proceedings of the 7th International Conference of the Learning Sciences, Bloomington, Indiana (pp. 1088-1090), Mahwah, NJ: Lawrence Erlbaum Associates.
- Strauss, S. (1982). U-shaped behavioral growth. New York: Academic Press.
- Suchman, L. (1987). Plans and situated actions: The problem of human machine communication. New York: Cambridge University Press.
- VanLehn, K. (1996). Cognitive skill acquisition. Annual Review of Psychology, 47, 513-539.
- White, S. S., Mueller-Hanson, R. A., Dorsey, D., Pulakos, E. D., Wisecarver, M. M., Deagle, E. A. (2005). Developing adaptive proficiency in special forces officers. (ARI

Research Report 1831). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Wineburg, S. S. (1998). Reading Abraham Lincoln: An expert-expert study in the interpretation of historical texts. Cognitive Science, 22, 319-346.

Zhang, J. (1997). The nature of external representations in problem solving. Cognitive Science, 21(2), 179-217.

APPENDIX A

In the model underlying the design of the diagnosis task, diseases had certain tests which were always positive, some which were always negative, and some which could be either positive or negative with equal likelihood. Figure A.1 shows an example reference case. All reference cases were of a similar form.

Patient Name	<u>D. Ruiz</u>	
ID Number	<u>73946</u>	
Presiding Physician	<u>Dr. Stanley</u>	
<u>Standard diagnostics:</u>		
Blood Pressure	<input checked="" type="checkbox"/> High	<input type="checkbox"/> Normal
Temperature	<input type="checkbox"/> High	<input checked="" type="checkbox"/> Normal
Heart rate	<input type="checkbox"/> High	<input checked="" type="checkbox"/> Normal
Patient complains of:		
<u>Dizziness, headache</u>		
<u>Standard Blood Screen Results:</u>		
<input checked="" type="checkbox"/> Pos	<input type="checkbox"/> Neg	Anemia
<input checked="" type="checkbox"/> Pos	<input type="checkbox"/> Neg	Elevated white blood cell count
<u>Perform Jacobson Test-Battery Results:</u>		
<input type="checkbox"/> Abnormal	<input checked="" type="checkbox"/> Normal	Reflex test
<input type="checkbox"/> Abnormal	<input checked="" type="checkbox"/> Normal	Visual acuity test
<input type="checkbox"/> Abnormal	<input checked="" type="checkbox"/> Normal	Treadmill test
Diagnosis:	<u>Arthotitus</u>	

Figure A.1. An example reference case.

Table A.1 shows the relation between each of the diseases and the eight medical tests which had positive or negative outcomes. The ninth medical test, Patient Complaints, was designed to be universally non-diagnostic across all diseases. This allowed for a simpler analysis, as well as an opportunity to see the frequency with which participants would order a completely uninformative test. Patient Complaints has been omitted from the table.

Table A.1. Associations among diseases and medical tests.

	Arthotitus	Bronson's Syndrome	Cromwell's Disease	Dendrosis	Ergotaxis	Fresomia	Gendontis	Hypomalia
Blood Pressure	+	-				+		
Temp.	-		+	+				
Heart Rate		+	-	-	-	-		
Anemia				+	+	-	+	-
Elev. White Blood Cell Count	+	+	+	-	-	-	+	+
Reflex Test	-	-	-	-	+	+	-	-
Visual Acuity	-	-	-				+	+
Treadmill Test	-	-	-	+	+	+	-	-

Within the cells in Table X, '+' indicates that the given test is always positive for that disease and '-' indicates that the given test is always negative for that disease. Empty cells indicate that the given symptom is positive and negative with equal frequency.

## APPENDIX B

Given the goal of narrowing the search space as quickly as possible, different medical tests are of different value, and their relative value changes as diseases are eliminated. For example, both the Elevated White Blood Cell Count and Treadmill Test tests are likely to eliminate more diseases than the Blood Pressure test initially, but once the result of Elevated White Blood Cell Count is known, Treadmill Test is of no additional value. Given a set of remaining diseases, it is possible to calculate the expected value of the number of diseases that each medical test will eliminate. The expected value corresponds to the average outcome value of a probabilistic event over many trials, and it is found by computing the sum of the probability of each possible outcome multiplied by the value of that outcome. In this case, we multiply the number of diseases that would be eliminated if the test were positive times the probability that it will be positive, and add that to the product of the number of diseases that would be eliminated if the test were negative times the probability that it will be negative. In this way, the relative value of each test can be calculated at any moment in time, and the optimal test can be identified.

When a participant diagnoses a patient, they order a series of tests. For each test, we compare the expected value of their choice to the highest possible expected value, and we compute a ratio. This ratio, which we will call the Test Ratio, is used in the final formula below. The average of the Test Ratios is a reasonable measure of optimality, but because the set of tests that a participant orders for a given diagnosis may or may not narrow the number of possible diseases down to a single disease, as is required by the goal of accuracy, the final measure of optimality must be weighted to account for incomplete searches. For each diagnosis, the measure we called the Weighted Optimality Ratio (WOR) is computed as the sum of the Test Ratios,

divided by the number of tests ordered, plus the number of diseases remaining in the search that still must be eliminated ( $Diseases\ Left - 1$ ). Equation 1 shows this formula.

$$WOR = \left( \frac{\sum (TestRatios)}{NumberofTests + (DiseasesLeft - 1)} \right)$$

**Equation 1 – Formula for the Weighted Optimality Ratio (WOR)**