Anderson, J. R., & Schunn, C. D. (2000). Implications of the ACT-R learning theory: No magic bullets. In R. Glaser (Ed), *Advances in instructional psychology: 5*. NJ: Lawrence Erlbaum Associates. (1-34)

Discusses the ACT-R theory of memory and information processing. Defines declarative and procedural memory. Refers to procedural as *productions*. Production rules are condition-action units that respond to various problem-solving conditions with specific cognitive actions. Declarative memory can come from either in a *passive, receptive mode* (encoding from the environment) or from *active, constructive mode* (storing the result of past mental computations). Says there has been lots of work trying to determine if *generating* the knowledge is more valuable than being told. His summation of other work is that it only succeeds if there is redundancy of encoding. The generation process produces multiple ways to retrieve the material. Because of difficulties with generation and possibility of mis-generation is could be preferable to tell the knowledge.

Procedural memories are created through *analogy*. This means a goal is required and an example of the solution. Just giving an example does not guarantee that a person can create a production rule. They need to understand the example and to deploy it; they need to see that the example applies to the new situation. He goes further to discuss that storing information once is not enough. It must also be used many times - up to 40 times for the same task over and over.

The rest of the theory goes into details about retrieval and uses various equations to map out the probability of a person using a memory. These equations match charts of multiple people's recall of information compared to how long it's been since they used it and how many times they practiced it before. They fit nicely and include information such as how long it's been since you last used it, how often you used it before, and the strength of the association. So time since use and amount of practice can be reasonably determined; however, strength of association seems rather vague. But it's useful to see the functions that fit because one can see how the first time you practice it has the largest effect and eventually additional practice maxes out its effect. Forgetting is the same way, the better you knew something, the harder it is to forget and after a certain amount of time, the amount you forgot levels off.

Next they try to apply ACT-R to teaching. They spend some time deciding if education's motivation is long-term knowledge to create a better public or short-term success as indicated by in class assessments. They discuss how a lot of research and testing of learning looks at quick to learn items such as mnemonics as indicators of success; however, if a person practiced language on a regular basis, these mnemonics cease to be relevant. If a person only learns mnemonics, then tests have shown that over time a person forgets more. To simplify declarative knowledge is easier to forget than procedures but procedures take more time to build and are very specific until lots of practice in slightly different contexts helps a person build a broader and firmer procedure. Discuss Situated competencies such as Lave's work and how this does not support the idea that broad competences are not possible, rather it argues that narrow competence is easier to acquire than broad competence. Broad generality of application requires a great deal of practice in a broad range of situations. Demonstrated in Anderson, Simon and Reder 1996.

Practice is key, however, it is not helpful if the correct things are not being practiced. The ACT-R theory doesn't say anything about what or how to choose the correct items to practice. That is up to the instructor. Various studies have shown that students who lag in different subjects are missing some basic information to support their understanding of a subject. An example of this is teaching grade school students about the number line. Palinscar and Brown (1984) produced dramatic improvements in student's reading comprehension by teaching students about and having them practice asking questions, summarizing and clarifying difficulties. Apparently these students were not there (physically or mentally) when these skills were taught. This is why tutoring is so effective; a person can give individual feedback and spot deficiencies on a per student basis.

"This implies that there is a real value for an effort that takes a target domain, analyzes it into its underlying knowledge components, find examples that utilize these components, communicates these components, and monitors their learning. Unfortunately, cognitive task analysis receives relatively little institutional support. In psychology, there is little professional reward for such efforts beyond those concerned with basic reading and mathematics. The argument (which has been received from many a journal editor) is that such task analyses are studies of characteristics of specific task domains and not of psychological interest. For experts in the various target domains (e.g., mathematics), the reward is for doing advanced work in that domain, not for analyzing the cognitive structures underlying beginning competence. In education, such componential analyses have come to have a bad name based on the mistaken belief that it is not possible to identify the components of a complex skill. In part, this is a mistaken generalization from the failures of behaviorist efforts to analyze competences into a set of behavior objectives. Thus, there is a situation today where detailed cognitive analyses of various critical educational domains are largely ignored by psychologists, domain experts and educators."

Benezet, L. P., (1935) The Teaching of Arithmetic I, II & III: The Story of an experiment. Journal fo the Nathional Education Association V24, N 8 PP241-244, V24, N 9 pp 301-303 & V25, N1, pp7-8

This series of articles is about an experiment with grade school children where they removed all maths from the curriculum until 7th grade. By all math he means memorizing multiplication tables etc… The lessons focused on Reading, Reasoning and Reciting (three R's). Reciting meant expressing themselves about what they'd read – not repetition. The author tries to claim that math is damaging and that students should only be learning how to read. For his evidence he goes to classrooms and asks questions like half of a pole is stuck in mud and water. Half the pole is in the mud. The half that is not in the mud is 2/3 in the water and only 1 foot of the pole is in the air. "How long is the pole?" When he asks the students who've been in his curriculum he begins by saying, "How would you go about figuring this out?" A discussion begins about what to do without numbers. Then they eventually work it out through discussion. When he asks the other groups he never says how would you do this. He just stops at "How long is the pole?" Students begin throwing out numbers and he praises students when they get wrong answers (not saying explicitly that it is right but giving expressions of pleasure in their answers). Then when one girl stands up and points out how to do it or discrepancies in previous answers he frowns at her and tells her to prove it. She does. I'm not saying that his curriculum is without merit; however, his carefully laid out evidence does not provide any support. He is teaching metacognitive processing in his new curriculum while the other is teaching rote memorization and discourages engagement of any other sort. So he has created something useful but does not indicate why. His conclusion simply tells a story of asking a question to some 8th grades that they reasoned out well (this is the whole class so who knows how many actually are doing this) and then he read them the responses he got from the same grade five years before. The students made fun of the other classes reasoning and picked out the errors (this is after they've successfully solved it with his leading questions).

Berardi-Coletta, B., Buyer, L.S., Dominowski, R. L. and Rellinger, E. R., (1995) Metacognition and problem transfer: A Process-Oriented Approach *Journal of Experiemental Psychology: Leanring Memory and Cognition*, Vol 21, No 1, 205-223

Interviewing students and asking them to verbalize their solution process can affect their problem solving. The paper is trying to narrow down what actually is affecting the student's problem solving during interviews. Good lit review about interview/verbalization effects. Carefully describe metacognition as not just metacognitive knowledge (knowledge of one's own self as a problem solver) but actual processing as you solve a problem. They believe that the meta processing questions help students solve the problem and help them learn enough to transfer what they've learned to a new problem. They use the tower of Hanoi problem and then the Katona's Card Problem so it may be a slightly different type of problem but sounds like fairly good problem solving – shorter, less steps. They required people to actually solve the problem within certain time limits or they were thrown out of the study. They started with five groups: Silent, Think-aloud, problem solution (questions about goal, rules and state of the problem), if-then (if I move here this will happen) and then metacognitive (how are you deciding your next move, how do you know this is good). The second experiment had silent, problem and metacognitive and the third only had metacognitive without verbalization vs. silent to see if it's just the thought process or the actual act of verbalizing a response. Their data shows that it's the thought process rather than verbalization. Overall biggest impact is metacognitive and least is silent or think-aloud. However, there is also the problem group, which fall between (but not statistically significantly different than either of the extremes – silent/think—aloud vs. meta processing. My survey is mostly problem oriented or think-aloud.

Tower of Hanoi: Three wooden pages are anchored 3 inches apart. There are six discs, ranging form 1.5" in diameter to 4 inches. The goal of the problem is to move the pyramid of discs from the start peg to the goal peg in as few moves as possible. First, you can only move one disk at a time. Second, you can never place a larger disk on top of a smaller disk.

Katona card problem: Eight cards are dealt. The first card will be dealt face up onto the table, the next card will be dealt, face down, to the bottom of the deck. The next card will be dealt face up onto the table, the next card will be placed face down on the bottom of the deck, and so on, until all the cards have been dealt face up onto the table. You are to figure out the order in which the cards have to be arranged at the outset so that as the cards are dealt, they will appear Ace, 2, 3, and so forth.

Timing: Think-aloud and silent groups spend about the same amount of time per move to solve the problem. The final task did not involve any talking yet the process groups still spent more time per move. There is a significant difference in time spent for process & problem groups compared to the control groups 2-3 times as long. Makes sense because more thought is being required other than solving problems because the student also has to think about answers to the questions they are being asked plus time for the interviewer to ask question. What I really find interesting is that it does take longer per move for the process groups on the transfer task since they are not talking. Have they been trained to think process ideas or is it to do well on the problem (which they all do better than the control groups) one needs to be thoughtful about each move. Actually they could be the same thing. The control groups could just be making moves to see what works while the process groups have made some sense out of what works and why some are trying to think about these ahead of time. The total time to solve the transfer task was less for the process groups because the time graph is per move. So more time per move because they are carefully thought out and I'd imagine many of the silent or think-aloud subjects are doing trial and error. Probably have some routine to their trial and error but may not even know it. I'd like to see them ask all subjects at the end to write a description of the successful way to attack the Tower of Hanoi problem. It'd be interesting to see if the process groups had thought out plans while the others may not have plans or at least are not able to verbalize what it is that they are trying to do when solving the problem.

In Experiment 3 they say they gave the subjects 6 seconds to think about their answer to the meta-process questions because it was the mean time per move in Exp 1. Experiment 1 had times of 14 and 15 seconds per move during training. It wasn't until the transfer task where they were not asked these questions that they went down to 6 seconds!

Experiment 4 shows that the think-aloud group does take longer than silent group did for final task where they are not talking. So think-aloud hindered their learning? Turns out groups are so small that even the think-aloud group taking over twice as long is not statistically significant. Why even report this data?

Some inconsistencies:

1. Experiment 3 says they give students 6 seconds to think about metacognitive questions because it is the mean from Metacognitive group in Experiment 1. Actually Experiment 1 has a mean of 6 seconds for the transfer task where the students are not asked questions or given any instruction. The times per move for the training tasks where they are asked questions is 14 and 15 seconds.

3. They have started talking about total time to solution and comparing it to the first experiment but Exp 1 never shows total time or discusses total time only time to solution.
2. Figure 7 shows times for metacognitive group that should be times for silent group. Either that or all the statistics in the body of the paper are backwards but that would ruin all their conclusions.
3. They say Figure 8 is the same result as in previous experiments but I plotted previous over the top and they are not. The metacognitive group time per move actually matches the problem group (not metacognitive) in previous study. (This is because they did not give enough time for students to think as mentioned in #1 above) Silent is the same shape as previous Silent group in Exp 1 but shifted down 1 second. They spend several paragraphs talking about the significance of how the metacognitive group spends more time per move during training but then they do not in the transfer task. They required them to sit there for 6 seconds between each move plus they took the time to ask their question. The difference in time per move between the two groups is only 5 seconds. So the extra time per move is less than the time to ask their question plus the forced break. I'm not saying that the metacognitive group didn't' do better in the end, they did, but the authors state a whole bunch of things that don't fit their data.
4. Exp 4 only has 15 students in total split into 2 groups and not all of them even solved the problem. They didn't have enough students so they didn't' want to throw anyone out.
5. They state a couple of times in their conclusions for this Exp that it shows similar results to Exp 2 "both in terms of ability to solve at all and in trials to solution given the ability to solve" How can they say this when anyone who couldn't solve it successfully in Exp 2 was thrown out?

Final paragraph "This implies that problem solving, in general, has to be viewed in terms of processing skills, not the content of one's knowledge base. ...Information processors that are continually acquiring data in more or less efficient ways, the efficiency being determined largely by the presence or absence of metacognitive processing." There is more than knowledge base and metacognitive skills that help people solve problems, however, I have to agree that the efficiency of building the knowledge base and their ability to solve problems is improved greatly by strong metacognitive skills. I can't agree that being a good problem solver requires good metacognitive skills though.

Bunce, D. M., Gabel, D. L., & Samuel, J. V., (1991). Enhancing chemistry problem-solving achievement using problem categorization. *Journal of Research in Science Teaching, 28,* 505-521 The effects of an explicit problem-solving approach on mathematical chemistry achievement. *Journal of Research in Science Teaching, 23,* 11-20.

This paper describes a study in which the researchers implemented a curriculum focused on teaching general chemistry students how to solve problems. The students were trained to follow a series of problem solving steps with hopes that they would improve their ability to successfully solve mathematical problems in chemistry.

Results showed no improvement in problem solving success with the trained students. Furthermore, nearly one half of the students reported that the problem solving steps were too time consuming. Only 24-44% of students actually implemented the problem solving steps on exams.

Bunce, D. M., Gabel, D. L., & Samuel, J. V., (1991). Enhancing chemistry problem-solving achievement using problem categorization. *Journal of Research in Science Teaching, 28,* 505-521.
In this study, the researchers teach students how to categorize mathematical problems in chemistry (for example, "stoichiometry") and then examine how this categorization affects their success in problem solving. The results show that teaching categorization skills does not alter a student's ability to solve single-concept problems in chemistry. However, it does increase a student's ability to solve problems with more than one concept, and it enhances their achievement on unannounced evaluations. The paper has an extensive introduction and some good arguments for conducting research in problem solving.

Ceci, S., Barnett, S., and Kanayak, T. (2003) Developing Childhood proclivities into adult compentencies. Chapter 3 *The Psychology of Abilities, Compentencies, and Expertise* Edited by Robert J. Sternberg, Elena L. Grigorenko. Cambridge University Press.
References twin studies. Say twins that are separated at birth do diverge during adolescence based on environment; however, after teenage years slowly converge to same IQ level. Theorize that once they are away from their parents, their genetics determine the environmental choices that they make so that they settle to the same level in adulthood. Lots of arguments about multiplier effects where once you gain certain knowledge, you're more comfortable to seek an environment where you can use that knowledge and learn more. Also talk about how the system is chaotic so a slightly different starting point will result in very different results. Lots of arguments and not much support.

Chi, Michelene T.H. 2006 Two Approaches to the Study of Experts' Characteristics in The Cambridge Handbook of Expertise and Expert Performance edited by K. Anders Ericsson, Neil Charness, Paul J. Feltovich and Robert R. Hoffman. (21-30)
Summary of definitions and studies of skills and shortcomings of experts. Defines Novice, Initiate, Apprentice, Journeyman, Expert and Master. Defines two areas of study that of absolute expertise and the relative approach – the comparison of experts to novices. In some places talks about exceptional experts and routine experts. To me in most places these routine experts are really just people who can solve exercises in their domain. Discussion on experts skills and shortcomings. Ways experts excel: 1. Generating the Best; 2. Detection and Recognition; 3. Qualitative Analyses; 4. Monitoring; 5. Strategies (forward etc...); and 6 Cognitive Effort (easy). Ways experts fail: 1. Domain-limited (really only discusses recall); 2. Overly Confident (contradicts discussion under monitoring); 3. Glossing Over (could be considered a strength about filtering relevant information); 4. Context-Dependence

within a Domain (Just shows the context makes up part of the necessary knowledge); 5. Inflexible (performance mode); 6. Inaccurate predictions, judgment and advice (contradicts same stuff in monitoring again. Separate skill. Also too much of an expert); and 7. Bias and functional fixedness.

The descriptions of these studies gives me the impression that the people they are studying are quite good at what they do but does that make them expert problem solvers? From the results of their studies that indicate that they are not as good as novices at particular things such as context dependence, inflexibility, and functional fixedness seem to all be studies focused at identifying their knowledge outside of their area but in some cases the questions are close enough that they put these experts into performance mode or simply just test the limits of their knowledge. Would like to read many of these to see if experts are always doing these things or if it's just on average they were worse. Claims such as domain-limited is based on studies of memory which is not problem solving. So quite a big claim based on the research cited.

Chi, M. T. H., Feltovich, P.S., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cogntitive Science*, 5, 121-152.

Several studies to verify their hypothesis that experts categorize problems using physics principles while novices use surface features of the diagrams. Diagrams were not given with the problems. The "problems" were typical back of the chapter tasks. The idea is that everyone forms a *problem representation*, "cognitive structure corresponding to a problem, constructed by a solver on the basis of his domain-related knowledge and its organization", when solving a problem. The authors say that a person takes the info from the problem and attaches it to a *canonical object frame* (info you already know) that "augments the information about an object stated in a problem with associated information from the knowledge base."

The first study had 8 novices (students who'd finished one semester of intro mechanics) and 8 experts (graduate students). These 16 people categorized 24 problems any way they wanted. Experts took 30% longer to sort the problems. First a discussion of 4 problems for each the experts and novices. This discussion does not include the problem statements, only a diagram that could be used with the problem. Remember that the diagrams were not supplied with the problems to the subjects. Analysis is done by showing the grouping, # to have used that grouping, average number of problems in the group, total problems categorized that way (# who used the group * average # of problems per group) Most discussion uses this total number of problems categorized in this group. I think it implies things that are not there. They say that experts use the physics principle to describe the category. From reading the categories that are supplied, this is truer for the novices but not completely. A large factor could be that the experts have a more consistent grasp on the vocabulary. Novices are not blatantly using surface features. The most notable categories for novices include a category of inclined planes, pulleys and don't know. The experts have a category called vectors. Also, note that the authors do not mention how accurate the categorization is for either group. Notably the number of experts who choose particular categories were 6,6,6,6,5,4,2,2,2,2,2. The above groupings only accounted for 76% of the groupings. The novice groups were chosen by 7,6,4,2,2,4,5,4,2,3,2,4,2 and accounted for 73% of the groupings. To me this demonstrates that eight subjects are not enough to see repetition and clear patterns.

Study two has one novice, one expert (grad student), one professor and one intermediate (4^{th} year major). In this study 20 problems were chosen because the surface features were crossed with applicable laws. These were HRW problems or problems that were devised by some "Electrical Engineering major with substantial physics experience". This means the diagram could be the same but a different principle is more efficient for different problems; however, the example problems shown could be solved with two of their three possible principles. There were several problems noted that this was the case. The subjects were asked to classify them. Notably the expert classified three problems wrong. The physics professor used 6 categories (based on solution procedures) instead of the three that the authors' claim existed (the grad student also used three). The "advanced novice" used both words that described the principle and the physical features.

Neither study had any sort of analysis done to see if the results were significant. I have been able to get stronger anecdotal evidence from my classes of 30 students. These two studies are followed by a long discussion where the authors discuss two possible reasons for the above results. Using these results to support a complete description of what is going on in the solver's mind while sorting problems. They claim that when a representation is chosen, this restricts the rest of the problem solving procedure. Why didn't they let the students solve the problems after they categorized them? Then they could see if they used these principles the whole way through or if they figured out the right one by the end etc… Or maybe interviewed the students to ask why they choose the categories that they did choose. Either of these steps wouldn't have added much work since they only used a few subjects.

Third study had two experts and two novices. They were given 20 category labels and were asked to say everything they could about problems involving each label. They were limited to 3 minutes per category. The experts were more likely to mention solution type processes first followed by important physics features. Novices were more likely to mention and discuss only the important physical features.

Study four had two professors and two 'A' students from Intro Mechanics. The same 20 crossed (trick) problems were used from study two. The subjects were asked to explain their basic approach to each problem and what features cued this information. The experts gave physics principles that they would use. The results were more consistent and more of what they had expected would happen with the sorting task. There were three problems that were not identified as expected. (Authors call this almost perfect.) One of the experts also identified the problems after reading on average only 20% of the problem. The experts also identified principles based on derived knowledge (something not stated in the problem explicitly – something basically part way through solution). The novices thought basic approach meant what they would do. So they said "first, I figured out what was happening…then I, I started seeing how these different things were related to each other…" The novices did not give their basic approach until they had read the entire problem in all cases. So these were exercises for the experts and problems for the novices.

In summary, entire article demonstrates the differences between solving exercises and problems. No information from these studies can be gained on actual problem solving procedures other than possibly the way a person first identifies what to try. Very little is given on this because there were no specifics given. Students also never actually engaged in solving the problems. The "experts" were sometimes grad students and sometimes professors. Some of the novices were top undergrads so could have been great "problem solvers".

The surface feature versus deep structure of Chi is interesting but what I think is more interesting is that everyone did it by topic. I think we (anyone who's ever had a physics course) are thoroughly trained to sort physics problems this way by the text books that present things sorted by topic. The student who sorted by problem difficulty may have had another goal in mind. What if you gave the problems to professors during finals and ask them to sort. If they've been creating an exam, maybe sorting by difficulty would be a very useful thing also.

de Jong, T. and Ferguson-Hessler, M. G. M. (1986). Cognitive Structures of Good and Poor Novice Problem Solvers in Physics. *Journal of Educational Psychology*, 78 (4) 279-288.

Detail of study described in "On the quality of knowledge in the field of electricity and magnetism"

De Leone, C. J. and Gire, E. (2007) Student Spatial Resaeoning and Physics Problem Solving Presented at the American Association of Physics Teachers 2007 Summer Meeting.

17 students in lab given ROT (Rotations test from Bodner and Guay). Correlation between worse score and more concrete details provided in diagrams. One student had 20/20 on ROT and showed no diagram. Problem was Tarzan type problem. There was no correlation between students who showed Tarzan's path in their diagram and their ROT score.

diSessa, Andrea A., Wagner, Joseph F. (2005) What coordination has to say about transfer. Transfer of Learning from a Modern Multidisciplinary Perspective edited by Jose Mestre. Information Age Publishing 121-154.

This work seems applicable to problem solving in several ways. diSessa says it's worth distinguishing parts of the causal net that determine existence and relevance of particular information from those that do the actual information determining. An example of determining trip length from a train schedule is provided. Most people know they can figure out the length of the trip from the arrival and departure times but not all can do it. Then if they have to figure out plane trip lengths that include time zone changes, fewer can do that; but, they do know that the information is there and matters.

Force or acceleration make good candidates for coordination classes while theorems, laws and other intrinsically relational cognitive entities do not make obvious candidates for coordination classes, as they entail multiple kinds of information and relations among them. Coordination class is a concept that provides a way to gather information (i.e. readout strategy) about the world. A well-formed coordination class requires *span* and *alignment*. Span is being able to apply the idea across many contexts. Alignment means when it is applied in different situations, the information obtained is the same in each situation. When a person uses different strategies and cognitive operations to apply a concept they call this a *concept projection.* If a person cannot apply the concept in different contexts, with the same or different strategy, then they lack span. The authors believe that it may be quite common to apply a concept using different strategies in different situations (does not have to be done with the same strategy in each context). If alignment has been achieved, then concept projections in different contexts will reliably determine the same information.

There are only two cognitive processes they feel are necessary for constructing a coordination class: *incorporation* and *displacement.* Incorporation is tying in new experiences and displacement is dismissing prior knowledge that may not fit in a particular context but is the initial idea that comes to mind when first thinking about the problem. There is also a difference between determining the existence and relevance of information (*read out strategies*) from doing the actual information determining (making inferences and sometimes predictions) (*causal net*). A person must successfully read out the information needed and then use the causal net to determine force or duration from a particular situation. Finally, if needed, predict an outcome based on these inferences. They point out that it's useful to determine if learning difficulties lie in the causal net or in the readout strategies.

An excellent point is given about the difference between readout strategies and the causal net. Metz (1993) shows that children understand that certain information is relevant before being able to make adequate use of it. "The distinction between inferences of existence and relevance on the one hand, and specific inference of information on the other, generalizes this developmental observation. i.e. A person may recognize a conversation as being in French, but cannot understand it because they do not understand French.

An example of an incomplete coordination class was given with understanding words. A person may be able to understand when they hear a word but not be able to read the same word. There are other examples and discussions throughout the paper that demonstrate that building a well-formed coordination class requires the acquisition of many skills; in this case, reading. With the train schedules the *causal net* must include knowing that one can and be able to subtract two times to infer a duration.

Another good point is in reference to conceptual theories that place naïve concepts even handedly with well-developed expert concepts. A naïve concept would not qualify as a coordination class. Studies show that these naïve concepts lack span and alignment. They can only be applied in limited contexts and don't always determine the same information (force in one case and acceleration in another). Instead naïve ideas become a part of the cognitive structure that eventually produces a "high-class" concept or a well-formed coordination class. Authors propose that an application of the coordination class theory is to provide a descriptive frame for tracing the development of a concept as it begins to meet the criteria of a coordination class.

When discussing the use of coordination classes says that the processes of determining information is most likely mediated through other coordination classes. For example force, mass and acceleration would each be their own coordination class. Mass and acceleration must be determined and then multiplied to determine force. Empirical work has demonstrated that coordination classes feed off of each other and are formed simultaneously with other relevant classes. "Overall it seems that a particularly important mode of learning is the mutual bootstrapping of a cluster of coordination classes. However, we have at present no generalities to offer about coordination clusters."

A coordination class is only a fully constructed, complete, piece of knowledge. diSessa says that what we need is a theory of coordination without the emphasis on class; an understanding of partially formed coordination classes, what is missing (i.e. alignment) and how to fix it. When looking at completeness, diSessa mentions the possibility that a layer of meta-knowledge may allow one to generate new concept projections so that the class becomes adequate for all contexts. "…it seems certain that experts do not "have" all relevant concept projections, but, rather they have resources to *generate* some of them at need."

Surrounding this entire discussion of coordination classes is transfer and discussions of why research does not see it as often as they expect. He begins by defining transfer as three Classes, Class A, B and C. *Class A transfer* is the type of transfer that is presumed in many psychological experiments. It assumes that knowledge is *well prepared* and does not require further learning to apply. This means that it can be tested in relatively short periods of time. *Class B transfer* may not occur in short periods but presuming persistent effort, knowledge is *sufficiently prepared* to be applied (in a few hours or days in real-world problems) using learning and other resources that might be available. *Class C transfer* is how *relatively unprepared* subjects use prior knowledge in early work.

Many theories say that when something is learned the main concepts have been abstracted from the examples and a person is only left with the concept. This work notes that abstraction is a central part of other transfer work but coordination class uses a different mode for wide applicability, which is where an extended range of context-specific capabilities are collected and coordinated.

Conclusions about blocks to transfer are quite solid. Fluent transfer is precisely the expected outcome for an adequately constructed coordination class. It is just that preparing such a concept is difficult, even if the difficulties are, in some cases, transparent. Accumulation and coordinating a wide range of knowledge into a coordination class simply cannot be easy or quick. Many psychological experiments that have found little transfer were expecting something that should not be expected. Given the limitations of the methods used in many transfer studies, little can be concluded from subjects' demonstrated success on a transfer task. On the flip site Type C transfer is quite common but blind and unreliable.

Having a concept might be impossible to assess in a single context and there are many states of partial construction that interact in highly specific ways with particular contexts. I need excellent assessments of partial states of knowledge construction before questions can be answered about when and how students do/do not use particular ideas.

When considering that it would be nice to know if the same information is always read out from different situations using the same knowledge, diSessa notes that "determining the actual knowledge used by an individual, rather than the output of that knowledge, is an extremely difficult methodological problem, especially with the primary data concern correctness." The authors do not expect the same knowledge to be used. When discussing schema theory and knowledge abstraction they point out that those theories place the ultimate value on analogical comparison, which requires construction of parallel representations of different situations. Coordination class theory would only predict this ability to create parallel representations of different things at the highest level. Meaning only after a rather well-formed class has been created. Rather coordination class theory expects students to use different strategies (concept projections) to 'infer' the same information from different situations. Instead of abstraction, students need to compile more knowledge (by learning how to use the concept in different contexts) so that they have a larger knowledge base to draw from as appropriate for different situations.

As his discussion of "By what processes is transferable knowledge created" he begins to touch on a few specific process skills while considering possible explanations resulting from empirical application of his theory. These include: perception of the relevance of particular aspects of a situation or even of a coordination class itself; different cognitive demands entailed by different contexts; resources adequate to *generate* new concept projections; he can see that a coordination class is relevant but either can't determine the relevant information or determines the incorrect information (failure of alignment); confidence

As with Anderson, diSessa calls for more research on analyzing successful students implementing a coordination class across context to a depth that distinguishes, as an example, different concept projections. However, he realizes that "It is much easier to see a distributed element when it first appears and feels to the subject like an insight than when it is routine and fluidly invoked." (my comment: declarative memory vs. procedural) Also states it would be very valuable to know how different learners might overcome failures, such as not recognizing the relevance of a concept or unable to determine correct information, and how different such solutions vary from learner to learner. "…we need much more information about particulars than can possibly be carried in global judgments of "has it, or not." Only when excellent assessments of partial states of knowledge construction are available will questions concerning when and how students do or do not use particular ideas be much more transparent."

diSessa suggested that a theory of coordination is what is needed and to determine what is missing in the coordination cluster for it to become a well-formed coordination class. He suggests alignment or span might be what is missing. There are two specific processes mentioned in reference to how the coordination class is formed; however, the theory doesn't evaluate what processes are used when applying a coordination class. This would tie in nicely with understanding what is missing in a partially formed coordination class. The analysis by Mestre et al. and another by Wittman is using the theory to discuss how students evaluate a particular situation. It seems quite awkward because they discuss what students do using this vocabulary of knowledge, yet the theory does not include processes other than described above for constructing a class (incorporation and displacement). Later in the paper diSessa does touch on some processes such as perception of relevance, confidence or resources to *generate* new concept projections. Other than these little bits, the main focus is strongly about knowledge and that it's knowledge that matters not how one uses it. I really think both are necessary and I'm not sure why he doesn't even mention this.

When applying coordination classes empirically, diSessa & Sherin had one student; Wagner always talks of one but mentions more than one at the start; and then Mestre actually has multiple subjects. diSessa & Sherin hypothesize that students' weaknesses are found in the causal net rather than in readout strategies. He notes that Mestre et al. were just as likely to find problems in the read out strategies. While discussing possible reasons for this, he doesn't even conjecture that it's because he and Sherin had one student! I find it surprising that he considers something as complicated as forming a thorough understanding of force or other concepts to be understood through the careful observations of one student. When teaching and working with students, it seems clear that their misunderstandings and weaknesses vary.

I also think it's awkward to ignore the division of declarative and procedural knowledge since these are distinct types of memory. Not only are these different types of knowledge, but they are stored and accessed in the brain differently.

Docktor, J., Heller, K., Heller, P., Thaden-Koch, T. and Li, J. (2007) Robust Assessment instrument for Student Problem Solving. Presented at the American Association of Physics Teachers 2007 Summer Meeting.

She's working on a Rubric not an instrument. Designed to evaluate written work such as Homework problems or worked examples. She used experience and inexperienced as her definitions rather than experts and novices. Spoke with her outside of her talk and she seems quite reasonable. Frustrated with people asking her why her rubric doesn't have this skill or that. She's very clear on the fact that it does not cover all skills, only the ones that Minnesota focuses on teaching. What she has made is specialized and does a very thorough job of evaluating a specific type of problem.

Dweck, Carol S. (1999). Essays in Social Psychology. Self-Theories; Their role in motivation, personality and development. Psychology Press: Philadelphia, PA.

Performance Goals vs. Learning goals.

Gave a bunch of 5^{th} graders the same task. Half were told they would be evaluated on how they did and the others were told they would learn useful things from the task. Many students with performance goal showed a clear helpless pattern in response to difficulty. Some condemned their ability and the problem solving deteriorated. Most of the Learning goal students showed a clear mastery-oriented pattern. They did not worry in the face of failure and remained focused on task.

Next they took half the performance goal students and told them they had high ability in this area and told half that for now they had low ability in this area. They did the same with the learning goal group. For students with performance goals, those who were certain of their high ability held on in the face of difficulty. The students who thought their ability was lower fell into a helpless response. For students with learning goals the message made no difference.

Helpless vs Mastery response:

Dweck also studies students helpless versus mastery response. She used a questionnaire that can identify which students show persistence in the face of difficulty. The two groups were given 12 problems. 8 they could solve and 4 that were much above their level. The two groups showed equal ability on the first 8 problems. When the helpless group got to the hard problems they gave up, said disparaging things about their abilities and then said they couldn't solve the first 8 again. They also tried to change the subject and talk about their successes such as: I'm going to be an heiress or they'd change the rules, I'm picking brown because I like chocolate cake; preschool level responses. The mastery group kept persisting but could not solve the final 4. When asked if they could solve the first 8 again they said of course and some thought it was a ridiculous question. Another study had workbooks that taught students some set of ideas. Then tests were given at the end. Students who missed questions were told it looked like they needed to review and then try again in an attempt to take all the performance pressure off.

There was a second set of booklets that had a confusing passage in the beginning. It was a passage about imitating people and it had nothing to do with the content of the lessons. It was written in a very confusing way. The students were then tested on these booklets. The helpless and mastery students scored the same on the non-confusing booklets 76.6% and 68.4% (not statistically different). One the confusing lessons the helpless group scored 34.6% compared to 71.9%.

Fixed versus Malleable Intelligence

The idea of fixed intelligence or entity theory is that intelligence is a fixed trait. A person has a certain amount and that's that. Malleable intelligence or incremental theory says that a person's intelligence increases as they learn new things. They are not claiming that there are not differences among people in how much they know or in how quickly they learn things. But that everyone is capable of learning more.

This idea is the basis for most of the behavior that she has observed. Students who are entity theorists will act helpless in the face of a challenge while incremental theorists will just keep trying. There are studies that show if students have confidence then they do better in school but Dweck says that is not always the case. Students who have high confidence but are entity theorists will still fail in the face of a challenge. They do just fine from one school year to the next as long as there isn't a big change. However, if things get much harder, they fall apart; whereas, students with high confidence who are incremental theorists always do well. Students with low confidence who are incremental theorists often bloom in junior high or high school. One possible reason is that the new challenges not only don't seem insurmountable since they know they can learn but they also have low beliefs about the skills they currently have so they believe they need to put in some effort to increase these skills.

During a study with college students they created two fake pyschology today articles about an 8 month old named Adam. In one article Adam was described as a genius because of his unusual level of vocabulary and attention span. In the next article his behavior was explained as being a result of his environment. Students who read the entity theory article were very unlikely (13.3%) to choose to take a tutorial about some problems they just did poorly on but the students who had read the incremental theory article wanted to take the tutorial (73.3%).

She studied grade school, junior high and college students; she found in each instance that student's theory of intelligence entity or incremental, correlated with their choice of goals. Students who were entity theorists emphasized performance goals and were more likely to blame their failures on low ability, feel distressed and ashamed about their academic performance. They also found that the entity theorists under achieved what was expected based on their SAT scores and that after four years, their self-esteem was lowered.

Discusses "stereo-type" threat which was defined by Aronson, Quinn and Spenser 1998 and Steele & Aronson 1995 and Steele 1997b, a study by Steele 1997a and Aronson (Steel & Aronson, 1995) where African American student response to stereo type threat is studied. Traditionally these students perform very poorly compared to their Caucasian counterparts with equivalent college board scores. (Aronson and Fried, 1998) studied students at Stanford. They reason that the transition from high school to college, especially a predominantly white college such as Stanford can be a very challenging one. They showed a short film to both African American and Caucasian students presenting scientific explanations, researchers' testimonies, neurological graphics and research findings to the effect that every time people meet a challenge, exert mental effort, and learn something new, their brain grows neurons and they become smarter. The film was accompanied by a lecture and the students were required to write a letter to grade school students explaining their new view of intelligence and how it expands with work. At the end of the term Aronson and Fried compared the grades earned by all students who had seen the incremental film to the grades of students who had not. Of the students who had not seen the film, the GPA's of the Caucasian students was significantly higher than those of the African American students. For the students who had seen the film, the gap between the majority and minority students was appreciably reduced. In addition, the African American students reported enjoying school and were more academically oriented than their peers in the control group.

Ericsson, K. Anders (2003). The Search for General Abilities and Basic Capacities – Theoretical Implications from the Modifiability and Complexity of Mechanisms Mediating Expert Performance. *The Psychology of Abilities, Compentencies, and Expertise* Edited by Robert J. Sternberg, Elena L. Grigorenko. Cambridge University Press. (93-123)

Ericsson frames everything to show that with practice people merge not diverge. He says that they trained students to memorize up to 80 digits rather than the typical 7-9. They say analysis says research showed students created a mechanism for rapid retrieval from Long Term Memory. How do they know that happened and not just improving STM?

After initial discussion of previous theories the article discusses how expert performance is acquired. Says it typically takes 50 hours for a person to acquire an automated level of a skill such as typing or playing tennis. After that point, if your engage in an automate fashion you won't improve. You have to engage in deliberate practice to improve at anything. You have to make mistakes to be motivated to improve. It's typically necessary to have a teacher or a coach who can design training that pushes you just over what you are able to do. Independent practice in between is also necessary. Studies show that ice skaters who are improving spend more time on jumps that they can't do while medium level ice skaters spend their practice time with jumps that they already can do; can engage in deliberate practice about 5 hours a day; and need a day of rest to recover. This is why novelists write in the morning and recover the rest of the day.

He argues that because the body can adapt to most things (kidney can grow 70% in a week if you lose one) trying to study grade school or college students is not going to help with understanding of abilities. He needs to study experts if interested in performance limitations. Most psychological tests that are designed to measure ability get continually better results from people who practice them.

Erricson says he hasn't seen evidence that people have limiting capacities and that so far anyone who tries will become an expert in a particular area.

Etkina, E., Van Heuvelen, A., White-Brahmia, S., Brookes, D. T., Gentile, M., Murthy, s., Rosengrant, D. and Warren, A. (2006). Scientific abilities and their assessment *Physics Education Research 2*, 020103-(1-15).

Eugenia spoke at the group meeting about the research that is reported in this reference. They are doing very careful work in teaching labs that have to be designed by the students to demonstrate certain physical concepts. The students are graded on the design and write up of their labs. After a few labs, the students ability to design a carefully thought out lab is greatly improved. They were also writing clear descriptive write-ups that communicated their design and procedures. Students are also encouraged to focus on significance of their answers (experimental uncertainties) but improvement in those areas is weaker. The group has designed a very thorough rubric for grading the labs. Many important planning skills are included with a well thought out scale for grading these skills. The group is somewhat disappointed that the students do not seem to transfer overall better scientific reasoning into the classroom. From the talk it seemed to me that they want to believe that having students plan these experiments improves their reasoning ability in general. I see it as improving the specific skills of planning and monitoring to a point. They are learning how to plan out a solution path and to look at the assumptions that they make when doing this. It does not specifically work on their knowledge base, or their metacognitive skills, or beliefs or qualitative analysis. They do however focus heavily on the use of free-body diagrams. This seems to provide some improvement in problem solving.

Ferguson-Hessler, M. G. M. and de Jong, T. (1987). On the quality of knowledge in the field of electricity and magnetism. *American Journal of Physics,* 55(6) (492-497).

They studied people's organization of knowledge structure as schemata or hierarchical. They contend that explicit use of problem solving schemata will help teachers make their tacit knowledge visible to the students, and consequently, will enable students to realize that characteristics of problem situations are used to determine which principles to apply.

They studied knowledge structure of good versus poor novice problem solvers. Motivated by other work showing that knowledge structure is the difference between experts and novices. Experts say "there is relatively little to remember".

Three types of discipline specific knowledge are defined for use in this paper. *Declarative Knowledge*-ampere's law, *procedural knowledge*-choosing a closed path and *problem situations*-long straight current. They suggest that knowledge organization for successful novices needs to be in the form of schemata that include all three types of knowledge listed above organized together so that various problems of the same type can be successfully tackled. Later on in more advanced courses students will learn enough about the subject to make the connections that create a knowledge hierarchy of electromagnetism found in experts. The research was done by having 47 students sort 65 separate cards which had individual bits of knowledge from the schemata of Ampere's Law written on them. Students were asked to sort the cards into piles where cards in each pile were more strongly connected to one another than cards on other piles. The piles were then analyzed and compared to exam results. The coefficients of correlation were up to 0.54. A separate analysis was performed – "hierarchical cluster analysis" of the cards to independently determine the characteristics of the piles of cards for high performing students 70% and up to low performing students 30% and below. The piles created by the high performers agreed with the problem schemata in most cases and the piles for the lower students showed little agreement with the schemata. Finally they had students label each pile. The results here were consistent with Chi's work. The good novices labeled using things such as "related to induction" while the poor novices used labels like "containing the word field".

The paper closes with a bunch of discussion on teaching implications. The authors say one should show the characteristics of schemata but not explicitly teach the details since students learn much better if they're forced to reason for themselves. They quote other work that has tried but says no more about it than it met with varying success.

One thing they do say that I disagree with is that students aren't lacking the knowledge only the structure. Their argument for this is that when sorting the cards only 2% of the cards of good students and 4% of the cards for poor were placed in the unknown pile. If categories are sorted with names such as "contains the word field", then they can sort them without any knowledge of the idea.

Ferguson-Hessler, M.G.M. and de Jong, T. (1990). Studying Physics Texts: Differences in Study Processes Between Good and Poor Performers *Cognition and Instruction,* 7(1) 41-54.

This paper looks at students' study procedures. It found that good students, as measured by exams in class, have different methods of studying. In contrast to previous work they found that the both types of students are equally active but that good students do a better job of restructuring the material. They had students read 10 pages of text with red dots spaced evenly throughout. When the student reached a dot, they explained what they had done since the last dot. They were also allowed to take a small amount of notes (1 6x8 sheet). These interviews and notes were coded for types of studying processes: Superficial processing, Integrating, Connecting, and other. Each of the processes was broken into subcategories. The good students had more instances of Integrating and fewer of superficial processing. Connecting and other activities were nearly equal. They were also coded for knowledge type. Poor students had higher frequency of declarative knowledge and lower situational and procedural. Looking at the notes showed similar results but the small amount of notes kept the data from being statistically significant. They also did think-aloud interviews with students while they solved problems. The students had been instructed to study the text and work problems as if they were studying for an exam. The think-aloud data showed an average of 4 problems worked by good students and 2.2 for poor.

Finegold, M. & Mass, R. (1985) Differences in the Processe of Solving Physics Problems between Good physics Problem Solversand Poor Physics Problem Solvers. *Research in Science & Technological Educations, 3* (59-67)

In this study the researchers asked teachers to identify which of their students were Good Problem Solvers (GPS) and Poor Problem Solvers (PPS). They talk about first order knowledge base which consists of facts, concepts, laws, constants and formulae while the second order knowledge base consists the ability to select a suitable strategy and plan a course of action, and the ability to carry out the solution by implementation of the appropriate knowledge and skills. The authors depend on ideas from Polya, Larkin and Reif. Based on the hypotheses contained in this prior literature the authors hypothesize that a good problem solver:

1. translates the problem statements more correctly and more exactly.
2. plans their solutions more fully and in greater detail before carrying them out than do poor solvers who tend to solve without planning.
3. completes the solution to a problem in less time.
4. spends relatively more time on translation and planning.
5. Poor solvers may use physical laws and quantities which are not applicable, possibly on the grounds that they have been useful in other problems that seem similar.
6. makes greater use of the physical reasoning.
7. makes greater use of algebraic manipulations.
8. is more likely to evaluate and check their solutions than are poor solvers.

The research supported 1, 2, 3, 4 & 6 only. The authors say that the study suggests that the work of poor solvers can be improved by emulation of good solvers. I don't see where or how this is suggested by this study. The authors mention a pilot study where they tried this without any success.

The study had 4 teachers who rank ordered their students based on solving skills. The teachers were told to make sure that all students had first order knowledge and then to rank them only on their solving skills. GPS had 90% or higher final class grade while PPS had 60% or lower final class grade. The study assures us that the students all definitely had first order knowledge. They never state how they can be sure of this.

The teachers and students were then asked to volunteer for the study and the researchers chose 8 GPS and 8 PPS to study.

Fisch, S.M., Kirkorian, H. and Anderson, D. Transfer from TV. *Transfer of Learning from a Modern Multidisciplinary Perspective* edited by Jose Mestre. Information Age Publishing 371-393.

Discussion of transfer from the perspective of what children comprehend from TV. Nicely written and seems to come to many solid consistent conclusions. Says that children's ability to transfer material depend on three things.

1. Comprehension of the educational content in the program
2. The nature of the viewer's mental representation of the content.
3. The transfer situation – the novel problem or solution to which the content is applied.

Their reasoning for the above fits well with the idea of coordination classes. There's info about what affects comprehension such as the narrative being similar to the educational content (cognitive load) and that repetition of the material is useful. Repetition can be exact repeats or showing the same educational content in terms of different narratives. The authors say "Through repeated practice in a variety of different contexts, the mental representation of the underlying content is forced to adapt in subtle ways to each new context, yielding a representation that gradually becomes more detached from the specific contexts presented, so that it can be applied more easily in new situations.

Foster, Tom

He created a problem solving rubric for the context rich problems in Minnesota. Rubric is very simple: Was their physics approach appropriate (right principle), Did they use the right variables and apply the formula correctly, Did their steps follow one another and finally was their math right? Students were graded down if they used means-end analysis.

He also spent a lot of time developing 21 features for grading problem solving difficulty. If a problem has 5 of them it's the right level. This was a very detailed analysis but still somewhat superficial. Approach: Cues, Choice of Principles, Non-standard application; Analaysis: Excess or missing info, vague/special conditions, more than two parts/five or more terms per equation; and Math: Algebra required, cal or vector algebra.

Gick, M.L. (1986). Problem-solving strategies. *Educational Psychologist,* 21, 99-120. Review of problem solving literature. Strongly states means-end is novice. Review is nicely laid out and sounds quite solid unless you have actually read the material being referenced. In all cases that I'm familiar with the literature, Gick states the authors' hypothesis as supported fact. In some cases going beyond the original articles' hypothesis.

Gick states that Chi et al "have shown that whereas novices' schemata for physics problems are based on superficial similarity, experts' schemata are based on solution principles." In my opinion card sorting is not a task that "shows" that students even have schemata associated with these physics problems much less telling us something about those (if they exist) schemata.

When discussing Experts and Novices states with complete assurance that novices use means-end analysis because their schemata are based on similarity of objects while experts schemata are based on solution procedures. In physics experts always work forward and novices work backwards. "This sequence is required because their schemata are based on similarity of objects"

Gick describes a problem solving process where a solver constructs a representation, then if a schema is activated, solves the problem. If not, then they search for a solution and then implement a solution. She says this is a key difference between novice and expert solvers. Just looks like the difference between and exercise and a problem to me.

Gick's conclusion about all expert novice studies including those in computer science and those in political science that experts have organized structures such as schemata that allows for use of sophisticated strategies. Even though much of the research she references outside of physics speaks of specific skills such as decomposition or effectively breaking a problem in sub goals or isolating a factor and attempting to eliminate it etc…

Further parts of Gick restate all speculation at an equal level to research. There doesn't appear to be any actual consideration about whether different studies fit together or not. In fact, there are places Gick contradicts herself because she is discussing different articles that have different speculations about the research.

Gick does discuss general skills as existing but mentions that often a person cannot use them successfully without a good understanding of the domain specific knowledge as well. The general skills listed (called strategies) include analogy, decomposition, breaking a problem into parts, planning and error-detection heuristics. "Further research will determine the general strategies that are domain independent, those that are facilitated by domain-specific knowledge, and the interaction of domain-specific knowledge and general strategies. This work should be given a high priority because of the important theoretical concerns surrounding the role of general strategies and domain-induced efficient processing (Block, 1985; Glaser, 1984; Newell, 1980; Sternberg, 1985)."

Greeno, J. G. (1980). Trends in the Theory of Knowledge for Problem Solving in *Problem Solving and Education: Issues in Teaching and Research* edited by Tuma, D. T. and Reif, F. Lawrence Erlbaum Associates; New Jersey.

"A person may not have learned exactly what to do in a specific problem situation, but whatever the person is able to do requires some knowledge, even if that knowledge may be in the form of general strategies for analyzing situations and attempting solutions."

Harper, K. & Hite, Z, Freuler, R., Demel, J., Foster, T. (2007) Do the Problems Assigned Cultivate Real-World Problem-solving Skills? Presented at the American Association of Physics Teachers 2007 Summer Meeting.

Talked about their great class of really brilliant students that they teach really well. Then talked about a way to rate difficulty of problems with a 3 by 3 grid of # of possible answers, with the amount of info supplied. Then said they found that almost all the problems given in this great class were not good problems based on their rating scheme so they are writing new problems.

Hatano, G. & Inagaki, K. (1984) Two Courses of Expertise. *Research and Clinical Center for Child Development Annual Report, 6* (27-36)

This paper describes to process of cognitive development that leads to spontaneous expertise. "Starting with little or no documented declarative knowledge, or rules, children-through accumulated experience - acquire domain - specific knowledge enabling them to solve various problems in the target domain." It has many ideas in common with coordination class theory and it's interesting because they have a footnote about not using the word structure so as to not get confused with Piaget's structure of co-ordination. I need to read Piaget again.

"The processes of expertise are based upon the accumulation of experience, which consists largely of solving problems in a given domain. In achieving expertise, individuals, supervised by more capable members, solve increasingly complex problems in the domain, using relevant prior knowledge which is, in turn, gradually enriched and integrated."

The authors discuss how people gain a conceptual understanding of something and how it takes going beyond routine procedures. There are three criteria that they see as necessary to create the situation where a person would go beyond routine expertise and gain further conceptual understanding which leads to adaptive expertise. Adaptive experts "i.e. those who not only perform procedural skills efficiently but also understand the meaning of the skills and nature of their object" this seemed pretty simple but coupled with the definition of expertise from the beginning it means you must be able to create a new variation to be an adaptive expert. They also go on to say this means that an adaptive expert has a conceptual understanding of what they do so that they can modify the skill according to changes in constraints. Their definition alone is just a bit vague but all their examples follow this requirement of creating new solutions. "…adaptive, e.g., "invent" other procedural knowledge." For example, a farmer observes lots of natural covariations while growing rice. "Because of this knowledge, tentative though it is, an experienced farmer can deal effectively with various changes in constraints, like unusual weather or plant disease. Eventually, the farmer may even serve as a consultant for less experienced farmers, and as such, can be legitimately called an adaptive expert" (Hatano, 1982 "cognitive consequences…")

From procedural to Conceptual knowledge

Development of a *conception.* Without a conception, one can't go beyond the original version of a skill (procedure) except by trial-and-error or empirical minor adjustment. With a conception, meaning is given to each step and provides criteria for selection of possible alternatives for each step within the procedure. To develop a conception once needs

1) *Data* – must observe changes between actions and consequences.
2) *Model* – Avoid term structure so as not to get confused but this is what they mean – mental structure. "preconceptual knowledge-even if only a tentative and implicit one. Without this model, it is impossible to determine what variables are to be chosen for consideration from among an almost infinite number of candidates." Can be obtains through perceptions, a vague "image" or derived indirectly especially if invisible. Can be borrowed through analogy.

These are reciprocally selective: the observed data suggest what model should be adopted, and the adopted model constrains what kind of data are to be observed.

They assume that human beings have intrinsic motivation for understanding and that knowledge acquisition is endogenous – external feedback only serves as a cue for interpretation; internal feedback is brought about by reorganizing pieces of prior knowledge.

I'm not sure I agree with the 1^{st} part of this assumption. I know plenty of people who do not care to know the "why" as long as they can do the procedure as best they can. To them, knowing the why is not necessary.

Generalized consequences of Routine Expertise

Routine experts can become quite skillful - "people unhesitatingly call them experts, since their procedural skills are highly effective for solving everyday problems in a stable environment." However, it is only useful when the same materials and devices are available. This has been carefully documented (Anderson, 1981) so they want to focus on the consequences.

1) transfer of training, that is, abacus training made 3^{rd} graders' paper-and-pencil addition/subtraction of multi-digit numbers faster and more accurate primarily through the shared component skills of basic computation. Did not improve their understanding of carrying and borrowing, it reduced the number of "bugs," that is, the consistent application of wrong algorithms.

Comment [WA1]: Reason sims help understsanding.

2) Often produces as byproducts strategies or consolidated sequences of behaviors by which the skill can be more efficiently performed. So a routine expert can show a capacity very different from ordinary people on tasks that are apparently different. Such as readers of Japanese can quickly infer the meaning of unfamiliar *kanji* compound words by combining prototypical meaning of the component *kanji*. Also works on artificial words.

3) May produce new mental devices convenient for performing a given task. Abacus masters (and even lower intermediate operators) use a mental representation for storing digits so they keep their rehearsal buffer free.

So routine expertise can have "generalized consequences" but not thorough understanding, but through well established pattern of processing. Refernce Hatano 1981 and Scribner and Cole 1981.

What differentiates routine from adaptive expertise?

Not enough evidence to answer but humans must try to understand the world. To do this one must systematically investigate variation in action upon outcome. The factors they tentatively propose that encourage one to engage in experimentation:

1) System has built-in randomness, thereby motivating one to modify the skill to some extent.
2) No vital importance or usefulness so people tend to produce minor variations, often playfully.
3) The degree of understanding that is valued. An understanding oriented culture.

Use the term grand expert for the one that takes 1000s of hours of effortful practice.

Hunt, Earl (2006) Expertise, Talent and Social Encouragement in *The Cambridge Handbook of Expertise and Expert Performance* edited by K. Anders Ericsson, Neil Charness, Paul J. Feltovich and Robert R. Hoffman. (31-38)

Discusses Intelligence, Cognition and Experience through Gf, Gc and g. Gc is Crystallized intelligence (knowledge), Gf is fluid intelligence (reasoning skills, ability to detect patterns) and g is supposed to be general intelligence inferred from positive correlations between Gc and Gf. Once a task is learned Gf becomes less important but does depend on the type of task. Exemplified with two versions of an Air traffic controller's test. One basically required memorizing not-too-complicated set of rules and the other required the solver to develop orderly patterns of traffic in the area near a terminal. On task one Gf correlated with success at .45 to begin with and after training .30. For the second task Gf correlation went from .40 to .55.

Findings from Industrial-Organizational Psychology (Training time vs. intelligence). This section focused on Armed Forces Qualifying Test (AFQT). Personnel with high scores reached asymptotic performance in a year while those with low scores reached it in three years. After three years of service the correlations were much weaker. AFQT is considered a measure of Gc.

The next two sections focused on motivation. You must want to become an expert and social Encouragement and Expertise. Does society support you becoming an expert? Show pay distribution for different professions. Top 10% of Financial and Business advisors get 3.5 times the pay of the average in that field. Mathematicians get 1.7 times the average, lower than their control, which was High School Teachers since those raises are based solely on years of experience and not performance.

Kohl, P. and Finkelstein, N. (2007) Patterns of Multiple Representation Use in Expert and Novice Physics Problem Solvers Presented at the American Association of Physics Teachers 2007 Summer Meeting.

Korsunsky, Boris 2003 Cognitive Mechanisms of Solving Non-Trivial Physics Problems Dissertation Harvard Graduate School of Education

Breaks problem solving into *rigid knowledge* and *bisociation* . Definition of *bisocition* ability to relate one's rigid knowledge to a new, unfamiliar problem situation. In practice his definitions are the same as Mayer's, knowledge vs. processes. Defines his theory of BARK (Bisociation and Rigid Knowledge). Concludes that bisociation is more important for problem solving than rigid knowledge.

Good Lit review. Nice definitions. Breaks review into lit on good problems, cognition of Solving and teaching to solve.

Good problems sections discusses the lack of decent problems in text books He says what makes a good problem is that it is based on realistic situations, be ill-defined and contain "tricks" unknown to the students. Forces students to think "out of the box" and will gradually help students acquire "expert" problem-solving approaches.

Cognition of Solving Sections on misconceptions that he thinks should be tied closer to problem solving lit and math. Then goes into a section on experts and novices. Pointing out that expert and novice are typically not even defined.

Teaching to Solve. This section mentions that some research calls for cookbook strategies for problem solving. He believes this is the result of typical textbook problems not requiring expert skills. Makes fun of several different procedures. Some of which also criticize plug and chug but then they themselves say find the equation, substitute…

Recognizes the need to carefully design problems versus exercises and provides a clear but long example of a high jumper to show that not all problems are problems for the expert.

Larkin, J. H. (1979). Processing Information for Effective Problem Solving *Engineering Education* (285-288).

Larkin reviews two experts and a novice solving problems. The problems are back of the chapter type with a small twist to each problem. The two experts are physics faculty while the novice had completed one quarter of physics, the first student was a good student and received an A in the course. This student was chosen "…in order to avoid the obvious possibility that the experts would solve problems differently simply because they knew more relevant physical principles." The novice solved the problems slower than one expert but faster than the other and solved all of them successfully.

There were two differences between the experts and novices. The first was that the experts performed a qualitative analysis of the problems first while the novice immediately started using equations to solve the problem. Correct equations but nevertheless, he never spent time discussing the physical situation where the experts always did. Larkin says he has years of experience watching the same sort of difference between experts and novices so he feels this is fair to assume based on 2 experts and 1 novice.

The second difference was that the expert's knowledge appeared to be chunked while the novices was in pieces. The evidence for this was to graph the time between one principle to the next. The novice hesitated between each step. The experts, once they started with a particular principle, just breezed right through. The graph shows a few pairs of principles with a fair amount of time between (a nice flat line) and then 40 pairs at 10 seconds apart. The novices graph is very close to an exponential curve. Basically a really nice screen? plot for experts and a total mess for novices. This data was compiled from 5 problems for each solver.

Larkin tries a test of teaching these principles. She uses 10 students and carefully teaches them the principles for solving circuit questions (Kirchoff's Laws). Once she's convinced that they thoroughly understand these principles she divides the group in half. 5 students continue practice on the principles and 5 are instructed on general problem solving methods. They are taught the importance of qualitative analysis – discussing a physical picture of what is happening in the problem and are taught a couple of analogies such as water to current and height to potential and then he taught them how to group the principles on one chart to facilitate chunking. Then she interviewed the students while they solved new problems. With all the students, if they did something incorrect while solving the problems, she stopped them and told them to try again that was incorrect. The experiment group had three students solve all three problems and two who solved two of the problems. In the control group four of the five solved one problem and the other didn't solve any.

Larkin sites an example in the classroom where she taught a strategy (described in earlier work). The students who were taught the strategy drew more diagrams, made more intelligent use of algebra, more evidence of planning, strikingly greater use of relevant principles, less wandering into blind alleys and slightly better success in achieving correct solutions (no data given).

Larkin closes by suggesting that the one point of the article is that it seems promising to improve a problem solving skill if one

- Observes in detail what experts do in solving problems.
- Abstract from these observations the processes that seem most helpful.
- Teach these processes explicitly to students.

Larkin, J. H. (1980). Teaching Problem Solving in Physics in *Problem Solving and Education: Issues in Teaching and Research* edited by Tuma, D. T. and Reif, F. Lawrence Erlbaum Associates; New Jersey.

Larkin (1980) considers three examples of general strategies. First, means-ends analysis (defined below) has been demonstrated to be an effective strategy when solving an unfamiliar problem (Simin and Simon, 1978). Second is a type of planning where the original problem is replaced with an abstracted version in which certain central features are retained and then this solution is used to guide the solution to the original problem. This strategy has been modeled by Larkin et. al (Larkin, 1977; McDermott and Larkin, 1978) and shown to be effective means of solving many different kinds of problems; Third the use of goals and subgoals. Greeno (1976) has developed a successful model for solving geometry problems using subgoals. Larkin (1980) concludes "there may indeed be some general strategies (i.e., some major features) that are seen in skillful problem solving in a variety of disciplines. However, these strategies cannot be implemented without a considerable amount of domain-specific knowledge." Probably a reasonable conclusion; however, it's important to remember that all studies that she sites have no more than three subjects, sometimes only one, or they are simply computer programs that are able to solve basic math and physics problems. She also cites cases of people teaching these strategies but no research is mentioned.

Larkin continues with a discussion of functional knowledge units, factual and procedural knowledge.

Lemke, M., Sen, A., Pahlke, E., Partelow, L., Miller, D., Miller, D., Williams, T., Kastberg, D., Jocelyn, L. (2003). International Outcomes of Learning in Mathematics Literacy and Problem Solving: PISA 2003 Results From the U.S. Perspective. ***Education Statistics Quarterly, 6*****, 4.**

Leonard, W. J. , Dufresne, R. J. and Mestre, J. P. (1996). Using qualitative problem-solving strategies to highlight the role of conceptual knowledge in solving problems. ***American Journal of Physics, 64*****, 12 (1495-1503)**

This article reports on a calculus based course where the instructors focused on qualitative problem-solving strategies in an attempt to improve students conceptual understanding of problems. The instructors site common research that indicates that despite our fondest hopes, students do not gain a solid understanding of major ideas by reading the textbook, listening to lectures and solving problems. They believe part of the blame lies in the way we model problems.

"When modeling problem solving for students, although we are usually careful to state verbally the principle or concept being applied to solve a problem, we often only *write down* the equations by which the principle is instantiated. Students, therefore observe that it is the manipulation of equations that leads to solutions; their perception is that principles are abstractions that bear little relevance to obtaining answers to problems."

Additionally, students will continue to avoid the difficult task of attempting to understand the deep meaning of concepts in favor of more practical goals, such as becoming proficient at manipulating equations to obtain answers to problems.

Research in problem solving indicates that one of the differences between expert and novice problem solvers is the way they tackle problems and the level at which they identify the concepts that can be used to solve problems. Put more explicitly novice problem solvers begin by immediately manipulating equations in an attempt to isolate the unknown while experts look for the underlying principles and then determine how these principles can be implemented mathematically. When identifying principles, novices tend to use surface features while experts look deeper.

With these ideas in mind the authors developed what they termed a qualitative problem–solving “strategy.” The strategy includes three steps:
1) the major principle(s) or concept(s) that can be applied to solve the problem;
2) a justification for why the principle(s) or concept(s) can be applied; and
3) a procedure by which the principle(s) or concept(s) can be applied to arrive at a solution.
Or more simply put the “what, why and how” of a problem’s solution.

This strategy was implemented in a 150 student calculus based physics course. The course was traditionally instructed with the exception of how problem solving was modeled in lecture. The “strategy” was presented along side the “solution” and clearly labeled. Homework solutions included at least one example of this each week. Students were encouraged to discuss the strategy in lecture and to use it in homework; however, homework was neither collected nor graded.

Three different tasks were used to evaluate the effectiveness of this teaching method: 1) A strategy writing task, in which students wrote strategies on their own during an exam, 2) a problem-categorization task, in which students selected principles appropriate for solving problems and 3) a recall task, administered months after finishing the course, in which students wrote down what they perceived to be important ideas, or principles, used to solve physics problems.

The strategy writing task was only done with the course which had the strategy method of problem solving implemented into the course. Therefore this task does not serve as a means for comparing the new instruction to traditional methods. However, the authors did find this to be a useful window through which students’ understanding of what, why and how conceptual knowledge is applied to solve problems. What I found to be a potentially (see Question 1) important point is that the students’ failure to provide a justification for why a particular principle can be applied was a very common omission in students’ strategies. Other times students incorrectly explained the strategy; however, these omissions or incorrect explanations did not always impede the students’ ability to write a correct solution to the problem.

The next task was to categorize problems based on the principle, which would provide the most efficient solution. This task was given to a traditionally taught course that was taught by different instructors the following semester. Students in the strategy course consistently selected the correct principle more often than the traditionally instructed students. Traditional students selected the correct principle 48% of the time while strategy students selected the correct principle 70% of the time. The pattern of incorrect answers was consistent with both groups, i.e. the most common incorrect answer was the same for both groups on every problem.

Next the two groups of students, traditional and strategy were broken into quartiles based on their problem solving performance on the final exam on all problems other than the selection problems since they did not actually solve the selection problems. The authors point out that the 4^{th} Quartile (the weakest performers on the final) in the strategy class selected the appropriate principle as often as the 1^{st} Quartile students from the traditional class. They indicate that this is a significant achievement since other research indicates a positive correlation between ability to solve physics problems and their ability to select an appropriate principle. (Question 2)

The final task was a mini longitudinal study involving both the traditional class and strategy class. Students from the top quartile of each class were contacted and asked to come in some months after the course and identify the 7 most important principles or concepts covered in the physics I course they took. This happened 11 months after the strategy course and six months after the traditional. Only 13 strategy students participated while 22 traditional participated. The Traditional students identified Newton's three laws with only a few students identifying additional strategies while most of the strategy students named at least two other principles in addition to Newton's three laws. This seems interesting but is extremely limited because simply naming principles or writing the formula does not indicate an understanding of the concept.

Question 1) When students omitted the "why", I'd really like to see how often this happens because they really don't understand the why rather than they just omitted this step. Also I wonder when an incorrect explanation accompanies a correct solution if it is simply an indication of the students' becoming comfortable with the new vocabulary and new ideas or rather an indication of an incomplete understanding of the ideas and the less desirable conclusion that a complete understanding of concepts is not necessary for solving problems as many physics instructors would like to believe.

Quesiton 2) Table II

Table II. Principle selection score by quartile. (Quartiles were determined by overall performance on the final exam.)

	T-class ($N=376$)	S-class ($N=148$)
Quartile 1	64	83
Quartile 2	49	67
Quartile 3	44	69
Quartile 4	34	63

Table II does show a positive correlation with ability to select principles with problem solving performance for the traditional class; however, when looking at the Strategy class, that correlation is not clear. In fact, aside from the 1st quartile it appears that it is possible for students to be successfully taught principle selection techniques without improving their problem solving or maybe simply that unless traditionally instructed, principle selection cannot be used as an indication of problem solving ability.

Maloney, D. P. (1993) Research on Problem Solving in Physics. *Handbook on Teaching and Learning.*

Thorough review of physics problem solving. Fairly objective. In reference to Finegold and Mass "This study is important because it is one of the few that tried to hold the novice knowledge base constant and look for differences in general problem –solving skills."

Martin, L & Schwartz, D. (2008)

Martin wants to create two new factors, he calls them catalysts, for adaption. These are supposed to be in addition to the three Hatano and Inagaki describe (variation, no pressure, understanding valued).

1. Fault-driven adaptations. When the situation forces adaptation.
2. Prospective adaptations. When people have had a great deal of experience in a particular domain, they may engage the adaptive pattern even though they can still get by with the routine pattern.

Neither of these seem different to me. Both fit into Hatano and Inagaki's work.

Variation describes how a situation has enough variation (built in randomness) that makes original skill ineffective. One example is a farmer. A farmer in a greenhouse will not have enough experience with climate variations to handle growing outside. A farmer who's had years of experience outside can handle a change in climate that he's never experienced before because he understands his job well enough to make alterations to fit the never before encountered experience. This encompasses fault driven just fine. The problem situation is different in a way that forces a new skill.

Hatano's definition of these factors is "what factors encourage one to engage in such experimentation? Then he tentatively proposes three. Martin calls them catalysts. So these are the things that need to be in place to motivate or allow adaption to occur. So prospective makes no sense. It's not a catalyst or a factor that encourages experimentation.

Aside from the problem with the definition of a catalyst, adaptation by definition means a sort of reaction. Now put prospective on it and you're saying they react to something that hasn't happened yet? What Martin is describing as prospective adaption is different in all three examples. First deliberate practice specifically applied to chass grand masters. What he describes here is Hatano's variation. The grand master studies variation to get better. The next one is Schoenfeld's professor who spent the majority of his time working to understand the problem. The professor demonstrated a wider range of problem solving skills than what a novice would try. Since the professor is solving a new problem, this would be adaption but it's not a "catalyst". The catalyst or factor that encouraged him to engage was the situation, variation in the problem plus in a situation that encouraged this. It seems that Martin is calling any successful problem solving adaption and unsuccessful attempts routine when he says what Schoenfeld describes as experts and novices maps onto adaptive and routine. The students did not solve the problems or know how to so I would not classify this as either.

In addition using Schoenfeld's professor as an example contradicts Martin's definition of prospective adaption. Prospective adaption by Martin's own definition means that the person was doing something forward looking but not strictly beneficial. What Schoenfeld's professor did, metacognitive processing, would fall under strictly beneficial because he's the only one who successfully solved these problems. Saying that a novice who "read, make a decision quickly, and pursue that direction come hell or high water" is only doing what is strictly required is not true. They are doing what they know to do and only what they know and it turns out not to be enough. The professor does what he knows needs to be done when tackling a new problem and it barely works for him.

His third example is Kirsch and Maglio reporting about Tetris players and how they move the pieces aside so they could gather information about its orientation easier. This is different than both of the previous examples. This seems to be part of the Tetris players' routine. You might say since the layout at the bottom is different every time you play the game, it's adaption but it's part of what the expert player does not a catalyst or factor that encourages one to experiment.

I think Martin is confusing problem solving skills or understanding with the factors needed to be able to adapt. He's seeing adaption in action and then trying to characterize the adaption rather than what led to it. Hatano and Inagaki say, the greenhouse farmer is not equipped to deal with climate change because they don't have enough experience. To be an adaptive expert you must have an understanding of your routine skills and what each skills purpose is. When you have that, then you are equipped to adapt. Note that Hatano never says that you will adapt but that this part is necessary to adapt. Then when you do adapt, they hypothesize that it also takes one of the three factors (variation, no pressure or understanding-valued culture) to motivate it. I also want to say that adaption requires creativity in addition to understanding your job and having factors to motivate. Hatano does not directly say that but all of his examples do include that.

The study in this paper is great. Looking past the adaptive expertise framing the study is demonstrating that it takes experience before you're willing to use problem solving skills that take a little more time at first. They call it going to the tool pane and are specifically looking at the use of representation. Either lists or more sophisticated trees/matrixes. They used efficiency metrics (time, accuracy, search) to evaluate success. The tool of representations was used because it is measurable in comparison to other cognitive processes.

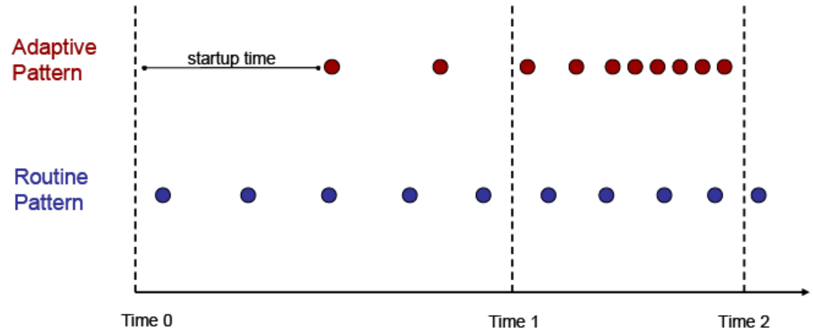


Figure 1. Two approaches to a recurrent task.

The study had three groups of students who were having to learn about different diseases and then diagnose new cases:

16 Undergrads	Continuous Cases - Teaching – Novel Cases.
16 Undergrads	Intermittent Cases – Teaching – Novel Cases
8 Graduate Students	Continuous Cases – Teaching – Novel Cases

Continuous access cases were where the students add access to the reference cases at all times before and during diagnosis. Intermittent access cases were when students could look at reference cases as long as they wanted except while diagnosing at which time they had to set them aside. There were 12 reference cases two for each of six diseases. There were 10 cases that needed to be diagnosed. The Novel Cases had four references cases and two new diseases. There were five new diagnosis that needed to be completed and patients could have any of the six original diseases or one of the two new ones. All students had full access during the novel phase. The Teaching phase required students to teach another person how to diagnose the diseases. They had five minutes to prepare and five minutes to explain.

The undergrads were only explained as being from a paid subject pool at a highly selective University (Stanford) and the graduate students were pursuing or had recently received their PhD in science or engineering at the highly selective Univ. There was mention later of them being at least 3rd year I think.

Everyone was able to successfully diagnose the cases so measurement was of time it took to diagnose, use of representation (none, list or tree) and choice of tests that had to be ordered before they were able to diagnose the patient. There were 7? tests that could be ordered. They called their measure of accuracy the WOR.

What they found was that undergrads rarely used reps unless they were in the intermittent case with the additional memory burden. All grad students used the reps, most trees. Grad students not only used reps, most moved from using a list to a decision tree. While teaching almost everyone used a rep. Then when they went back to solving cases, the students who did not use a rep, then did to teach, did not to diagnose the novel cases. This was the interesting part. Conclusion discusses how transfer studies that claim students do not transfer because they don’t recognize the transfer problem as being the same, well that is not the case here.. They just don’t see the time benefit. I did notice however, that the use of rep graph for the novel cases is ONLY if they make a new rep or modify an existing one. So it’s possible they used what they’d made for teaching or used what they’d made for the first part. The discussion suggests that this is not the case but it’s never explicitly addressed.

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

	Continuous Undergraduate	Intermittent Undergraduate	Continuous Graduate	Totals
Matrix/Tree	1	2	6	9
List	2	12	2	14
No Representation	13	2	0	15
Totals	16	16	8	40

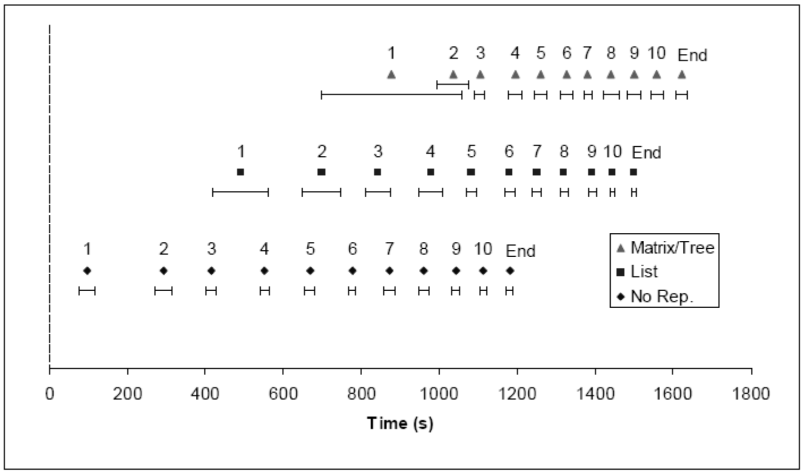


Figure 6. Speed of performance on the original cases. Error bars indicate +/- 1 standard error of time per problem.

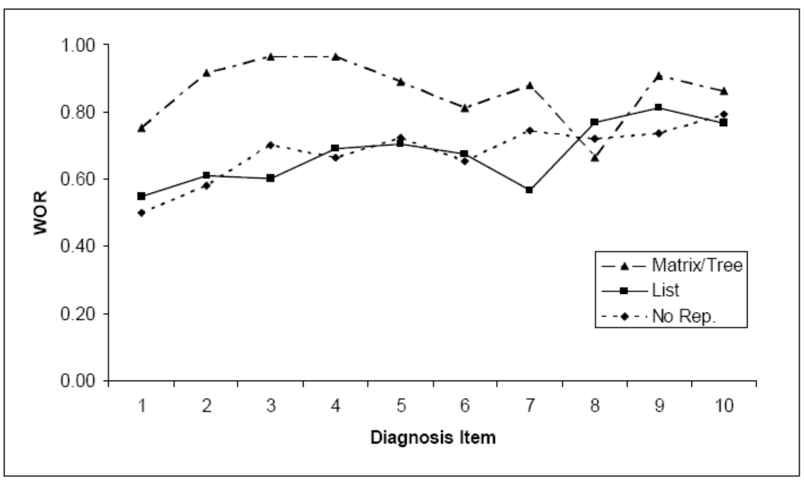


Figure 7. WOR across original cases, by type of representation.

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

	Continuous Undergraduate	Intermittent Undergraduate	Continuous Graduate	Totals
Matrix/Tree	2	6	5	13
List	8	6	0	14
No Representation	6	4	3	13
Totals	16	16	8	40

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

	Continuous Undergraduate	Intermittent Undergraduate	Continuous Graduate	Totals
Matrix/Tree	1	4	6	11
List	2	8	1	11
No Representation	13	4	1	18
Totals	16	16	8	40

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

	NOVEL CASES	
ORIGINAL CASES	No Representation	Representation
No Representation	15	0
Representation	3	22

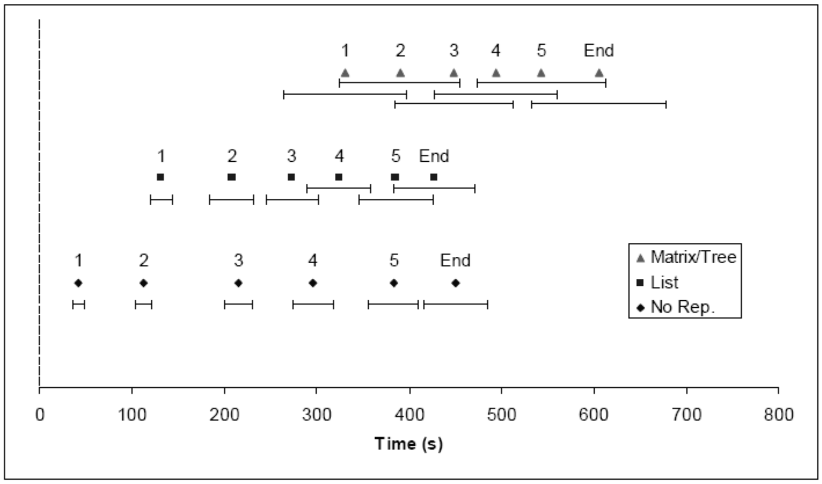


Figure 8. Speed of performance on the novel cases. Error bars indicate +/- 1 standard error of

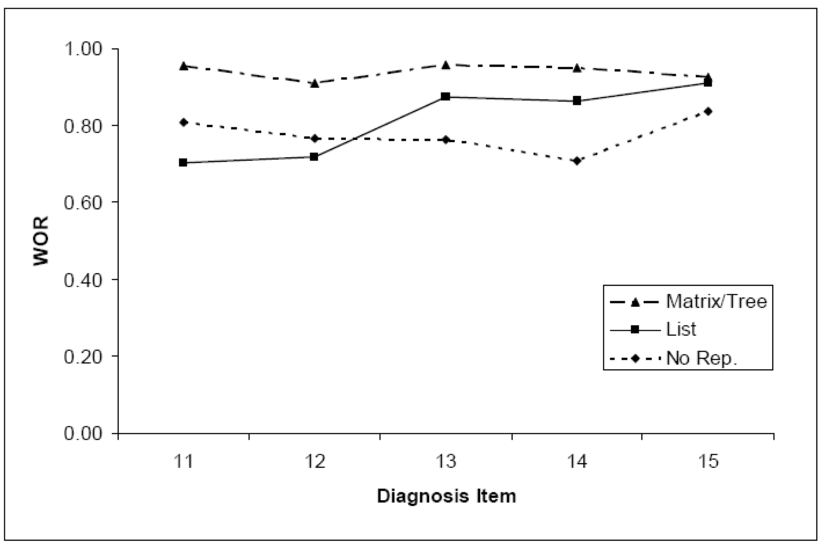


Figure 9. WOR across novel cases, by type of representation.

Mayer 2003 Thinking and Learning Chapter 9 section on Case based learning

Augues a person's experiences (knowledge) in a particular domain so that they can undergo analogical problem solving. Normally based on analogies to a person's own experiences but if augmented with case based learning, they can now tackle situations with knowledge from these cases they've learned about rather than just their own personal experiences.

Mayer 2003 Thinking and Learning Ch 12 Teaching by fostering problem solving strategies.

This chapter covers Mayer's four strategies for teaching problem solving. 1. Focus on a few well-defined skills. 2. Contextualize the skills within authentic tasks. 3. Personalize the skills through social interaction and language-based discussion of the process of problem solving. 4. Accelerate the skills so that students learn them along with lower-level skills.

He refers to Mayer 1999. However, when reading about project Odyssy he says it targets component skills and the skills are taught within the specific context that students are expected to use them, teachers and students model and discuss the process of problem solving and the higher order skills are taught before all lower level skills are mastered.

Mayer uses this program and others to support the idea that component skills are better taught within the domain. He does not say that problem solving is domain specific (as one may easily imply), just that it makes more sense to teach the skills within the domain that you intend to have the students apply the skills. You have to be able to apply a skill to use it and learn it. How can you learn a skill without making use of it? So of course it needs taught within an applicable context. So the way to test this is to find a skill that is useful in more than one domain. There are tests of transfer between domains; however, for various reasons they have not been successful with the exception of Schoenfeld.

Looks like all four of Mayer's principles for teaching thinking skills are in reaction to previous incorrect research conclusions. The first is to tackle the idea that intellectual ability is made up of component skills rather one single thing called intelligence. In this section he discusses Binet who believed it was component skills and that Binet has tests that correlate with school success. I'd like to see Binet's tests of problem solving skills. When discussion teaching component skills he keeps dodging any specific skills. He says that the list of component skills vary from one subject matter to the next. Only specifying planning and monitoring.

Learning and Instruction does reference Schoenfeld. It uses his work as an example that problem solving can be taught. He says although numbers are low it's evidence that Polya-like heuristics can be taught. His example of Schoenfeld's problem solving is almost identical to the example in his problem solving review from Wertheimer 1959 with the parallelograms except Schoenfeld uses a frustum and pyramids.

Next is problem solving skills are general and can be taught outside of other subject areas. This is dangerous because as I've discussed above when arguing against this, one can easily come to some other very different conclusions. I don't know yet if Mayer is saying this but to me it's important to learn things in context and that's the main problem with learning problem solving skills on their own without the subject matter to apply them to. However, he specifically says it is only possible to learn them within the subject matter because problem solving is domain specific. This implies that skills are different for different domains and that learning a problem solving skill while writing essays is not going to help you at all with algebra problems. However, if you learn meta-processing skills while writing essays, why can't you apply these to algebra? I just can't imagine that a person who's learned really strong meta-processing skills while writing essays wouldn't be ahead on algebra problems. Maybe they wouldn't immediately use these skills on their own but it shouldn't take much at all to help them see how it's useful in math as well. Meta-processing maybe isn't a fair skill to discuss so how about planning. Once someone has learned how to take the time and ask relevant planning questions when writing, I'm sure it'd take much less work and convincing to teach them to do the same with word problems as another student who's never received any planning instruction in any area.

In the section about intelligence being general or specific he cites tests of memory in different areas of learning. This is not a good way to test general skills because memory is content dependent. It depends on what you have to remember and how it relates to what you already know. Studies do show that these tests correlate and hence they have the g-factor. I'm sure there is some skill that is common among different memory tests, maybe a person's ability to pick things up the first time or the way they organize things but a stronger influence would be how the new information connects to what they know. So this is not a good way to measure problem-solving skills. What are some other tests of learning? I don't know exactly what these would be but I cannot picture any test of learning that does not depend on what you already know. Mayer goes further to say that cognitive skills such as planning a solution in mathematics and planning an essay may not have much in common because they are domain specific.

3. Process and not just right answer

His third point is to teach process and not just spitting out the right answer.

4. Higher level skills vs Lower level

His fourth point is that high-level skills can be taught before or along with low-level skills. After reading his example here it seems much more reasonable than in the review. Here it's really just pointing out what I've found also. These so called lower-level skills are not required before higher. In this book so far at least, he's not advocating that it's better to teach higher level skills at the same time (except for the purpose of maintaining students' interest in school). He's mainly emphasizing that it is not necessary to teach low level first. So are they really low level? No in my opinion it's the difference between knowledge and cognitive process. In fact, in most schooling it seems emphasis is on knowledge acquisition and not on cognitive processing. The same problem that Benezet was seeing, but not getting, when he tried to abolish mathematics in 1935.

He points out that teachers focus on teaching knowledge with the assumption that solving problems or writing essays requires these skills so students are learning them automatically while doing the assignments that were designed to enhance their knowledge. But students do not always know what they need to be doing. They are looking for the most efficient route to the answer and may not even realize that it could help to ask some specific planning questions. Although CLASS results show that they do know some of this is important but so much of class work is not actually problem solving so it's actually not necessarily useful for a lot of classes because teachers are focusing on the acquisition of more knowledge and not on thinking processes.

The first is Productive Thinking Program where students solve mysteries. Students are asked to formulate their own questions, then generate ideas to explain the mystery. After this, Jim and Lila give theirs. Jim and Lila make mistakes at first but with the help of the adults in the story, eventually figure out the mystery. There is a list of skills given for these problems. Things students are supposed to do while solving them. Many of these skills are things that are in my rubric. Studies of this program show that students in the program do better on similar types of problems even 6 months later; however very little improvement is seen in other domains. So Mayer concludes that it only helps with a few specific skills that aren't useful in many other areas. I wonder what the other tests were like? I'm sure these wouldn't help so much in solving physics or math problems with lots of content if the person doesn't know the content. Again content is the dominant skill – or rather an insurmountable hurdle. But these skills would help if given a physics problem that has lots of steps and requires more than content knowledge. However these types of problems are not always valued in math and science so they are unlikely to be what was used when testing the transferability of this program.

Mayer, Richard E. (2003) What Causes Individual Differences in Cognitive Performance? *The Psychology of Abilities, Compentencies, and Expertise* Edited by Robert J. Sternberg, Elena L. Grigorenko. Cambridge University Press. (263-274)

Mayer summarizes the various points of view. Says Cognitive ability is the result of knowledge which is built from a genetic predisposition (ability) towards studying a certain topic.

He discusses the definition of cognitive performance and clarifies the difference between what an expert can do on a consistent basis versus creative performance which is somewhat erratic. Makes me think it's the difference between solving exercises and solving problems. Not enough info here on creative performance to be sure.

There are arguments in here that genetic predisposition determines what sort of activities people are attracted to. Then they do the activity and when they become better at it. It makes it easier to read so they read more then become even better readers. Or they do another sort of activity, become better and seek the company of others who are even better at it, which motivates the person to work at improving their level. These are called multipliers.

Mayer, Richard and Whitrock, Merlin (2006). *Ch. 13 Problem Solving.* Handbook of Educational Psychology, Second Edition. Edited by Patricia A. Alexander and Philip H. Winne. Erlbaum; New Jersey (287-303)

Short but concise description of problem solving precedes the review. He points out that most real problems are ill-defined but that this part of the definition is not personal (does not depend on the solver). He also uses routine vs. non-routine to discern a problem from an exercise. Instead of using Skill, MetaSkill and Will he discusses cognitive processes and knowledge. Processes: Representing, Planning/monitoring, Executing and Self-regulating. Knowledge: Facts, Concepts, Strategies, Procedures and Beliefs/metacognitive knowledge. Interestingly he says "I am not good at math" is a belief and falls under Metacognitive knowledge. Also mention of conditional knowledge – knowing when and why to use existing conceptual and procedural knowledge.

When discussing Redish or some others work, a good point from Mayer is "Although executing is sometimes emphasized in classroom instruction, the major difficulties of most problem solvers involve representing, planning/monitoring, and self-regulating" (Mayer, 2003).

Now discusses types of promoting problem-solving transfer. Example of Gestalt teaching students to find the area of a parallelogram by either showing the calculation or by taking the triangle off on end and putting on the other so you have a rectangle and then finding the area of the rectangle. Says this is meaningful learning or productive thinking because students do the same on tests of finding area of parallelogram but second group does much better on test with finding area of other figures. I think this sounds like he's giving them knowledge that can attach to what they know already rather than an abstract idea that just has to be memorized. So making sense of why it works a certain way – then students can apply it elsewhere.

Next section is about Mayer's SOI model of selecting, organizing and integrating information. It explains exactly why he believes triangle method of parallelogram solving works. Pretty much as described above.

Automaticity methods are useful such as memorizing times tables so that working memory is freed up for problem solving tasks such as devising and monitoring a solution plan. Uses the term higher-level problem solving tasks. I have said that there isn't a hierarchy but I want to be careful of what is meant by that. Control methods vs. automated knowledge – hard to not say control methods (meta skills) are not higher level. It's just that higher level does not mean you have to have one before you get the other. It just means it's a different thing that is more sophisticated maybe?

Teaching students to automate decoding of passages while reading by having them read out-loud over and over – the students were better and comprehending passages after automating these procedures. Similar work in math showed automating component skills such as recognizing congruent angles allows students to progress from effortful performance to automatic performance. This frees up capacity to focus on the problem solving. Work by Anderson & Schunn, 2000 p 26 says students can achieve the same level of competence in one third of the time as traditional education. I wonder what they did since my math training was all systematic practice of lower level skills. What are they doing differently and what are they comparing it to?

Other studies show that calculator use helps students in realistic mathematical problem solving; presumably because they can devote their attention to high-level processes rather than low-level arithmetic.

Schema Activation methods: Advance organizers, pre-training, and cueing. All very similar used to cue ideas that students already have so that new information can be attached to current knowledge. Usually concrete rather than abstract ideas are cued.

Generative methods: Elaborative, Note-Taking, Self Explanation and questioning. Used to encourage integration. Students who have to create analogies or summaries of material that they have either read or heard about do better.

Guided discovery: In the 60's push there was a for discovery but research has shown that guided discovery is necessary for learning. Discovery allows for integration but does not prime selecting process. Behavioral activity (hands-on, discussion and free exploration) is not what is needed it's cognitive activity (selecting, organizing and integrating knowledge).

Modeling Methods: Learning by example better than learning by doing. More effective if the worked examples clearly specify the subgoals. Mayer has done work on case-based learning. Apprenticeship seems similar to above. Can be productive but only if all three cognitive processes are encouraged.

Next section talks about teaching thinking skills directly. Mayer says it's more successful when

1. Focus on a few well-defined skills.
2. Contextualize the skills within authentic tasks.
3. Personalize the skills through social interaction and language-based discussion of the process of problem solving.
4. Accelerate the skills so that students learn them along with lower-level skills.

There isn't reasoning given behind any of the about requirements. He gives a brief overview of four examples of different attempts at this. Some were successful and some weren't. Why is it better to teach these skills before the basic knowledge of the domain is automated? One of his examples was in economics and in that case the basic skills already were automated and that program was quite successful. In the other programs the skills had not yet been automated. Yet in these other three, there were weaker results – at least it appears that way from this review. This was the first section without reasoning for each of the conclusions. My first inclination would be that trying to master the content at the same time as the problem solving skills would be distracting – add to cognitive load. Unless he's thinking that if they know the base skills they will be less likely to engage in cognitive processing because they think they know? Says there is little research so these guiding principles seem to make sense.

Says consistent evidence that thinking skills courses promote transfer mainly to similar problems within the same domain. No reference for this and no previous work cited where they tried to test transfer to other domains and failed. I'm assuming this statement is based on something like that but what was tried?

Says one of Educational Psychologists greatest successes is teaching domain specific thinking skills. Essay writing is greatly improved when students are required to outline the essay before starting or required to answer questions such as why am I writing this, who is it for,

As Mayer stated in the beginning of the teaching thinking section. So many teachers assume students are doing or learning these thinking skills automatically to be able to solve problems. Turns out they need shown how to think about things and what is important to focus on when attempting to solve problems. If shown or required to do these necessary steps, it helps. Students are always looking for the most efficient route to the end. Many times without realizing what they've sacrificed by taking the shortest route. Or in some cases, maybe not even knowing other routes exist.

At the very end there is one paragraph on metacognitive skills. It says, "Becoming an effective problem solver requires the development of self-awareness of one's thinking processes."

This review is missing metacognitive processing skills. He has courses that teach thinking skills but none of them talk about metacognitive skills. The description is general enough that a person could easily slip Schoenfeld's work in that section or even in the apprenticeship section but the description of apprenticeship does not fairly represent what Schoenfeld does. He does briefly mention metacognitive processing in one sentence and it's in the table but that's it. No where in the teaching section is it covered.

Mestre, J. and Feil, A., (2007) Bait-and-Switch: Problem Solver Reaction to (Secretive) Problem Switch. Presented at the American Association of Physics Teachers 2007 Summer Meeting.

Interviewed intro and grad students on intro type physics problems. Problems were on a computer screen. Student gave first look and when ready to explain they had to look away and while explaining what they would do, the problem was switched. Jose presented data that implied that no one caught on during the first round but then experts were more likely to catch on second round if the switch made a difference in the physics but not in the surface features. Novices did not notice either way. The theory was they'd notice surface changes such as color changes but they did not. A video showed data that no one caught on during the first round, but Jose showed a video clip with a graduate student doing a double take and saying that the problem had been changed and that he felt that it had been changed on him the first time also. Maybe the data for the first round only indicates reactions during the first round.

Pretz, J. E., Naples, A. J. and Sternberg, R. J. (2003). Recognizing, Defining, and Representing Problems in *The Psychology of Problem Solving* edited by Davidson, J. E. & Sternberg, R. J. Cambridge University Press; New York (3-30).

Presents a problem cycle. When the steps are completed it usually gives rise to a new problem and then they must be repeated.

1. Recognize or identify the problem.

2. Define and represent the problem mentally.
3. Develop a solution strategy.
4. Organize his or her knowledge about the problem.
5. Allocate mental and physical resources for solving the problem.
6. Monitor his or her progress toward the goal.
7. Evaluate the solution for accuracy.

Gives a review of classes of problems and then spends the rest of the chapter focusing on recognition, definition and representation. The specific variables that affect people's ability to solve problems are analyzed with respect to these three particular aspects of problem solving. The variables that are analyzed knowledge, cognitive processes and strategies, individual differences in ability and dispositions, as well as external factors such as social context.

His classes of problems include nice definitions of well-structured and ill-structured problems. He actually does not recognize the distinction between metacognitive processing and cognitive processing in general. He lists 1.) Recognizing the existence of a problem, 2.) Defining the nature of the problem, 3.) Allocating mental and physical resources to solving the problem, 4) Deciding how to represent information about the problem, 5) Generating the set of steps needed to solve the problem, 6) Combining these steps into a workable strategy for problem solution, 7) Monitoring the problem-solving process while it is ongoing, and 8) Evaluating the solution correspond to the first second and fourth metacomponents, which are used in the planning phase of problem solving.

One of the three aspects of problem solving (actually 1 & 2 of 1st list and 1,2 & 4 of 2nd list) may be very difficult to do but the other parts easy. What makes a problem hard depends on the particular problem. Sometimes the biggest challenge is simply identifying the problem. You can either be given the problem, recognize a new problem or create a problem. Typically students are given a problem. Recognizing a problem would be something like finding a hole in the current research such as all studies have been of men. In this case the hard part was recognizing the problem, once that's done the solution is straightforward, repeat what's been done with women. Creating a problem is much more complicated. At times once you've created it, you still have the challenging of making the existence of the problem clear to others. There was no example given to help clarify the meaning of this.

The author discusses the shortcomings of education and how children are only given well-defined problems. Not only is the solution path clear (if you have the knowledge) but the definition of the problem is provided. Students rarely if ever have to define their own problems.

Price, E. and Gire, E. (2007) Physics Majors' Modes of Thinking During Problem Solving Presented at the American Association of Physics Teachers 2007 Summer Meeting.

How do students pick which concept to apply to physics problems? Example problems were ball drop and tarzan. In their studies (don't know numbers) students would prefer kinematics but would switch to conservation of energy at times because it's easier to remember but not because they thought it would be best. When asked they choose kinematics because it's what they learned first so believed it must be more important.

Redish, J. (2003). Teaching Physics with the Physics Suite. Wiley & Sons Hoboken, NJ

Redish classifies the "hidden curriculum" as including expectations, metacognition and affect. Within expectations is a students feelings about science and how they interpret what they hear. Metacognition is thinking about their own thinking (*reflection)* consciously – "these two sentences don't make sense" and subconsciously via confidence "it just feels right". Affect includes emotional responses including motivation, self-image and emotion.

Hammer variables from Hammer, D. (1996). More than misconceptions: multiple perspectives on student knowledge and reasoning, and an appropriate role for education research, *American Journal of Physics 64* 1316-1325
Independence takes responsibility for constructing understanding vs. takes what is given by authorities without evaluation. Coherence believes physics needs to be considered as a connected consistent framework vs. believes physics can be treated as unrelated facts or independent pieces. Concepts stresses understanding of the underlying ideas and concepts vs. focuses on memorizing and using formulas without interpretation or sense making.

Much of his definition of metacognition *reflection* is exactly what I think of as meta-processing. "Thinking that reflects on the thinking process itself." He gives some examples "including evaluating their ideas, checking them against experience, thinking about consistency, deciding what other ideas might be possible, and so on"

Motivation under affect includes *internally motivated* – self-motivated by an interest in physics and a desire for learning *externally motivated* – motivated to get a good grade, *weakly motivated* – taking it because it's a requirement but only are concerned with passing *and negatively motivated* – want to fail. Maybe to demonstrate to a controlling parent or mentor that they are not suited to the subject.

Discusses self-image and gender threat. References Steele, C. (1997) A threat in the air: How stereotypes shape the intellectual identities of women and African Americans *American Psychology 52* 613-629 Gave a group of math majors a test that was slightly above their ability. He told one group of students that this test is just a trial and wanted to see how they would do and told the other group that the test showed gender differences. No threat both groups got 16-18 correct, gender threat women got 5 correct and men 27 correct.

Carpenter 1983 Carpenter, T.P., Lindquist, M.M., Mathews W. and Sliver, E.A. (1983) Results of the third NAEP mathematics assessment: secondary school *Mathematics teacher*, *76* 652-659. 45,000 13 year olds "An army bus holds 36 soldiers. If 1128 soldiers are being bused to their training site, how many buses are needed?" 70% did division correctly but 23% gave the correct answer of 32 while 29% gave 31 remainder 12 and 31 given by another 18%. Expectation: Mathematical manipulation is hat's important and what is being tested.

Redish, Joe (2007) Physics Dept Colloquium

Designed for the non-PER expert. Showed Harvard results of not knowing why we have seasons. Talked about neural connections. We have lots of knowledge grouped together so three charge problem is easy for us. Students take an hour. Showed the gorilla video. Kept saying students miss the gorilla when we lecture to them and don't give them time to put everything together. Presented the problem solving of the 3 charge problem differently than in the paper. Claimed students had correct idea 4 times before TA suggested making a diagram and then students put ideas together and didn't lose them again. In the paper it presents transcripts and it didn't quite work out that way. Presented epistemic games in a very unclear way. A lot like the first paper. Just throwing things out there in the middle of a lot of exciting things. Kept talking about how useful group problem solving is for helping students build the complete knowledge that they need about physics. Joe also says he'd teach physics majors differently and that they are ready for lectures because he knows they'll take notes and then go home and make sense out of what he's told them. But he'd still do group problem solving.

Did point out that the class he learned the most in was a modern physics course with the most horrible instructor ever. I had the same experience. Always thought it was because I read the book carefully since I got nothing from his notes or lectures. In other classes I depended almost completely on lecture notes. Books have descriptions of why and how while lecture notes are mostly definitions and equations. Maybe should give physic majors a bunch of disconnected information that makes no sense during lecture and tell them to go home and make sense out of it? Essentially give them puzzles to go home and solve that force them to read (not just read once) carefully the text to figure out how it works. I suppose essay assignments would also function in the same way.

Both Joe's work and Tominoro's paper speak of a group of students as one person. When Joe talks about what the students did and says 4 times they came to the fact that the charge is negative. Part of the transcript that I saw shows that same girl coming back to it and other girls then pulling them away. Joe presented all this stuff as if it were one entity putting all these ideas together. Just like Tominorio's Games. He picks a game that they play but the game is defined by their progress and there are times when if you look at one person at a time different games are happening at the same time. Then if you combine their actions, yet a third comes about? Or they claim they switch games but it's really just who is talking. . Yet they describe the game as switching or change the definition of the game based on complete discourse which involves two people talking and not paying attention to each other. Does not seem like a useful way to understand thinking. Maybe group dynamics but that's not what they are claiming. If you do it their way, you have to talk about the level of understanding and confidence or ability to adapt and how all this plays out and the problem gets solved. They say it takes an hour for groups to solve these.

I'm sure there are always groups with students like I was and I always see others where they figure out the math because they know the rules – not because they actually think of all the physical implications that a physicist does. Then this person after spending less than 2 minutes to figure it out, spends another 5-10 showing other group members and they move on. How is his group problem solving and games useful then? We can do the same thing in lecture situations and show a few students who really excel in that situation. It's not because that person learns best with lecture. It's because that person has their own skill set (including motivation) that allows them to learn if presented with the materials and goal. I believe Redish does have overall data showing the success of these courses (at least he has me thinking this) but he never presents anything that shows how his work is successful over an entire course?

Really focuses on how to build physics knowledge. He may work on other skills but I don't see him encouraging metacognitive processing skills, planning, maybe analysis a little. Pretty much just things that fall in the skill area.

Ross, B. (2007) Problem Solving and Learning for Physics Education. 2007 Physics Education Research Conference Proceedings.

Talked about the basic understanding of memory and reasoning from a cognitive science viewpoint. He also pointed out some of the additional richness of physics problems and how that could add to current understanding.

He is very clear about there being a number of specialized systems that have different purposes and at least somewhat different means of representing and retrieving knowledge. A critical distinction is between procedural and declarative memory but that there are additional distinctions as well. Declarative is facts "knowing what". Procedural is "knowing how". It's not the factual information of how the procedure is done it is the procedure. Tying your shoe is procedural. People can't explain how to tie their shoe without doing it or imagining it. This is procedural. Lots of studies support this: behavioral studies, neuroscience showing a difference in brain regions and amnesias and computational models that are able to account for many different findings using both types of representations.

Reasoning theorists argue whether reasoning has different systems. Ross believes the evidence tends to favor at least two if not more. System 1 is a heuristic processing system that works unconsciously in an associative manner, pulling to mind what relates to the current info being processed. System 2 is more analytical, deliberative processing, such as one might see in a novice figuring out how to fill numbers of a problem into equations. System 1 is usually thought of as a faster initial processing that provides an answer or passes information to System 2.

This provides an example of why this information is important for physicists. A common problem when teaching is "…a failing student would express dismay since s/he had understood the way to solve the problem perfectly when I explained it in class. I had to point out that the exam did not test their understanding of my solution, but their ability to generate their own solution. This mismatch in processing is a common and underappreciated influence on transfer."

Content in thinking. Abstract thinking is hard. Describes how everything we learn is couched in a concrete example. Principles are taught and then examples given. We hope they will clarify the principle and extract only the main ideas from the example. Instead the concrete information is part of this knowledge. To begin with examples are remembered and eventually these become categories or principles in memory. So a novice uses analogies to solve problems and an expert finds categories and uses abstract principles. These are not so different. Just different ends of a spectrum of knowledge about a principle. There's always some structure left in the abstract principles of experts. Because it's useful and Ross suggests that it should still be there. Inclined planes usually fit with certain categories of physics principles. An example of a math teacher doing an escalator problem. A person is moving up a down escalator. While solving the problem she talks about moving upstream because all of the problems she practiced with, to learn the principles necessary to solve this problem were associated with rivers and boats. Does point out that as educators we cause some difficulties by always using the same types of structures to teach certain principles.

Memory retrieval Cues – Memory – Retrieved Info (go back to cues maybe) – Problem Solving. Uses this to describe the differences between experts and novices solving problems. Novices use structure to cue an example or formulae then go from there. Experts use structure and the category that this structure cues to cue a much larger range of earlier problems, many formula and a variety of problem categories. The category has become a very available memory from much earlier use compared to examples. Given this retrieval one is likely to see category based problem solution perhaps falling back to analogies or equations for difficult or unusual problems. Specifies that cues are not the problem elements, but the problem as interpreted by the problem solver. "This representation is critical and has huge influences on problem solving at least partly due to what knowledge is retrieved."

Ross, B. (2007) Problem Solving and Learning for Physics Education. Presented at the 2007 Physics Education Research Conference.

"One of the only places in the world where context doesn't matter is mathematics."

We always overestimate. We give students principles and examples and want them to look at the intersection but they don't. Roses and tulips are harder to work with than roses and vases. Thoughts: Harder to transfer when procedure is learned in context. diSessa type B transfer gives a person time to parse and process how a previously learned idea applies. On the other hand, its harder to learn without context to attach it to. Students are not taught how to categorize but categories are laid out. So students try to pick up on this.

Savrda, Sherry L, (2007) Stabilization: An Alternative Model of Problem Solving Presented at the American Association of Physics Teachers 2007 Summer Meeting.

Refers to other research that claims four primary factors of problem solving: categorization, goal interpretation, resource relevance and complexity. It suggests a fifth superordinate factor called stabilization. Stabilization happens once a solver reaches a point where they are settled in and just need to do calculations. This may happen several times during the problem if something occurs to throw them off and they have to reorient.

Schoenfeld

Uses the same sorts of graphs that Valerie and Noah use to show problem solving behavior. Does a nice job of recognizing real problem solving by an expert rather than getting mixed up with exercises. Uses metacognitive process questions throughout his teaching. Just like the ones in Berardi-Coletta.

1. What (exactly) are you doing? (Can you describe it precisely?)
2. Why are you doing it? (How does it fit into the solution?)
3. How does it help you? (What will you do with the outcome when you obtain it?)

Schoenfeld, Alan H. (1985) *Chapter 1 A Framework for the Analysis of Mathematical Behavior* Mathematical Problem Solving. Academic Press, Inc. Harcourt Brace Javanovich; Orlando…(Ch 1 11-45)

Introduces four categories of knowledge and Behavior:

1. Resources: Knowledge – factual, procedural and propositional. Key phrase is "capable of bringing to bear"
2. Heuristics: Rules of thumb or strategies such as exploiting analogies or working backward. My view: more knowledge. Pretty much domain specific but a little of it can be used for physics problems but only those that require math so maybe still domain specific.
3. Control: Major decisions regarding planning, monitoring and assessing solutions. Good control allows them to exploit their resources, without it they squander their resources. He says control is often called metacognitive skills in psychology lit but he includes all planning and metaprocessing (ability to determine reasonableness etc…). Similar to the metacognitive skills that Mayer used in '92 work. Anything that determines what they do next or how they use what they know (except belief part I guess).
4. Belief systems: Mathematical worldview. Beliefs can determine how one chooses to approach a problem, which techniques will be used or avoided, how long and how hard one will work on it and so on. Beliefs establish the context within which resources, heuristics and control operate.

Schoenfeld, Alan H. (1985) *Chapter 2 Resources* Mathematical Problem Solving. Academic Press, Inc. Harcourt Brace Javanovich; Orlando…(45 - 68)

Knowledge

Schoenfeld, Alan H. (1985) *Chapter 3 Hueristics* Mathematical Problem Solving. Academic Press Inc. Harcourt Brace Javanovich; Orlando…(69 - 96)

Demonstrates various heuristics and finds three reasons that prior attempts at teaching these strategies have been unsuccessful. 1.) That the general title working backward, or "exploiting analogy" are very broad summary statements that aren't useful. Have to teach the explicit strategies underneath Quotes Polya 1dknowing to use the reigth?? strategy, 2 knowing the appropriate versions of it for that problem, 3 generating appropriate easier, related problems, 4 assessing the likelihood of being able to solve and then exploit each of the easier problems 5 choosing the right one, 6 solving the chosen problem and 7 exploiting its solution. Learning to use the strategy means learning all of these skills. 2.) One cannot expect too much of heuristic strategies. One's success in any domain is based on a foundation of one's resources in that domain, and even a good mastery of heuristics cannot be expected to replace shaky mastery of subject mater. 3.) Depends not only on access to the strategies but good "executive" decision-making.

Means end analysis actually recommended as a useful strategy by Simon 1980. In teaching problem solving, major emphasis needs to be directed towards extracting, making explicit, and practicing problem-solving heuristics 0 both general heuristics, like means-end analysis, and more specific heuristics, like applying the energy conservation principle to physics. P94

Schoenfeld, Alan H. (1985) *Chapter 4 Control* Mathematical Problem Solving. Academic Press, Inc. Harcourt Brace Javanovich; Orlando…(97 - 144)

Control

"A Vygotskean perspective suggests that the "internal dialogues" of competent problem solving result from their having internalized aspects of the cooperative problem solving sessions in which they had been engaged. Conversely, a paucity of cooperative problem-solving experience might (in depriving the student access to overt models of control behavior) hamper the development of individual control strategies."

Reference to Brown, Bransford, Ferrara and Camione, 1983 discusses these things. Schoenfeld says "The research indicates that the presence of such behavior has a positive impact on intellectual performance. That its absence can have a strong negative effect – when access to the right knowledge is not automatic…"

Schoenfeld says that Brown et al. also say regulation of cognition includes *planning* (predicting outcomes, scheduling strategies, various forms of vicarious trial and error etc.), *monitoring* (testing, revising, rescheduling one's strategies for learning) and *checking* outcomes (evaluating the outcome of any strategic actions against criteria of efficiency and effectiveness).

If this is regulation of cognition then one could take their definition of monitoring to mean stepping outside of the solution every once in awhile to see if it's a good idea or if something else can be done. Mayer uses the same words - planning, monitoring - and adds executing but then says self regulating is separate. Says monitoring is evaluating the appropriateness of the solution method. This can be interpreted as the same if it didn't have the preceding sentence about regulation of cognition.

Schoenfeld, Alan H. (1985) *Chapter 5 Beliefs* Mathematical Problem Solving. Academic Press, Inc. Harcourt Brace Javanovich; Orlando…(Ch 5)

Belief System: One's "mathematical world view", the set of (not necessarily conscious) determinants of an individual's behavior.

About self
About the environment
About the topic
About mathematics

"Purely cognitive behavior – the kind of intellectual performance characterized by discussion of resources, heuristics, and control alone – is rare. The performance of most intellectual tasks takes place within the context established by one's perspective regarding the nature of those tasks. Belief systems shape cognition, even when one is not consciously aware of holding those beliefs."

Schoenfeld, Alan H. (1985) ***Chapter 6: Explicit Heuristic Training as a Variable in Problem –Solving Performance*** **Mathematical Problem Solving. Academic Press Inc. Harcourt Brace Javanovich; Orlando…(189-215)**

He presents a small study at Berkley with 7 students. The control group was given 20 problems that contained 5 possible strategies. The problems were mixed up so that the students did not get the same strategy twice in a row. They were given two weeks to work on the problems. First session included a 10 minute lecture. The control group was told that this was to help them with their problem solving skills and told them the mechanics of what they'd be expected to do. Next and during the rest of the sessions they would try to solve them, then look at the solution and then listen to a tape recorded explanation. The experimental group had the same problems but in order of strategy so they could focus on one strategy per session. Their tape recording at the beginning of the first session told them the purpose was to learn the five strategies and they were given a sheet summarizing the five strategies that they kept at hand during all sessions. In addition to the solution for each problem there was information on the strategy used to solve it boxed in the margins. The tape recordings for each solution also started with "now notice blah←???" so that it showed them how to pick the appropriate strategy.
During the posttest the students were stopped at five-minute intervals and told "stop, take a deep breath, and look over your work. Then decide whether you want to continue in that direction." Or " stop, take a deep breath, and look over the list of strategies. Then decide whether you want to continue in that direction."

All students had fairly extensive backgrounds in math so had previously received instruction on all types of problems presented.

The experimental group did learn. One third of the five problems they appeared to gain understanding of the strategies. The question is why? I see a couple of reasons. 1. The strategies were explicitly stated and the students were trained, one at a time, how to use the strategies and then given a slightly different problem at the end to apply the strategy too. Hopefully helping them stretch themselves a little further than just rote learning. 2. During the posttest, students were interrupted every 5 minutes reminding them to look at the strategies to see if they were on the right track. Not only does this stop them for a bit, it tells them what to look at. Trying to force metaprocessing. From the data, it really appeared to change their direction but only for the experimental group. The author mentions it making all the difference for two of the four but does not mention whether it did or did not affect the other two students in this group. In the control group, the students did pause sometimes and would change what they were doing but never changed their heuristic strategy. So a big part can be accounted for due to the warnings. However, there still is improvement beyond what was caused by warnings so training also helped.

When comparing to my work it seems the skills that were being assisted were metaprocessing, picking out useful information, remembers previously noted facts (different scenarios for this due to different training structure), planning how (since same strategy used each time) connects steps and pieces, keeps problem framework in mind (different levels of necessity in each group) tie in info from others (easier for experimental group since info was pre organized for them).

Schoenfeld, Alan H. (1985) *Chapter 7: Measures of Problem-Solving Performance and Problme-Solving Instruction* Mathematical Problem Solving. Academic Press Inc. Harcourt Brace Javanovich; Orlando…(216 - 239)

"The testing literature has offered few methods, whether for purposes of research or for use by teachers, of directly examining the procedures used by individuals as they attempt to solve problems. Virtually all available examinations of problem-solving performance have used *product* measures rather than *process* measures. That is the tests they employ focus on the correctness of the answers the students produce to problems, rather than focusing on the procedures that the students use in trying to solve them."

Includes references of reviews of commercially available and researched problem solving examinations. All are wanting and all focus on products. Not validated for processes.

Schoenfeld, Alan H. (1985) *Chapter 8 Problem Perception, Knowledge Structure, and Problem-Solving Performance* Mathematical Problem Solving. Academic Press, Inc. Harcourt Brace Javanovich; Orlando…(242 - 269)

Schoenfeld, Alan H. (1987) *Chapter 8 What's all the Fuss about Metacognition?* Cognitive Science and Mathematics Education. Lawrence Earlbaum and Associates; New Jersey. (189 - 215)

Three types of Metacognitve skills

1. Your knowledge about your own thought processes. How accurate are you in describing your own thinking?
2. Control, or self-regulation. How well do you keep track of what you're doing when (for example) you're solving problems, how well (if at all) do you use the input from those observations to guide your problem solving actions?
3. Beliefs and intuitions. What ideas about mathematics do you bring to your work in mathematics, and how does that shape the way that you do mathematics?

Schoenfeld, Alan H. (1985) *Chapter 9: Verbal Data, Protocol Analysis and the Issue of Control* Mathematical Problem Solving. Academic Press Inc. Harcourt Brace Javanovich; Orlando…(270 - 344)
Discusses protocol analysis. Possible problems with interviews and how interviewing might disrupt the solving process. Points out that interrogating a person about why they decided to do this or that might train them or at least change what they do. Lots of discussion that comes to the conclusion of more recent literature that they call here noninterventionist methodology. Further discussion on coding. In 70's coding got very complicated to the point of a 2 page dictionary of actions such as read the problem, drew a figure, hesitated, started putting info together, yield an equation…. Later work narrowed down to "coding scheme for heuristic processes of interest" focusing on five heuristic processes related to planning, four related to memory for similar problems and seven related to looking back. Kantowski, 1977 Journal in Research for Mathematics Education Analysis was statistical and compared different procedures with product success to see what is important. *Coding is generally on the tactical level. Schoenfeld says no research has focused on strategic decisions and their impact on PS performance.*

"The most important events in a problem session may be the ones that do not take place – for example, when a student does not assess the current status of a solution or the potential utility of a proposed approach, and as a result goes off on a wild goose chase that guarantees that the problem-solving attempt will fail.

Schoenfeld proposes analyzing problem-solving using: Read, Analyze, Explore, Plan, Implement and Verify. Each episode is analyzed separately. Attempts were made to analyze the transitions between episodes and the management of these and it proved to be completely unwieldy so they just analyze each episode into the six possible processes listed above. The purpose is to study control behavior. Mentions local and global assessment.

Says there is lots more that needs to be done. Especially answering questions regarding characterization of monitoring, assessing and decision-making processes. Calls these executive level decisions. Points out that assessment is not necessary and actually gets in the way unless something untoward occurs.

Schoenfeld, Alan H. (1985) Chapter 10 The Roots of Belief Mathematical Problem Solving. Academic Press, Inc. Harcourt Brace Javanovich; Orlando…(353 - 387)

Schulz, L. (2007) Naïve Physics/Savvy Science: Causal Lerning in Very Young Childeren … and the Rest of Us Presented at the 2007 Physics Education Research Conference.

Why are children so good at learning but so hard to teach?

Schwartz, D. (2007) Socializing Learning and Transfer. Presented at the 2007 Physics Education Research Conference.

Started with examples of why verbal protocol is not adequate for measuring what students know. Showed an example of the number line and our ability to respond to two numbers and determine which is larger. He used numbers ranging from 1-10 and then threw in negative numbers up to 20. Showed that there are different regions where we react quicker yet if we were asked to say where we would react quicker, we wouldn't be able to. Measured reaction based on pressing a button on the side that the larger number appeared. For the audience he had us squeeze our fingers together on the right side if the right number was higher. Data shows that Kindergartners are slowest but by 7th grade students are as fast as adults.

The point of this section seems to be the need to measure not just facts and concepts through verbal interviews but to also look at perceptual (procedural) knowledge as well. To do this, verbal does not always suffice. That is what this number test is able to measure.

Discusses how just the mere thought of social interaction enhances learning in an innovation situation but not in efficiency. Testing the effectiveness of the teachable agents. Some students are told that this is a real person in the other room providing feedback through the computer (Avatar) and others are told that it's just a computer (Agent). The students who think it's an Avatar show better learning than the students who believe it's only an Agent (computer). A measure of body heat is taken during the testing and students in the Avatar situation have an increase in body temperature when they are interacting with the Avatar but the students in the computer situation do not.

Schwartz, D. L., Bransford, J. D. and Sears, D. (2005). Efficiency and Innovation in *Transfer Transfer of Learning from a Modern Multidisciplinary Perspective* edited by Jose Mestre. Information Age Publishing; North Carolina (1-52).

This paper is a nice argument for a new focus on the type of assessment used to evaluate transfer. Transfer literature includes a variety of seemingly conflicting perspectives. Some argue that transfer is rare; others argue that transfer is ubiquitous; still others worry that transfer is an unworkable concept. The authors of this paper argue that all of these perspectives are pieces of the truth. The problem lies in how transfer is evaluated.

"Classic Definition of Transfer"*: "the degree to which a behavior will be repeated in a new situation.*

The authors then proceed to motivate an alternative definition of Transfer to include: *preparation for future learning.*

This leads into the two types of assessment: "sequestered problem solving" (SPS) and "preparation for future learning" (PFL).

Transfer itself is then broken into two categories; "transfer-out": *Learning used to solve a problem (apply acquired knowledge to solve a new problem with different surface features.)* and "transfer-in": *Learning used for learning. (How prior knowledge affects future learning.)*

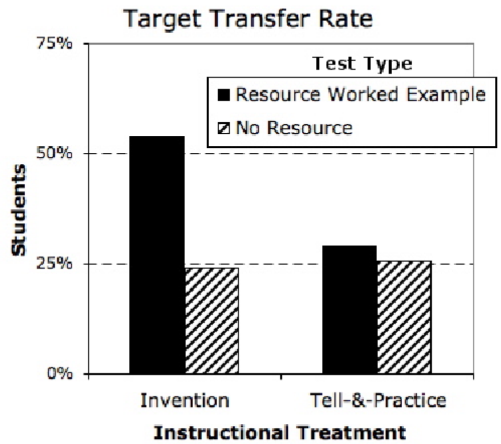
Broudy's three kinds of knowing: "Replicative": *Regurgitation of fact,* "Applicative": *See "transfer-out"* and "Interpretative": *what one notices about new situations and how one frames problems.*

Innovation: *Student creates their own solution to a completely unfamiliar problem*
Efficiency: *Refers to routine solutions where one solves problems (if they can really be thought of as problems after this) of a similar nature in a similar environment.*
In situ theorists: *Structure and development of human psychology emerges from practice.* "Standard" Cognitive Theorists"
Favorite definition: "resistance to premature assimilation": *High confidence coupled with low "competence" is a dangerous combination for the prospect of future learning.*

The two types of assessment explain the apparently contradictory views of transfer literature. SPS "makes us look dumb←???" and PFL "make us look smart. This is nicely illustrated with many examples of studies such as the Burgess' Eagle Challenge. 5th grade students and college students each had to create statewide recovery plans to increase the number of bald eagles in their state. This assessment is considered an SPS type. An extension to this study involved telling the participants to create a list of questions they would ask to help them work out the problem. The questions revealed quite a difference in the two groups. The 5th grade questions were somewhat superficial but still useful while the college students' questions were much deeper and more useful. The Challenge was later used with K-12 principles who showed even better problem solving abilities with respect to their use of resources and their willingness to let go of their initial ideas. This "tentativeness" was not seen in the college students who blissfully muddled on with their original ideas not even considering they could be based on incorrect assumptions.

Typical assessment was also viewed in the lens of Broudy's three kinds of knowing. He says that most assessments focus on replicative and applicative knowledge which encourages schools to focus on memorization and procedural training. What we're missing is assessment on interpretive knowing. Replicative and applicative knowledge are useful ways of knowing that can be brought to bear in problem solving but existing schemas and knowledge base are not always adequate. This is where interpretation becomes important. An example is an instructional technique used to learn about place value and bases in mathematics. The story about counting soldiers does not provide the knowledge to answer questions directly but sets the stage for subsequent learning.

The next section of the paper details three studies that illustrate the differences between instructional techniques. In all three cases SPS type assessment showed equivalent disappointing results for transfer; however, PFL assessment revealed instructional techniques that included innovation and efficiency were far superior to other forms of instruction that used innovation or efficiency but not both. An example with 15 classes of 9th grade students. Two weeks of equivalent instruction and then on the last day they received different treatments.



Invention Condition: Invent own solution for comparing high scores. Tell and practice: They were shown how, then practiced while teachers answered questions and corrected.
Posttest had a worked example that included summary measures (which had not been seen by the students previously) for only half the students.

The idea is to form their theory that the most effective learning is done with both innovation and efficiency at the same time. Teaching the two separately may be effective; however, not nearly as effective as the two together.

They conclude with a push for their "working smart" assessments.

Comments

Throughout the paper little comments about motivation and attitude can be found. This is not addressed directly but is mentioned on a regular basis as an underlying important idea to consider.

The authors point out that so much research is focused on routine-problem solving the point is lost about innovation. "There are a host of useful and sophisticated problem-solving routines. Nevertheless, they still are often taught as script-like, mechanical routines-often because this is the only way to show effects when they are assessed through the efficiency-oriented lenses of applicative problem solving…. Learn to break free of old routines and discover new ideas on their own." Such as no more problem-solving research trying to turn novices into experts in one semester.

Students regularly transfer knowledge of popular culture into school tasks. In contrast, many lament the large number of transfer failures. Schwartz believes this is because the failures are studying transfer out where people fail to apply acquired knowledge to solve a new problem with different surface features.

"Learning to learn (or general problem solving skills) …is important, but it tends to stress content-independent strategies for learning information presented by others. And it underemphasizes the important role of content knowledge for shaping people's interpretations of new situations."

Shekoyan, V. and Etkina, E. (2007) Epistemic Cognition in Physics Problem Solving: Experts and Novices. Presented at the American Association of Physics Teachers 2007 Summer Meeting.

Very interesting abstract: Epistemic cognition occurs when a person is solving a problem that does not have one right answer (an ill-structured problem) and thus she/he has to examine different possibilities, assumptions, and evaluate the outcomes. Epistemic cognition involves thinking about the limits of knowing, the criteria of knowing, identifying the assumptions made, looking at limiting cases, etc. and is used in ill-structured real-life and professional problems. How do we measure epistemic cognition? How does an experts epistemic cognition differ from a novices? To answer these questions we have conducted video-taped interviews with experts and novices (physics professors and students). During the interviews we asked the subjects to solve ill-structured physics problems. Using reflective questions, we encouraged them to review and possibly to refine their solution. They were prompted to search for criteria to validate different approaches. We present and discuss our analysis of the transcripts of the interviews.

I checked the term Epistemic Cognition from Kitchner, 1983 Cognition, Metacogntion…. The study was of interviews of experts and novices. They would prompt students to demonstrate epistemic cognition and then count the number of prompts it took to get them to do this. In his talk Vazgen defined the term nicely but then when he used it and showed examples it was simply of monitoring and problem decisions. They did not appear to be identifying the assumptions made or the limits of knowing as defined above. Somewhat hard to understand and follow so I may have misunderstood.

Sherin, B. (2007) Cognitive Science: The Science of the (Nearly) Obvious Presented at the 2007 Physics Education Research Conference.

Everyone has this sort of Nearly Obvious knowledge but we are not aware that we possess it. His example is the alphabet and the fact that almost everyone knows it in chunks. Some of which are common amongst everyone with the others in several possible combinations that are common. l m n o p is always together. Could be because of the alphabet song but not sure. Says cognitive scientists refer to this knowledge as implicit or tacit knowledge and that we don't have names for the specific bits that physicists have and use.

Simon, H. A. (1980). Problem Solving and Education in Problem Solving and Education: Issues in Teaching and Research edited by Tuma, D. T. and Reif, F. Lawrence Erlbaum Associates; New Jersey.

Sing, C. (2007) Problem Categorization in Teaching and Learning Presented at the American Association of Physics Teachers 2007 Summer Meeting.

First repeated Chi's study with her own problems. Grad students did well, intro students not as well but better than Chi's.

Asked TA's to categorize how they'd expect their students to do it. Finally instructed students to think if they had a sister or other relative who had just finished the course how they might do it. These instructions worked and they did it. She reported that they did a good job and showed bar graphs showing % if categories that were expert-like or based on deep structure. No data presented on what the category titles were or if all grad students did about the same or if some were more novice like and some less so. All data was aggregate and hard to see much of anything. Grad student feedback was that they appreciated the task and it helped them see the difference after they were experienced.

Sonnentag, Sabine, Niessen, Cornelia and Volmer, Judith (2006). Expertise in Software Design in The Cambridge Handbook of Expertise and Expert Performance edited by K. Anders Ericsson, Neil Charness, Paul J. Feltovich and Robert R. Hoffman. (373-387)

Entire article summarized using the division of studies that compare operational expertise which is based on years of experience versus high performance. Generalizations cannot be easily made between the two lower categories novices and moderate performers. Breaks out categories for studying programming to include: Requirement analysis and design; Program comprehension and programming; Testing and debugging; Knowledge; Communication and cooperation. Many similarities in results compared to other problem solving literature. Novices preferred a strategy in which plans are linearly implemented reflecting less varied knowledge. Studied experts in one language using a new unfamiliar language. Found that experts do have abstract transferable knowledge and skills. There were what appeared to be contradictory results from studies on transfer of knowledge with different languages; however, the study that did not show experts being stronger categorized people with length of experience as experts rather than high performers. The other studies looked at high performers and found that there were many skills that did transfer. They labeled these as adaptive experts as opposed to routine experts (Hatano and Inagaki, 1986).

Programming and Program Comprehension had an interesting note about high performing experts while explaining their programs translated what the program did into real world terms rather than only using the programming/domain terms. Testing and Debugging found high performers searched more intensively for problems and showed more information-evaluation activity. The experience studies showed differences that almost solely demonstrated knowledge base.

Communication and Cooperation is a big part of this article. Experts show adaptation to specific situational constraints. In structured meetings no difference was seen between low and high performers behavior; however, high performers behavior in unstructured cooperation situations mirrored their behavior when working on design problems in individual settings. High performers showed process-regulating behavior.

Lots of really good conclusions and future research questions. 1) Conceptualization of Expertise. Studies have not shown correlations with length of experience versus high performance. Even stronger in computer science than other fields because the field is changing so rapidly. What is important is adaptive skills and communication skills. 2) Issue of Causality 3) Task and situational characteristics. 4) Role of Motivation and Self regulation. 5) The question of how expertise develops.

Sweller, John (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science 12*, 257-285.

This article discussed the possibility that means end analysis retards students learning because of the additional cognitive load imposed. This was done by both creating a computer program to demonstrate the additional load of means end analysis versus open-ended (no goal) solving. Then an experiment was discussed where students solved various trig problems where either they knew the goal or they were just to solve everything. Then they were required to reproduce the solution. The reproduction was done to test what the students had learned since the idea is that means end analysis blocks learning.

The introduction presents previous work. It is clear from the introduction that the types of problems used are exercises for experts. "The same cognitive structures which allow experts to accurately recall the configuration of a given problem state also allows immediate moves toward the goal from the givens. … Novices, not possessing appropriate schemas, are not able to recognize and memorize problem configurations and are forced to use general problem-solving strategies such as means-ends analysis when faced with a problem." When discussing Chi categorization tasks they say that experts grouped problems based on "solution mode" "Novices, not having sophisticated schemas of this type must resort to surface structures when classifying problems." "In summary, the expert-novice research suggests that domain specific knowledge, in the form of schemas, is a major factor distinguishing experts from novices in problem-solving skill." This is a matter of definition. Rather than having discovered a difference between expert and novice problem solvers, the research has managed to demonstrate that the different people that they chose to label as experts and novices reflect the basis that they used to define experts and novices (experience). Using the word problem solving skill is not accurate since they have defined this on the fact that the experts do give the correct answers and the novices do not always give the correct answers. The experts are not problem solving so have not demonstrated their problem solving skills.

When discussing Sweller and Levine 1982, Sweller states "when given a conventional goal…" Later he says "in most circumstances, means-end analysis results in fewer dead ends reachedHe also says "if subgoals have been used, a goal stack must be maintained." In 1982 article he references Mawer and Sweller, 1982 where they used subgoals to distract the solver from means-end analysis and it helped them do a better job of learning the structure of the problem.

The computer program modeled means end analysis by finding an equation that had the goal in it and then calculating the unknowns. If an equation wasn't available where the only unknown was the goal then it continued finding equations until it could solve for the goal. The comparison program was one where the computer program simply solved for anything that it could not knowing the goal state.

Next a study is described where 24 students solve a series of 6 problems. The problems are all triangles that need sides calculated by solving for one side and then another with the use of trig functions. The goal group is told which side to solve for and it requires solving an interim side to get to the solution. The other group is told to solve for anything that they can but are specifically told not to solve for certain sides (these being the sides that are unnecessary to get to the solution specified for the other group.) It also says that students were not required to give numerical values, just the relations. This is at the end and appears to go with the no goal group but I think they really meant it for all students. If not, that would be a substantial difference. The second part of the task is that after they solve a second problem, they are asked to recreate the problem before it. So solve 1, then solve 2, now recreate 1, solve 3, recreate 2 etc… This recreation is supposed to demonstrate schema acquisition. I find this a serious stretch since there is no schema tested by memorizing the specific details of these triangles. There were very few differences found. The time to solution and time for recreation were the same for both groups. The number of errors in solution was the same. With the recreation there were less errors for the no goal group in angle position, side value and solution; however, the same error in segment labels, angle value, side position. So very minor differences. Sweller argues that "it appears that most subjects allocated excess capacity to the second task resulting in the usual performance difference occurring on that task rather than the primary one." since his results were not what he hoped for.

In the conclusion Sweller says that "Goal attainment and schema acquisition may be two largely unrelated and even incompatible processes." This statement can be true depending on the problem.

Sweller, John and Levine, Marvin (1982). Effects of Goal Specificity on Means-Ends Analysis and Learning. *Journal of Experimental Psychology; Learning, Memory and Cognition*. *8*, 463-474.

Comparison of students solving mazes when they knew the goal and when they didn't. The mazes were specifically constructed so that there was a pattern to learn. They wanted to see if students who didn't know the goal could learn the pattern easier than those who did. They hypothesize that means-end analysis blocks learning.

The mazes were finger mazes. Students were blindfolded and used their finger to trace out the path. If they were the students who were told where the goal was this was done by placing their left index finger on the goal while the right index finger tried to find the path. If they didn't' know the goal, then the left index finger was not used. There were ten subjects in each group. The path of the maze was a series of steps moving directly away from the goal until the far corner was reached and then the maze followed the perimeter of the border to the goal which was in the opposite corner as the end of the step part of the maze.

Experiments 1 and 2 used 20 and 24 subjects respectively. Half had their left finger on the goal and half did not. They found that the Goal group made more errors before finding the solution than the no goal group. Experiments 1 had 4.5 errors compared to 3. In the experiment 2 there were twice as many steps. This time the Goal group had 7.8 errors compared to 5. In their experiment they also retested the same students in an attempt to see if the Goal student learned the pattern as easily as the no Goal students. They found no significant difference.

Experiment 3 used a computer rather than a finger maze and 20 subjects. The computer screen showed the possible directions that were allowed at each location with arrows. After each move, the next set of arrows appeared. For the Goal group a solid box appeared where the goal was. In the bottom left hand corner of the screen while the maze stepped off into the opposite corner. With this maze only 1 Goal student solved the problem within 298 moves. 9 no Goal students solved the problem. They only had the subjects who had solved it the first time repeat. The 1 Goal subject solved it again without any wrong moves. 8 of 9 no Goal subjects solved it in the minimum number of moves (didn't say if the 9th solved but made errors or didn't solve it?). They repeated this experiment but added extra steps to the dead ends. 8 of 10 Goal subjects solved the problem and then solved it again with no errors. 9 of 10 no Goal subjects solved it. Five of these resolved with no errors. Two of the remaining four didn't solve it on the second try but the other two did after making errors.

Authors feel that these tests show that means-end analysis blocks learning. There seem to be other explanations. In the first examples students were provided with additional information for solving the problem. They used it and since the information was counter productive (led away from the correct path) it inevitably added to the required number of moves but students did solve the problem and had showed no difference on redoing the problem with the goal and no-goal groups. So one could say that each learned the pattern and the goal group also had to rule out the usefulness of knowing the goal.

The third experiment makes me think that it could be more serious. I suspect since the subjects didn't solve it in 298 moves, they were getting frustrated and just trying stuff. Evidence of this is the two goal students who did solve it the first time did not solve it the second time. If this is the case, something else could be happening here. Performance mode or something related could be happening here with the Goal students.

Experiment 5 uses numbers but is similar to the maze. The first number is 30 with two possible next moves either 29 or some random number between 30 and 99. If they choose the wrong number the next possible move given is 30 only. This way they know they have to go back. If they choose 29 the next possible moves are 30, 28 or another random number between 30 and 99. If they choose the wrong number then the only possible move after that number is 29 so back one move. If they choose 28 then they get three choices again 27,29 and random number. This goes on until they get to 15. Then the possible moves are 99 and 16. 99 is the goal state. 30 students do the test 15 know the goal is 99 and 15 don't know anything about the goal. As you can guess the students who know the goal take a little longer to figure it out than the students who don't. Again, they say the use of means end analysis blocks the students from learning the pattern. Means end analysis is a tool where you use the goal which is an important part of the problem and in every problem you give students the goal is a logical location that you have to find.

“The authors say that all problems require some moves away from the goal (they tend not to be classified as problems otherwise).” Experiment 5 is more interesting. It’s a maze the starts and ends in almost the same place but takes a round about path to get there. At each move there are arrows representing the possible moves. The arrow that shows the correct move is solid and the other arrows are hollow. So there is a clue to catch on to. 20 students. All 10 no Goal subjects solved it and 9 of the Goal subjects solved it. During the second trial, 7 no Goal subjects solved it without error and only 2 Goal subjects solved it without error. So most of the no Goal subjects caught on to the clue but only two of the Goal subjects apparently saw it. The authors conclude that “rule induction involves learning and means-end analysis does not. They claim further that “the more problem solvers know of the goal state, the less they learn of the problem structure during the solution process.” Early on in the article they refer to a study by Mawer and Sweller 1982 that found that numerical transformation problems can be solved by means-end analysis without the solver being aware of the transformation rule. In this study they provided sub goals along the solution process that helped the solvers readily induce the rules. Their claim is that these sub-goals distracted the solvers from the goal so they didn’t use means-end analysis. This adds more cognitive load and although they don’t use the terminology of cognitive load until later papers, they do say in the conclusions that the reason means-end doesn’t work is that it concentrates problem solvers attention on the difference between where they are and the goal state taking their attention away from rule induction. In most of their experiments, instead of guiding students to something useful, they were guiding the other students away from something useful.

Tominoro, J. and Redish, E. (2007). Elements of a cognitive model of physics proble solving: Epistemic games. *Physical Review*

Talk about how instructors could cue students to switch the e-game that they are playing to help them solve a problem. They talk about groups but not individuals. I see so many individuals who cannot be cued to map meaning to mathematics. Games could be most useful if one watched a particular student and saw that that student never engaged in a particular game. I have seen students who use recursive plug and chug or Transliteration to Mathematics but never any games that involve making meaning. Elem. Ed majors for example. Can one focus on teaching a particular game or is this too large of an idea to tackle at once. Like trying to teach sense-making rather than specific problem solving skills.

This paper had some good points about experts “Students in introductory physics courses are far from experts, so using scientists’ approaches to inquiry as a norm by which to describe students’ inquiry would not be appropriate. They also discuss the difference between the way a typical teacher would solve a problem. They say a teacher has tightly compiled resources (knowledge). However they also say richer collection of problem-solving strategies (i.e. an assortment of epistemic games for solving problems in physics) than most students. For the typical teacher, the problem statement immediately cues the appropriate epistemic game and tightly compiled resources. I would argue that the appropriate e-game for the teacher with the tightly compiled resources is not the same as the appropriate e-game for the student with the loosely compiled and not as complete set of mathematics and physics resources.

Page 45 discussion between Martha and Susannah has been labeled as the beginnings of Physical Mechanism but then changed to Recursive Plug-and-Chug. If I understand your classification of e-games correctly, I believe if Susannah had suggested $p = \square gh$ instead of saying the game switched you would have said it was Mapping Meaning to Mathematics. Is this true? If so, it seems a weakness in your definition if the correctness of the students discourse changes the e-game. In your definitions of e-games this did not seem to be part of the criteria. The idea of students creating understanding is part of your definition of the difference between games but not whether the understanding is correct.

The lack of clarity of game choice between Martha and Susannah also brings up another point that has surfaced throughout the case study section. To see a particular e-game, in most cases you have focused on only one student at a time. In the case of Darlene playing Transliteration to Mathematics, it is only Darlene doing this without understanding the sample problem. If one focuses on what Alisa is saying (harder to do without seeing the video tape) it appears that she is making meaning of this example and so is she playing a different e-game? Possibly a 7^{th} game not listed in this paper where one maps meaning onto sample problems? If I only look at one person at a time again, I can make more sense of Susannah and Martha. Martha is beginning to play Physical Mechanism or possibly Meaning to Mathematics. Then Susannah jumps in with Recursive Plug-and-Chug, avoiding all sense-making including listening to the TA. Martha is thinking and listening to the TA and Susannah while trying to make some meaning out of all of these inputs. So what game is Martha playing?

The ideas presented in this paper are all very interesting, thought provoking and provide a new way to view some of what students are doing while attempting to solve problems. Classifying students' problem solving into different epistemic games could be a constructive way for teachers to think about students actions while solving problems; however, it is somewhat limited. Even then there are many questions that could surface about various sections of discourse such as: (1) Which student is used to classify the e-game?; (2) Was one game started and then changed to another?; or (3) Has the choice of incorrect information slowed down the students' ability to make meaning clouding our ability to classify the appropriate game?

I do see how this framing could be potentially useful for researchers and curriculum developers if they could use it to think about how to cue appropriate games. Again, however, this gets tricky since the appropriate game varies with the problem and some problems require more than one game to be solved. I could also imagine situations where the appropriate game depends on the students' progress so far. This leads one to think maybe it's most useful for the teacher to consider in midstream; however, as mentioned above, it's awfully difficult to analyze the games in real time.

Voss, J.F., Green, T.R., Post, T.A., & Penner, B.C. (1983). Problem solving skill in the social sciences. In G. Bower (Ed.), *The psychology of learning and motivation, 17,* New York: Academic (165-213)

Ward, M. and Sweller, K (1990). Structuring effective worked examples. *Cognition and Instruction* 7, 1-39.

This article found that certain types of worked examples helped students solve exam problems and others were not. It was not explicitly stated; however, it seemed clear that no text book was offered to these students so they couldn't use the typical examples found in a text. 42 students were from two classes, each split in half. The topics were mirrors and lenses. The students were given a lecture and the materials from the lecture which gave the information and told how to use it. Then students tried two example problems in class and were helped whenever they asked questions. Since classes were split in half, all students received the same lecture together. Then they were given a set of 10 problems to take home. With the control group they were told to go practice the problems and then they'd be tested on them the following day. The experimental group's problems had one of each pair of problems worked out (showed ray diagram solution but no description). They were also told to go study the problems and then would be tested on them the next day. The first time this was done the topic was on Mirrors. The experiment was then repeated for the section on lenses but the control and experimental groups were switched. In both cases, mirrors and lenses, the experimental group did better by about 20% (60% vs 80%). The author thinks, "Worked examples reduce cognitive load by switching attention away from goal-directed search, thereby assisting in both schema acquisition and rule automation. "

During the first experiment there had been low practice problems (i.e. 8 problems on concave mirrors and only 2 on convex mirrors) plus two transfer problems. When testing, students were timed and did not attempt all problems. Most of the problems they didn't attempt were the transfer problems. So an effect was seen on the low practice but not on the transfer problems. The next experiment was the same as the first (32 students form one class – switched when topics switched) except the two transfer problems were required and given separately. The students had five minutes to solve each. In addition the first problem was collected before the second one was handed out. In this case there was a striking difference in performance on the transfer problems. The worked example students did much better.

The third experiment used kinematics worked examples (linear motion, projectile motion and collisions.) There were 17 students in each class a year apart. The worked examples alternated with topic and in the next year the alternation was opposite. In these cases, there was very little difference between performance. The solutions were designed so that they wouldn't mean anything without referring back to the problem statement. The author claims "attention must be directed to mentally integrate its various facets. This process of integration will require cognitive resources to be devoted to activities that, at best, are marginal to schema acquisition and rule automation." There may be additional cognitive load required with this sort of example but a more serious difference between the two types of worked examples given is the additional information given in the second one.

The worked examples stated the problem and then separately gave a solution such as

A car moving from rest reaches a speed of 20 m/s after 10 seconds. What is the acceleration of the car?

u = 0 m/s
v = 20 m/s
t = 10 s

$v = u + at$
$a = (v - u)/t$
$a = (20 - 0)/10$
$a = 2\ m/s^2$

In experiment 4 the worked examples had integrated solutions.

A car moving from rest (u) reaches a speed of 20 m/s (v) after 10 seconds (t): [$v = u + at$, $a = (v - u)/t = (20 – 0)/10 = 2\ m/s^2$]. What is the acceleration of the car?

Not surprisingly the results of Experiment 4 showed these integrated worked examples to be more effective for student learning.

The author claims "The more detail and assistance a worked example provides, the more difficult it is to format the problem with a unitary structure. Additional information intended to be helpful to students but not strictly necessary may be difficult to integrate physically with essential, core information, leaving students to accomplish the integration mentally, with deleterious effects." This is sound advice; however, it is ironic that the author's example of a worked example that is supposed to have additional information actually is *missing* crucial information. So the author's advice is sound but not applied correctly in practice. A more serious difference existed between his problems but he seemed unaware of this difference only focusing on cognitive load. This is the case throughout Sweller's work.

In the final experiment (6) Sweller took the worked examples from the Mirror and Lenses units and added a large textual explanation to the ray diagrams. In these cases he was able to hurt the students learning with the worked examples.

In the conclusion Sweller is trying to understand results by other researchers where better students are able to more fully and correctly explain in detail what is happening in worked examples and were more aware of their own failures to comprehend than poor students. Sweller claims this can be understood from a cognitive load perspective because better students have more cognitive capacity than weaker students. Motivation, expectations, knowledge structure etc… are much larger, more important features. Just from this brief description of what Chi and other researchers found when interviewing students using worked examples, it sounds like their knowledge structure, metacognitive skills and possibly reading comprehension would all be more likely explanations for these differences than cognitive load.

Wineburg, S. (1998). Reading Abraham Lincoln: An Expert/Expert Study in the Interpretation of Historical Texts. *Cognitive Science, 22* (319-346).

This paper starts out with a nice review discussing the usual chess studies and the early work in physics (Chi, Larkin, Simon & Simon). Glaser review says "the quintessential expert possessed rich networks of highly-elaborated knowledge and myriad problem-solving templates that smoothed the way for the filled processing of new information. This process went on with lightening speed and, compared to novices, relative ease." This description does not fit what the author or other work in writing and history say about experts in their areas. For example in written composition (Scardamalia and Bereitwer 1991), found instead is that compared to novices expert writers took more time executing tasks, detected more problems in their writing, agonized longer over revisions, and spent longer time puzzling about the "rhetorical space" of their compositions. "Rather than fluidity and rapidity, the writing process of these experts was characterized by nagging propensity for finding flaws at every corner. Similar results have been seen in history. In a comparison of high school students to historians (Wineburg 1991 and 1994), the historians, not the students, echoed pangs of doubt about their interpretations, second guessing themselves and appending strings of qualifications to their conclusions. Novices on the other hand, quickly formed interpretations and typically never looked back.

The author goes on to discuss Schoenfeld and his number theorist that he had solve the geometry problems that students with more knowledge on that topic couldn't solve. He terms these sorts of things as "adaptive" expertise rather than "routine". He says adaptive expertise (Hatano & Inagaki, 1986) speaks to the ability to apply, adapt and otherwise stretch knowledge so that it addresses new situations – often situations in which key knowledge is lacking. Viewed from this vantage point, expertise is less the rapid firing and deployment of knowledge than the ability to pick oneself up after a tumble, work through confusion, and reorient oneself to the problem at hand.

He says these two images of expertise – the nimble and quick problem solver and the resourceful and persistent *bricoléur* - may not be in conflict, but may speak to two different aspects of expertise. He goes further to reference studies of medical experts. The specialist whom you'd want for bypass surgery or the general expert such as a family physician who'd you'd rather have sitting next to you on the plane when you had a heart attack.

This study has two historians looking at seven documents relating to Abraham Lincoln's view on blacks versus whites. Three documents were from Abraham Lincoln spanning 21 years of his life, one from Stephen Douglas and three from historical Contemporaries. The idea was to understand what Lincoln actually believed about race and if this had evolved over time.

Both historians were top historians with strong degrees, full professorships and had authored many books. One was an expert in this area focusing on Abraham Lincoln and the Civil War while the other specialized in American History but not Abraham Lincoln, the Civil War or Reconstruction.

The author also did this same exercise with 14 college history majors and non-majors in a previous study. The comparisons to these 14 students are quite useful; however, many of their conclusions about the differences between the two Expert problem solvers are over generalized for only a study of two. For example they apply significance to the idea that one historian took longer than the other. Or that the content specialist drew more analogies than the general historian and used this 'fact' to demonstrate again how this is the opposite of scientific expertise and problem solving in the two domains is different.

In response to recognizing that one of the hardest and most important tasks that an expert historian has to perform is creating coherence between different pieces of text that are separated by context and time. He says "an important question in any study of expertise is how experts get to be that way. ….Snapshot studies such as this offer little…. But the study of expertise must also address a second key question: How is it that experts keep learning? Why do they continue to get smarter from encounters with materials and situations that leave other problem solvers unfazed (Holyoak, 12991; Perkins & Saloman, 1989)" This is a nice lead in to the comparisons with the college students performance on the same task.

The students had a variety of responses but the two trends that stuck out were the students that took Lincoln's words at face value without considering context or the passage of time. Other, more careful, readers recognized that they needed a context but rather than choosing a context of the 1860's they chose one from their contemporary social world. The historian who was not a specialist in this area had the same immediate reaction (we cannot help activation based on loaded topics – race) but was able to step back from this immediate response and place Lincoln's words into the context of the 1860's and what these words meant at that time. The specialist "spoke" the language of 1860 so had much less difficulty doing this but still had to look up a few key phrases to see if they were indeed phrases used at that time or just words.

In conclusion the author points out the non-specialist had extensive factual knowledge but that is not what stood out while observing his solution process. "Once he was immersed in the documents it was what he didn't know that came to the fore: his way of asking questions, of reserving judgment, of monitoring affective responses and revisiting earlier assessments, his ability to stick with confusion long enough to let an interpretation emerge. It was how he responded in the face of what he didn't know that allowed him, in short, to learn something new."

Zwart, J. W. (2007) Doing It Wrong So They Get It Right. Presented at the American Association of Physics Teachers 2007 Summer Meeting.

Gave students worked examples with mistakes. The only data is a student survey of what they believe it helped them do better.