



Story of a Research Program

John Sweller



Early Days

I was born in 1946 in Poland to parents who, apart from my older sister, were their families' sole survivors of the Holocaust. Very few family members who lived outside of Poland survived. One of those was my mother's sister, my aunt who lived in Adelaide, South Australia. She had become a dentist in Vienna and was fortunate that Nazi regulations did not permit her to practice her profession prior to the war. In 1938 she and her family left for Adelaide, Australia. My aunt was my mother's only surviving relative, and since Adelaide was almost as far from Europe as my parents could find, that was where my parents, my sister, and I landed in 1948.

My parents' native language was Polish rather than the commonly spoken Yiddish of Polish Jews, and so my first language, strictly speaking, was Polish. In practice, given my age, English supplanted Polish very rapidly and was my de facto first language. I certainly could understand and speak English prior to arriving at school. As is regrettably common among native

English speakers, it became my only language.

At school, I began as a mediocre student who slowly deteriorated to the status of a very poor student by the time I arrived at the University of Adelaide. (Most Australian students attend their home university.) Initially, I enrolled in an undergraduate dentistry course but never managed to advance beyond the first year. While I am sure that was a relief to the Dental Faculty, it also should be a relief to Australian dental patients.

Given the physical proximity of the teeth and brain, I decided next to try my luck at psychology. It was a good choice because my grades immediately shot up from appalling back to mediocre, where they had been earlier in my academic career. I decided I wanted to be an academic.

That decision was not as silly as it sounded. While I was no better at sitting for exams or obtaining good grades on normal assignments than I had ever been, on some occasions we were given research assignments requiring us to devise psychological experiments. Leon Lack, then a tutor in the Department of Psychology, instituted these. At that time, the University of Adelaide's Psychology Department was emphatically oriented towards experimental psychology. I seemed to have some skill at theorizing about psychological variables and devising experiments. It was the only academic skill that I had ever rated myself as better than average. As I advanced through my undergraduate years, my grades gradually improved. While they never reached any stellar heights, they now were sufficient to allow me to enroll as a PhD student, a

degree that under the Australian system concentrated entirely on research with no coursework. I was finally in my milieu.

In 1970, under the capable oversight of my supervisor, Tony Winefield, I commenced my research as an experimental psychologist studying learning theory. At that time, Behaviorism was dying and the cognitive revolution was beginning. I began my work on animal learning but decided fairly rapidly that applying cognitive principles to animal learning might not be productive and so rapidly switched to human problem solving. I have conducted my research on learning and problem solving ever since, leading to my career as an educational psychologist.



University of Adelaide Campus

On completing my PhD at the University of Adelaide in 1972 my first academic position was as a lecturer in educational psychology in the teacher training program at the Tasmanian College of Advanced Education in Launceston. I was unused to living in a small town, unused to living alone rather than with my family and, of course, unused to teaching rather than solely carrying out research. I liked the college, and Launceston is an exceptionally attractive town, but I needed to leave. I stayed for a personally difficult year before moving to Sydney for an equivalent position at the University of New South Wales (UNSW), where I have remained ever since.

I was at home in Sydney. After three years, I married my wife, Susan, a 1956 refugee from Hungary, and we have two daughters, Naomi and Tamara. I was happy

to live in Sydney permanently. At UNSW, I recommenced my research career, studying problem solving.

The Beginnings of Cognitive Load Theory

After several non-descript experiments, I saw some results that I thought might be important. I, along with research students Bob Mawer and Wally Howe, was running an experiment on problem solving, testing undergraduate students (Sweller, Mawer, & Howe, 1982). The problems required students to transform a given number into a goal number where the only two moves allowed were multiplying by 3 or subtracting 29.

Each problem had only one possible solution and that solution required an alternation of multiplying by 3 and subtracting 29 a specific number of times. For example, a given and goal number might require a 2-step solution requiring a single sequence of: $\times 3, - 29$ to transform the given number into the goal number. Other, more difficult problems would require the same sequence consisting of the same two steps repeated a variable number of times. For example, a 4-step problem always had the solution: $\times 3, - 29, \times 3, - 29$ while a 6-step problem required 3 iterations of $\times 3, - 29$. Accordingly, all problems required alternation of the two operations a variable number of times.

My undergraduates found these problems relatively easy to solve with very few failures, but there was something strange about their solutions. While all problems had to be solved by this alternation sequence because the numbers were chosen to ensure that no other solution was possible, very few students discovered the rule, that is, the solution sequence of alternating the two possible moves. Whatever the problem solvers were doing to solve the problems, learning the

alternating solution sequence rule did not play a part.

Cognitive load theory probably can be traced back to that experiment. My objections to the many variations of discovery and problem-based learning also have a similar source. While the puzzle problem-solving task used had no direct educational relevance because such tasks do not form part of any curriculum, the results seemed to say something about how students learned and solved problems. It was obvious to me that if I had simply informed students to solve each problem by alternating the two moves until they reached solution, they would have immediately learned the rule and would have been able to solve any problem presented to them no matter how many moves were required for solution. Of course, since these were problem-solving experiments, I had not informed participants of the alternation rule and most failed to discover the rule for themselves.

It seemed plausible to me that the same processes might apply when students were asked to solve problems in an educational context. We give students problems to solve in subjects such as mathematics with the expectation that they would learn to solve such problems. If my experimental results were generalizable to educational problems, that expectation may not be realised. Perhaps we should be showing students how to solve problems rather than having them solve the problems themselves?

Early Theoretical Issues

The next step was to test whether educational problems had the same characteristics as the puzzle problems that I had used. Despite being an obvious step, it was not one that I took. Before testing the hypothesis using educational problems, I decided to try to determine the cognitive mechanisms that caused my strange experimental results. Specifically, why could my participants easily solve their problems

but seem to learn little from the exercise? I needed to determine their problem solving strategy and needed to analyse why that strategy was preventing learning. It took several years during the 1980s to identify the relevant cognitive structures and functions with much of the work continuing to use puzzle problems.

In the 1970s, the study of problem solving had expanded and made considerable progress. The seminal work was carried out by Newell and Simon at Carnegie-Mellon University in Pittsburgh (Newell & Simon, 1972). They had identified the characteristic strategy that humans use when problem solving: means-ends analysis. A means-ends strategy requires problem solvers to consider their current problem state, the goal state, extract differences between the two states, find a problem-solving operator that can reduce the differences, and repeat the process until the goal is reached. This strategy requires problem solvers to process



University of New South Wales Campus

in working memory all of the information concerning problem states and problem operations simultaneously. I figured that since it was well known that working memory is very limited in capacity and duration, it is likely that when a means-ends strategy was used, nothing else can be considered. The result is that problem solvers can successfully solve a problem but learn nothing from the exercise if no

information is transferred to long-term memory.

Cognitive Load Effects

This process seemed to explain why my problem solvers could solve their problems but not discover the rule that they had solved all of them by alternating the two possible moves. If students solving educational problems used the same procedures, then the use of problem solving in educational contexts should be questioned. The function of problem solving using means-ends analysis seemed to be to reach the goal of a problem, not to learn, where learning was defined as transferring knowledge to long-term memory.

Goal-Free Effect. The goal-free effect derived from this reasoning. It was the first cognitive load theory effect although in some senses, the theory derived from the effect rather than the effect from the theory. Here is the reasoning that was used. If working memory during problem solving was overloaded by attempts to reach the problem goal thus preventing learning, then eliminating the problem goal might allow working memory resources to be directed to learning useful move combinations rather than searching for a goal. Problem solvers could not reduce the distance between their current problem state and the goal using means-ends analysis if they did not have a specific goal state. Rather than asking learners to “find Angle X” in a geometry problem, it might be better to ask them to “find the value of as many angles as possible”.

You can reduce differences between where you are and where you are going if you know that where you are going is to find a value for Angle X. You cannot find such differences if you are merely attempting to find as many angle values as you can. For example, you can work backwards from a goal such as “find angle X”. You cannot work backwards from the goal “find the value of as many angles as you can.” It is a different type of goal that requires a different problem solving strategy.

The initial work on the goal-free effect used puzzle problems and was carried out with Marvin Levine during my first sabbatical at the State University of New York at Stony Brook. I discussed my ideas with him and we devised the first experiments on the goal-free effect. We established that transfer effects were substantially enhanced by the use of goal-free, puzzle problems (Sweller & Levine, 1982).

I had written to several people asking whether I could visit them during a sabbatical. Apart from Marvin Levine, none were very enthusiastic with most making it clear that a visit from me would be a nuisance. In contrast, Marvin was enthusiastic. Susan and I arrived on Long Island for a one-year stay. It was our first trip outside of Australia or New Zealand. Every weekend, we would catch the train to Manhattan, going to museums and attending concerts and plays. We wandered all over Manhattan. At that time, Manhattan had a reputation for being dangerous but all the criminal behavior must have been occurring behind us because although we were offered drugs on a regular basis, we never saw anything violent. We did avoid the subway so perhaps that saved us.

Perhaps we should be showing students how to solve problems rather than having them solve the problems themselves?

On returning to Sydney, the next step was to provide evidence that the goal-free effect applied to educational problems not just puzzle problems. Elizabeth Owen, Bob Mawer, and Mark Ward, the first two mathematics teachers and the third a physics teacher, demonstrated the effect using geometry and physics problems (Owen & Sweller, 1985; Sweller, Mawer, & Ward, 1983). That work provided us with the first cognitive load theory effect using instructional materials.

Worked Examples Effect. The worked example effect, according to which learners who study worked examples perform better

on test problems than learners who solve the same problems themselves, also derived from the reasoning that conventional problem solving interfered with learning because it concentrated on reaching a problem goal rather than transferring knowledge to long-term memory. Papers by Graham Cooper demonstrating the worked example effect using algebra problem solving were published (e.g. Cooper & Sweller, 1987). More recently, in her PhD, Arianne Rourke demonstrated the worked example effect using students learning designers' styles (Rourke & Sweller, 2009). Juhani Tuovinen demonstrated an advantage of worked examples over discovery learning in his PhD that I supervised (Tuovinen & Sweller, 1999).

Sunah Kyun, a PhD student from Korea supervised by Slava Kalyuga and me extended the work on worked examples to learning about English literature for students studying English as a foreign language (Kyun, Kalyuga, & Sweller, 2013). We have had several PhD students working on cognitive load theory to teach English as a foreign language. Jase Moussa-Inaty, whose PhD was supervised with Paul Ayres (Moussa-Inaty, Ayres, & Sweller, 2012) and Yali Diao (Diao & Sweller, 2007) worked in this area studying native Arabic and native Chinese, respectively, learning English.

Difficulties Convincing Researchers of the Problem with Problem Solving

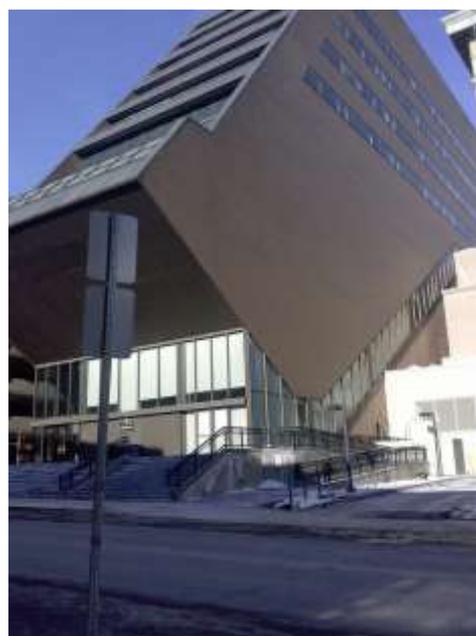
In 1984, I spent a few months on a sabbatical at the Learning Research and Development Center (LRDC) in Pittsburgh that, along with Carnegie-Mellon University, was the center for research into problem solving. I tried, far too ambitiously, to convince people that learning via problem solving was a dead-end. Predictably, I failed. At that time, writing computational models of cognitive processes was strongly emphasized and since Pittsburgh was the center for such activity I wrote and published a model, based on the goal-free effect, supporting the suggestion that problem solving imposed a heavy cognitive

load that interfered with learning (Sweller, 1988).

It was the worst possible time to be publishing papers calling into question the efficacy of using problem solving as a learning device. Our increased knowledge of problem solving, largely led by researchers in Pittsburgh, led many to suggest and most to accept that educational problem solving should be emphasized. For reasons that are unclear to me, randomized, controlled trials testing the effects of problem solving on learning were not used. The fact that evidence in support of the notion that problem solving was a relatively good way to learn was contradicted by the worked example effect, was treated as an irrelevant detail. Most of the field leapt enthusiastically on the problem solving bandwagon. The research on worked examples was treated either with hostility or more commonly, ignored, a state of affairs that lasted for about two decades.

Further Instructional Effects of Cognitive Load Theory

In the meantime, ignoring the issues the field had with worked examples, we needed to extend our knowledge of the



Learning Research and Development Center (LRDC) in Pittsburgh

worked example effect. The original experiments demonstrating that studying worked examples was better than solving problems had been carried out using algebra transformation problems such as, $a + b = c$, solve for a (Cooper & Sweller, 1987). The next and obvious step was to demonstrate that other areas such as geometry or physics problem solving also led to the worked example effect. We ran experiments comparing problem solving with studying worked examples using geometry or physics problems and found no statistically significant differences whatsoever. We were mystified. Why should studying algebra worked examples be superior to solving the equivalent problems but studying geometry or physics worked examples prove no better than solving the equivalent problems?

After several years we realized that the issue was not whether worked examples or problems were used but whether different instructional procedures increased or decreased working memory load. It was a lesson that we seemed to have to re-learn every few years. The issue could never be whether the use of worked examples was better than solving problems but rather, whether the particular worked examples used reduced unnecessary working memory load compared to solving problems. In the case of our algebra worked examples, the conventional format used to present a worked example did reduce unnecessary working memory load compared to solving a problem. In the case of geometry and physics worked examples, the conventional worked example format did not reduce working memory load compared to solving problems and so the worked examples were ineffective.

Split-Attention Effect. The issue with geometry or physics worked examples was split-attention. Learners studying worked examples in conventionally structured geometry or physics had to split their attention between multiple sources of information and then mentally integrate them. For example, if a geometry statement mentions Angle ABC, learners have to note

the angle and find it on the diagram. Until the statement and the diagram have been mentally integrated, neither can make any sense. This activity has to occur in limited working memory and the sole reason it has to occur is because of the conventional format of geometry worked examples. If instead, the statements are placed on the diagram or had arrows indicating the relations between each statement and the diagram, the worked example is physically integrated and working memory resources do not have to be expended to integrate the two sources of information. Extraneous cognitive load is reduced and learning is facilitated. By eliminating split-attention in the worked examples, the same worked example effect that was obtained using algebra problems can be obtained using geometry or physics problems.

The issue did not arise in the case of algebra problems because while the conventional way of presenting algebra worked examples does not incorporate split-attention, the conventional way of presenting worked examples in geometry and physics does incorporate split-attention. A worked example in algebra consists only of one line followed by the next line with a transformation. Learners do not have to split their attention between the statements and a diagram or between different categories of statements as they do in geometry or physics. We had been fortunate that our first attempts to test the worked example effect happened to use an area that conventionally did not incorporate split-attention, otherwise we may never have discovered the worked example effect. The effect applies to the presentation of all information, not just worked examples. Rohani Tarmizi, a Malaysian mathematics teacher enrolled in a PhD demonstrated the split-attention effect using geometry problems (Tarmizi & Sweller, 1988) while Mark Ward demonstrated the effect using physics problems as part of his PhD (Ward & Sweller, 1990). Paul Owens demonstrated similar effects in his PhD on music education (Owens & Sweller, 2008) as did Narciso Cerpa studying computer education

for his PhD (Cerpa, Chandler, & Sweller, 1996). Physically integrating disparate sources of information so that they no longer have to be mentally integrated reduces extraneous cognitive load and facilitates learning.

Modality Effect. The modality effect provided an alternative technique to physically integrating disparate sources of information. Seyed Mousavi, an Iranian PhD student supervised by colleague Renae Low and me, discovered the instructional modality effect (Mousavi, Low, & Sweller, 1995). When dealing with, for example, a diagram and text, instead of presenting the text in written form, it can be presented in spoken form. By using both auditory and visual channels, working memory capacity can be increased. Thus, the modality effect occurs when learning is facilitated by using both visual and auditory channels. It also is useful to provide markers on the visual information to indicate what the auditory information is referring to, as found by Hyunju Jeung from Korea in her PhD with Paul Chandler and me (Jeung, Chandler, & Sweller, 1997).

Transient Information Effect. When demonstrating the modality effect, care must be taken to ensure that spoken material is short and simple. Auditory material is transient with new information replacing old information, unlike written material which is permanent. Indeed, the permanence of written material and the transient nature of spoken material is presumably why we invented writing. Wayne Leahy, a previous PhD student of mine and now a colleague at Macquarie University, found that the modality effect is reversed when using long, complex spoken text (Leahy & Sweller, 2011). Any advantage of using both auditory and visual processors is negated by presenting lengthy, complex text in spoken form. Such text always should be presented in written form so that learners can easily return to given segments to ensure that they understand the text. Anna Wong and her PhD supervisor, Nadine Marcus, a previous PhD student of mine and now employed as an academic in the School of Computer

Science at UNSW, found similar effects using animation that is transient compared to static graphics that are permanent (Wong, Leahy, Marcus, & Sweller, 2012). Samuel Ng, a PhD student from Singapore who Slava Kalyuga and I supervised also studied the effects of transience due to animation when teaching physics (Ng, Kalyuga, & Sweller, 2013). These findings led to the transient information effect.

The transient information effect is interesting in that it explained some failures to obtain the modality effect. The modality effect had been discovered in the mid-1990s. It was frequently replicated over the years by many researchers but there were some notable failures and even reversals of the effect with a few studies indicating single modality presentation was superior to dual modality presentation. Of course, like all cognitive load effects, the modality effect will only be obtained when the instructional procedure reduces extraneous working memory load. A long, complex oral statement, because it is transient, will increase rather than decrease working memory load compared to a written statement, leading to the transient information effect and a reversal of the modality effect. As always, restructuring information is only effective insofar as it reduces working memory load. Just as the worked example effect cannot be obtained using split-attention information, the modality effect cannot be obtained using lengthy, complex, transient, oral information.

The transient information effect also has implications for technology-mediated presentation of information. Frequently, we introduce new technology because we can. For example, we may use spoken rather than written information or animations rather than static graphics because technology now allows us to use these techniques more readily. Commonly, the cognitive consequences of new instructional procedures are far more important than the medium used.

Redundancy Effect. In the late 1980s, Paul Chandler turned up to enroll in a PhD

and commence the most effective research collaboration I have had. His work had an immense influence on cognitive load theory, first as a student and then as a colleague. Several cognitive load effects were discovered during our collaboration. His PhD investigated the redundancy effect (Chandler & Sweller, 1991). Providing learners with any unnecessary information requires them to process that information

Redundancy is anything but harmless. Providing unnecessary information can be a major reason for instructional failure.

which can overload working memory. For example, one of my PhD students, Susannah Torcasio, found that providing beginning readers with pictures was redundant and interfered with learning to read (Torcasio & Sweller, 2010). Janette Bobis in her PhD found that including additional, complex diagrams during mathematics education could lead to the redundancy effect (Bobis, Sweller, & Cooper, 1994). Most people assume that providing learners with additional information is at worst, harmless and might be beneficial. Redundancy is anything but harmless. Providing unnecessary information can be a major reason for instructional failure.

As was the case for the split-attention effect, we found the redundancy effect by accident. In fact, the redundancy effect flowed from the split-attention effect. We had previously found that requiring learners to split their attention between a diagram and text interfered with learning compared to physically integrated diagrams and text. We erroneously assumed that all diagrams and text had the same properties. Instead, the logical relations between the two sources of information was critical. For the split-attention effect, the sources of information had to refer to each other and be unintelligible in isolation. For example, a geometry statement such as “Angle ABC =

Angle XYZ” is likely to be unintelligible without reference to the diagram. In contrast, some statements simply reiterate information that can be seen just by looking at a diagram. A diagram that shows blood flowing from the left ventricle to the aorta does not need a statement saying “Blood flows from the left ventricle to the aorta”. Such statements belong to a different category to statements such as “Angle ABC equals Angle XYZ” which are essential to understand the diagram. Redundant statements are unnecessary and processing them leads to an extraneous cognitive load. Instead of integrating them with a diagram, they should be eliminated due to redundancy, leading to the redundancy effect.

Compound Cognitive Load Effects

Compound cognitive load effects are ones in which different effects interact. The manner in which they interact, effectively provides limits to various cognitive load effects. All instructional effects have limits and those limits are just as important as the effects themselves.

Element Interactivity Effect. The first compound effect we found was the element interactivity effect. As indicated above, cognitive load theory effects depend on the imposition of a heavy working memory load by a task. Some instructional material, because of its composition, does not require extensive working memory resources. That material should not be expected to demonstrate any cognitive load effects. In order to determine whether information potentially might impose a heavy working memory load, we needed a measure of complexity. The measure devised was element interactivity. If elements of information interact, they must be processed simultaneously in working memory to be understood, imposing a heavy cognitive load. If they do not interact, they can be processed individually, with less cognitive load.

The concept of intrinsic cognitive load derived from this reasoning. It could be predicted that only when element

interactivity was high resulting in a high intrinsic cognitive load would cognitive load effects occur. Empirical work on the element interactivity effect confirmed this prediction. Unless learners find the information being processed complex and difficult to understand, cognitive load effects will not be obtained. Sharon Tindall-Ford, a PhD student of Paul Chandler and mine was instrumental in this work, especially as it related to the modality effect (Tindall-Ford, Chandler, & Sweller, 1997). The element interactivity effect suggests that no cognitive load effect can be obtained if element interactivity is low. Cognitive load theory applies to complex material that is difficult to understand.

Expertise Reversal Effects. The element interactivity effect and another important cognitive load theory effect, the expertise reversal effect are closely related because information that is high in element interactivity for a novice is likely to be low in element interactivity for an expert. Paul Chandler and I first discussed the expertise reversal effect while walking through a Sydney brewery. The brewery trained apprentices and we were there to discuss the possibility of running some training studies. Eventually, we did not run any experiments there and neither did we get any free beer. However, we needed to run some experiments testing the expertise reversal effect. Slava Kalyuga turned up to do a PhD and another great collaboration ensued. He also became a colleague in due course. Slava had almost completed a PhD in the Soviet Union before it collapsed along with Slava's hopes of completing his degree. Instead, he was exiled into the gulag of cognitive load theory from which he has never managed to escape.

The expertise reversal effect occurs when Instructional Procedure A is superior to B for novices with the superiority decreasing and eventually disappearing or even reversing with increases in knowledge levels. For example, studying worked examples may be better than solving problems for novices but with increased expertise, solving problems may be better

than studying worked examples. The reason is that while studying worked examples may be of assistance to novices, as expertise increases, worked examples may become redundant with learners needing to practice solving problems instead. Increased expertise reduces element interactivity and there is no need to reduce working memory load if few memory resources are needed during problem solving.

Alexander Yeung's PhD that I supervised with Putai Jin was one of the first to demonstrate the expertise reversal effect (Yeung, Jin, & Sweller, 1998). He studied the effect of explanatory notes associated with text. Maria Pachman, a PhD student supervised by Slava and me did work on the expertise reversal effect as it related to deliberate practice when studying mathematics (Pachman, Sweller, & Kalyuga, 2013). Kimberley Leslie, for her PhD supervised by Renae Low, Putai Jin, and me, also studied the expertise reversal effect when primary school students learn science concepts (Leslie, Low, Jin, & Sweller, 2012).

International Acceptance of Cognitive Load Theory

While cognitive load theory work was continuing in Sydney, similar research was beginning to gain some traction in the rest of the world, primarily Europe. It was largely ignored in Australia—where it continues to be ignored. The first large-scale interest in the theory occurred in Holland, led by Jeroen van Merriënboer and his brilliant PhD student, Fred Paas. They confirmed the worked example effect, and invented two new effects.

Completion and Variability Effects. The completion problem effect occurs when learners asked to complete the solution to a partially completed problem learn more rapidly than students asked to solve a problem without being shown any of the moves (Paas, 1992). The variability effect occurs when learners shown highly variable worked examples learn more than learners shown more similar worked examples (Paas & van Merriënboer, 1994). Yuan Gao, a PhD student who I co-supervised with

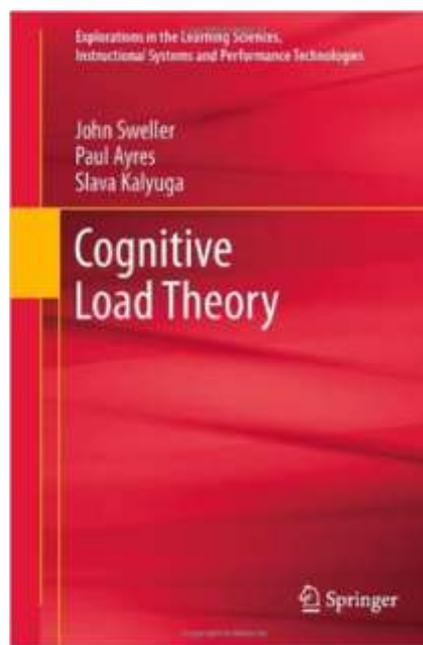
colleagues, Putai Jin and Renae Low, recently generalized the variability effect to learning to listen in a foreign language (Gao, Low, Jin, & Sweller, 2013).

Measuring Cognitive Load. Jeroen and Fred's most important work at this time was in devising a subjective rating scale to measure cognitive load (Paas, 1992; Paas & van Merriënboer, 1993). Subsequently, the measurement of cognitive load became an important sub-field in its own right with Detlev Leutner and Roland Brünken in Germany along with Jan Plass in New York contributing heavily to the field and in the process, transforming it (Brunken, Plass, & Leutner, 2004).

Guidance Fading Effect. Alexander Renkl from Germany carried out ground-breaking research. He became the world's leading expert on the worked example effect. His work on the guidance fading effect according to which learning is facilitated by gradually fading the assistance given to learners as they gain expertise has been critical (Renkl & Atkinson, 2003).

Domain-Specific vs. Generic Skills. In France, André Tricot advanced theoretical work on cognitive load theory and introduced it to the French-speaking world, along with Lucile Chanquoy (Chanquoy, Tricot, & Sweller, 2007). André and I collaborated on a paper concerned with the relative merits of emphasizing domain-specific as opposed to generic skills (Tricot & Sweller, 2014).

When it came to academic organizational flair, the Europeans, especially the Dutch, made me look like a rank amateur. They organized symposia on cognitive load theory, first at European conferences and then in the US and organized special issues of both American and European journals on the theory. Those special issues had a far greater impact on the field than individual papers. Alexander Renkl made contact with researchers at Carnegie-Mellon and that contact had more influence in acquainting Americans with the worked example effect than I had ever managed.



More Recent Cognitive Load Effects

The theory continued to generate new instructional effects, that will be described briefly below.

Isolated Elements Effect. The isolated elements effect, according to which interacting elements of very complex, high element interactivity information that are initially presented as isolated elements improve learning compared to having them presented in their “natural”, integrated format was demonstrated by Edwina Pollock, one of my PhD students (Pollock, Chandler, & Sweller, 2002). Paul Ayres devised important experiments on the isolated elements effect (Ayres, 2006). Paul Blayney, an accountancy academic at Sydney University who completed a PhD with me, demonstrated many of the conditions required for the isolated elements effect by relating it to the element interactivity and expertise reversal effects (Blayney, Kalyuga, & Sweller, 2010).

Imagination Effect. The imagination effect occurs when students asked to imagine concepts or procedures learn better than students only asked to study the same materials. Graham Cooper, Sharon Tindall-Ford, Paul Chandler and I carried out the initial work on this effect (Cooper, Tindall-Ford, Chandler, & Sweller, 2001). Subsequently, Paul Ginns, currently an

academic at Sydney University, further developed this work during his PhD with me (Ginns, Chandler, & Sweller, 2003), as did Wayne Leahy (Leahy & Sweller, 2005). Paul Ginns also was instrumental in publishing several meta-analyses of individual cognitive load effects (e.g. Ginns, 2005).

Collective Working Memory Effect.

Femke Kirschner and Fred Paas introduced the collective working memory effect to deal with collaborative learning (Kirschner, F., Paas, F., & Kirschner, P., 2009). When learners with different knowledge bases collaborate on a task, each individual, of course, has a limited working memory but by collaborating, in effect they are pooling their working memories. Providing the costs of collaborating are less than the effective increase in working memory due to pooling, performance should be increased compared to individual learning. Endah Retnowati, a PhD student of Paul Ayres and mine from Indonesia followed up on that work (Retnowati, Ayres, & Sweller, 2010).

Evolutionary Educational Psychology and Cognitive Load Theory

While the ultimate aim of cognitive load theory is to provide instructional effects leading to instructional recommendations, the theory itself continued to develop leading to new effects and controversies in the first 10 years of this century. Evolutionary psychology was beginning to gain a degree of prominence and some aspects of it were relevant to both the theoretical base of cognitive load theory and instructional design.

It became apparent that the information processing orientation used by cognitive load theory was analogous to the information processes that formed the base of evolution by natural selection. The suggestion that human cognition and evolution by natural selection were analogous was not new, with the ancestry of the analogy traceable back to Darwin. There seemed an obvious correspondence between: 1. information held in DNA and in long-term memory; 2. the transmission of

information during reproduction and the transmission of information between communicating humans; and 3. random mutation and random generate and test during problem solving. On the other hand, working memory, so central to human cognition, did not seem to have an obvious analogous function or process in evolutionary biology.

My wife, Susan who is a biologist, provided the necessary link (Sweller & Sweller, 2006). The epigenetic system plays the same role in evolutionary biology as working memory plays in human cognition. The epigenetic system uses the environment to determine biological structures and functions. For example, a person's skin and liver cells have exactly the same DNA in their nuclei but have vastly different structures and functions. Those differences are due to the epigenetic, not the genetic system. Both working memory and the epigenetic system act as a link between the information store (long-term memory and a genome) and the external environment. The analogy between biological evolution and human cognition proved to have instructional implications, discussed below.

There is a second, very direct role played by evolution by natural selection when determining instructional procedures. David Geary provided the relevant theoretical constructs (Geary, 2012). He described two categories of knowledge: biologically primary knowledge that we have evolved to acquire and so learn effortlessly and unconsciously and biologically secondary knowledge that we need for cultural reasons. Examples of primary knowledge are learning to listen and speak a first language while virtually everything learned in educational institutions provides an example of secondary knowledge. We invented schools in order to provide biologically secondary knowledge.

I had been vaguely aware of Geary's work but had not paid much attention to it because I assumed it was not relevant to my immediate research concerns. In 2006, Jerry Carlson and Joel Levin asked me to write a commentary (Sweller, 2007) on an article

written by David Geary (2007) to be published in an edited book. I agreed, which necessitated my reading his work with considerably more care and thought than I had hitherto managed. I was astonished. What Geary was proposing had the potential to change our field. It provided a resolution to issues that had seemed intractable to me.

Instructional Consequences of Evolutionary Educational Psychology

This is the context for the issues Geary dealt with. For many years our field had been faced with arguments along the following lines. Look at the ease with which people learn outside of class and the difficulty they have learning in class. They can accomplish objectively complex tasks such as learning to listen and speak, to recognize faces, or to interact with each other, with consummate ease. In contrast, look at how relatively difficult it is for students to learn to read and write, learn mathematics or learn any of the other subjects taught in class. The key, the argument went, was to make learning in class more similar to learning outside of class. If we made learning in class similar to learning outside of class, it would be just as natural and easy.

How might we model learning in class on learning outside of class? The argument was obvious. We should allow learners to discover knowledge for themselves without explicit teaching. We should not present information to learners – it was called “knowledge transmission” – because that is an unnatural, perhaps impossible, way of learning. We cannot transmit knowledge to learners because they have to construct it themselves. All we can do is organize the conditions that will facilitate knowledge construction and then leave it to students to construct their version of reality themselves. The argument was plausible and swept the education world.

The argument had one flaw. It was impossible to develop a body of empirical literature supporting it using properly constructed, randomized, controlled trials

altering one variable at a time. The worked example effect demonstrated clearly that showing learners how to do something was far better than having them work it out themselves. Of course, with the advantage of hindsight provided by Geary’s distinction between biologically primary and secondary knowledge, it is obvious where the problem lies. The difference in ease of learning between class-based and non-class-based topics had nothing to do with differences in how they were taught and everything to do with differences in the nature of the topics.

If class-based topics really could be learned as easily as non-class-based topics, we would never have bothered including them in a curriculum since they would be



learned perfectly well without ever being mentioned in educational institutions. If children are not explicitly taught to read and write in school, most of them will not learn to read and write. In contrast, they will learn to listen and speak without ever going to school.

Explicit Instruction. Coinciding with these theoretical developments, Paul Kirschner, an American who had transformed himself into a Dutchman and whom I knew from meetings at cognitive load theory symposia, suggested we collaborate on writing a paper advocating the use of explicit instruction rather than the minimal guidance commonly promoted. We wrote some drafts of the paper and Paul suggested that we should send it to Dick Clark for advice before submitting it to a publisher. Dick made it clear he liked the paper very much and gave some excellent advice for further improvements. With more work it became clear that Dick’s advice was becoming too extensive for a mere acknowledgement and so he was

included as a co-author. Thus began an extensive collaboration between the three of us that continues to this day.

The Kirschner, Sweller, and Clark (2006) paper had an immediate impact, unlike the 10-20 year wait before my empirical papers were noticed. It was polarizing with many opinions either strongly positive or strongly negative. Some idea of the reactions can be found in the edited collection of Sig Tobias and Tom Duffy (Tobias & Duffy, 2009), a book that derived from the several symposia on the topic of Constructivism generated by the original paper. Whatever the long-term influence of this work, it is notable that the term “constructivism” seems to have largely disappeared from the current research literature. Whether that disappearance was due to our efforts or other factors, and whether some of the replacements are any better than constructivism, are debatable topics.

There is a strong confluence between the work on evolutionary educational psychology and the issues associated with explicit instruction and minimal guidance. Humans are amongst the very few species that provide and obtain extensive information from other members of the species. Providing and obtaining information is a biologically primary skill. We are very good at it. Given this biologically primary skill, the suggestion that we should not explicitly provide learners with information is bizarre. I hope that the minimal guidance movement is an aberration that does not return.

Domain-Specific Knowledge. There are other implications that flow from the evolutionary educational psychology base of cognitive load theory. In the last few decades there has been a considerable emphasis on the acquisition of generic, cognitive skills such as, in mathematics, general problem-solving skills rather than domain specific skills. An example of a domain-specific skill in mathematics might be learning that when faced with a problem such as, $a/b = c$, solve for a , one should multiply both sides by the denominator.

Learning this procedure is important to solving a limited class of problems but useless when solving unrelated problems. While generic cognitive skills are far more important than domain-specific skills, because of that very importance, they are likely to be biologically primary and so do not need to be taught.

... build a program of research studying one variable after another rather than floating from one area to another.

André Tricot and I suggested that at least one of the reasons for the success of cognitive load theory has been its emphasis on teaching domain-specific rather than generic skills (Tricot & Sweller, 2014). Our field’s emphasis on the teaching of generic cognitive skills may be entirely misplaced, explaining the lack of a body of evidence supporting the teaching of such skills.

While generic cognitive skills can be learned but not taught, they can be used in the acquisition of domain-specific skills that need to be both taught and learned (Paas & Sweller, 2012). Learners who know how to use a generic cognitive skill may not know that it can be usefully applied when dealing with particular, domain-specific content. Simply pointing out to students that they should use a generic, cognitive skill may be beneficial when learning domain-specific content. Recently, Amina Youssef-Shalala, a PhD student of Paul Ayres and mine, Carina Schubert, a visitor from Germany, and I found that telling students that in the absence of domain-specific knowledge indicating how certain problems can be solved, they should use a strategy of randomly generating moves (Youssef-Shalala, Ayres, Schubert, & Sweller, 2014). For this strategy, students are told to make as many moves as they can without any reference to the goal of the problem. That strategy can be effective on transfer problems. While students did not need to be taught how to randomly generate moves because it is a biologically primary skill, they did need to be told to use the strategy in the

specific, biologically secondary domains that they were studying.

Some Current Work

Currently, while technically retired, I am continuing to conduct research and supervise research students. As one example, I have commenced collaboration with Tzu-Chien Liu from Taiwan who is carrying out important work on cognitive load theory. Several of my students and collaborators are working on relations between the worked example, element interactivity and generation effects and on element interactivity and the testing effect. While the concept of element interactivity has been around for 20 years and is central to cognitive load theory, cognitive load theorists have largely ignored it. Its centrality needs to be emphasized. As indicated above, cognitive load theory only applies to complex information that is high in element interactivity. It is not a theory of everything and cognitive load effects should not be expected using low element interactivity information.



Lessons Learned

There are several general lessons (not generic cognitive skills!) that I have learned over an almost half century of research. The main one is that age-old lesson that applies to many facets of life: *if you are confident of your ideas, persist*. In the case of research, ignore negative editorial

decisions and astonishingly ignorant reviewers. I have had my fair share of both. But as a frequent reviewer, I learned long ago that my reviews are just as incompetent as the worst. We all try our best but we are attempting to judge work that we are clearly inept to judge. The competent judges of our work are several generations hence. We do not have sufficient knowledge to properly judge current work but, of course, despite our incompetence, there is no one else available. *It will be up to future generations to determine the usefulness of our efforts.*

Contrary to what we might expect of researchers allegedly devoted to new ideas and new knowledge, we are incredibly conservative. The closer our ideas are to the prevailing zeitgeist, the more acceptable they will be. Most research papers support the prevailing views, whatever those views might be. *Therefore, do not hesitate to advance ideas conflicting with the current zeitgeist.* They may be ignored for a while but, if they do have merit, they are very likely to be ultimately recognized.

It is also useful to *follow up your ideas systematically and build a program of research studying one variable after another rather than floating from one area to another.* That may take some stubbornness, especially since the merit of one's work may not become clear for decades, but such a program of research has a better chance of obtaining eventual recognition than an identical number of unrelated studies.

I have at times, advanced suggestions that many felt were outrageous. Some of those suggestions now seem to be considered self-evident by many in the field. How did I change people's views? I do not think I did. Rather, people retired or died to be replaced by younger people who did not have to carry the burden of their own long history. *Societal renewal and change over the generations is as much a part of the human condition as individual resistance to change*, at least in some societies.

References

- Summaries of all of the work discussed above until about 2010 can be found in Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.
- Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology, 20*, 287-298.
- Blayney, P., Kalyuga, S., & Sweller, J. (2010). Interactions between the isolated-interactive elements effect and levels of learner expertise: Experimental evidence from an accountancy class. *Instructional Science, 38*, 277-287.
- Bobis, J., Sweller, J., & Cooper, M. (1994). Demands imposed on primary-school students by geometric models. *Contemporary Educational Psychology, 19*, 108-117.
- Brunken, R., Plass, J. L., & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual-task methodology: Auditory load and modality effects. *Instructional Science, 32*, 115-132.
- Cerpa, N., Chandler, P., & Sweller, J. (1996). Some conditions under which integrated computer-based training software can facilitate learning. *Journal of Educational Computing Research, 15*, 345-367.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293-332.
- Chanquoy, L., Tricot, A. & Sweller, J. (2007). *La Charge Cognitive*. Paris, France: Armand Colin.
- Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology, 79*, 347-362.
- Cooper, G., Tindall-Ford, S., Chandler, P., & Sweller, J. (2001). Learning by imagining. *Journal of Experimental Psychology: Applied, 7*, 68-82.
- Diao, Y., & Sweller, J. (2007). Redundancy in foreign language reading instruction: Concurrent written and spoken presentations. *Learning and Instruction, 17*, 78-88.
- Gao, Y., Low, R., Jin, P., & Sweller, J. (2013). Effects of speaker variability on learning foreign-accented English for EFL learners. *Journal of Educational Psychology, 105*, 649-665.
- Geary, D. (2007). Educating the evolved mind: Conceptual foundations for an evolutionary educational psychology. In J. S. Carlson & J. R. Levin (Eds.), *Psychological perspectives on contemporary educational issues* (pp. 1-99). Greenwich, CT: Information Age Publishing.
- Geary, D. (2012). Evolutionary Educational Psychology. In K. Harris, S. Graham & T. Urdan (Eds.), *APA Educational Psychology Handbook* (Vol. 1, pp. 597-621). Washington, D.C.: American Psychological Association.
- Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction, 15*, 313-331.
- Ginns, P., Chandler, P., & Sweller, J. (2003). When imagining information is effective. *Contemporary Educational Psychology, 28*, 229-251.
- Jeung, H.-J., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology, 17*, 329-343.
- Kirschner, F., Paas, F., & Kirschner, P. A cognitive load approach to collaborative learning. United brains for complex tasks. *Educational Psychology Review, 21*, 31-42.
- Kirschner, P., Sweller, J., & Clark, R. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential and inquiry-based teaching. *Educational Psychologist, 41*, 75-86.
- Kyun, S., Kalyuga, S., & Sweller, J. (2013). The effect of worked examples when learning to write essays in English literature. *Journal of Experimental Education, 81*, 385-408. doi: 10.1080/00220973.2012.727884
- Leahy, W., & Sweller, J. (2005). Interactions among the imagination, expertise reversal, and element interactivity effects. *Journal of Experimental Psychology: Applied, 11*, 266-276.

- Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology, 25*, 943-951.
- Leslie, K., Low, R., Jin, P., & Sweller, J. (2012). Redundancy and expertise reversal effects when using educational technology to learn primary school science. *Educational Technology Research and Development, 60*, 1-13.
- Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*, 319-334.
- Moussa-Inaty, J., Ayres, P. & Sweller, J. (2012). Improving listening skills in English as a foreign language by reading rather than listening: A cognitive load perspective. *Applied Cognitive Psychology, 26*, 391-402.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.
- Ng, H., Kalyuga, S., & Sweller, J. (2013). Reducing transience during animation: A cognitive load perspective. *Educational Psychology, 33*, 755-772.
- Owen, E., & Sweller, J. (1985). What do students learn while solving mathematics problems? *Journal of Educational Psychology, 77*, 272-284.
- Owens, P., & Sweller, J. (2008). Cognitive load theory and music instruction. *Educational Psychology, 28*, 29-45.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*, 429-434.
- Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review, 24*, 27-45. doi: 10.1007/s10648-011-9179-2
- Paas, F., & van Merriënboer, J. (1993). The efficiency of instructional conditions: An approach to combine mental-effort and performance measures. *Human Factors, 35*, 737-743.
- Paas, F., & van Merriënboer, J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*, 122-133.
- Pachman, M., Sweller, J., & Kalyuga, S. (2013). Levels of knowledge and deliberate practice. *Journal of Experimental Psychology: Applied, 19*, 108-119.
- Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction, 12*, 61-86.
- Renkl, A., & Atkinson, R. (2003). Structuring the transition from example study to problem solving in cognitive skills acquisition: A cognitive load perspective. *Educational Psychologist, 38*, 15-22.
- Retnowati, E., Ayres, P. & Sweller, J. (2010). Worked example effects in individual and group work settings. *Educational Psychology, 30*, 349-367.
- Rourke, A., & Sweller, J. (2009). The worked-example effect using ill-defined problems: Learning to recognise designers' styles. *Learning and Instruction, 19*, 185-199.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257-285.
- Sweller, J. (2007). Evolutionary biology and educational psychology. In J. S. Carlson & J. R. Levin (Eds.), *Psychological perspectives on contemporary educational issues* (pp. 165-175). Greenwich, CT: Information Age Publishing.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.
- Sweller, J., & Levine, M. (1982). Effects of goal specificity on means-ends analysis and learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 463-474.
- Sweller, J., Mawer, R. F., & Howe, W. (1982). Consequences of history-cued and means-end strategies in problem solving. *American Journal of Psychology, 95*, 455-483.
- Sweller, J., Mawer, R. F., & Ward, M. R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General, 112*, 639-661.

- Sweller, J., & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology*, *4*, 434-458.
- Tarmizi, R. A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, *80*, 424-436.
- Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, *3*, 257-287.
- Tobias, S., & Duffy, T. E. (Eds.). (2009). *Constructivist instruction: Success or failure?* New York: Routledge.
- Torcasio, S., & Sweller, J. (2010). The use of illustrations when learning to read: A cognitive load theory approach. *Applied Cognitive Psychology*, *24*, 659-672.
- Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review*, *26*, 265-283. doi: 10.1007/s10648-013-9243-1
- Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology*, *91*, 334-341.
- Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, *7*, 1-39.
- Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning & Instruction*, *22*, 449-457. doi: 10.1016/j.learninstruc.2012.05.004
- Yeung, A. S., Jin, P., & Sweller, J. (1998). Cognitive load and learner expertise: Split-attention and redundancy effects in reading with explanatory notes. *Contemporary Educational Psychology*, *23*, 1-21.
- Youssef-Shalala, A., Ayres, P., Schubert, C., & Sweller, J. (2014). Using a general problem-solving strategy to promote transfer. *Journal of Experimental Psychology: Applied*, *20*, 215-231.



About Acquired Wisdom

This collection began with an invitation to one of the editors, Sigmund Tobias, from Norman Shapiro a former colleague at the City College of New York (CCNY). Shapiro invited retired CCNY faculty members to prepare manuscripts describing what they learned during their College careers that could be of value to new appointees and former colleagues. It seemed to us that a project describing the experiences of internationally known and distinguished researchers in Educational Psychology and Educational Research would be of benefit to many colleagues, especially younger ones entering those disciplines. We decided to include senior scholars in the fields of adult learning and training because, although often neglected by educational researchers, their work is quite relevant to our fields and graduate students could find productive and gainful positions in that area.

Junior faculty and grad students in Educational Psychology, Educational Research, and related disciplines, could learn much from the experiences of senior researchers. Doctoral students are exposed to courses or seminars about history of the discipline as well as the field's overarching purposes and its important contributors. .

A second audience for this project include the practitioners and researchers in disciplines represented by the chapter authors. This audience could learn from the experiences of eminent researchers—how their experiences shaped their work, and what they see as their major contributions—and readers might relate their own work to that of the scholars. The first issue, prepared by Tobias as a sample chapter, was intended for illustrative purposes. Authors were advised that they were free to organize their chapters as they saw fit, provided that their manuscripts contained these elements: 1) their perceived major contributions to the discipline, 2) major lessons learned during their careers, 3) their opinions about the personal and 4) situational factors (institutions and other affiliations, colleagues, advisors, and advisees) that stimulated their significant work.

We hope that the contributions of distinguished researchers receive the wide readership they deserve and serves as a resource to the future practitioners and researchers in these fields.



Acquired Wisdom Series

Edited by

Sigmund Tobias

J. Dexter Fletcher

David C. Berliner

University at Albany

Institute for Defense Analyses

Arizona State University

State University of New York

Alexandria VA

Tempe AZ

Advisory Board Members

Gustavo Fischman, Arizona State University

Arthur C. Graesser III, Memphis State University

Teresa I. McCarty, University of California Los Angeles

Kevin Welner, Colorado State University

 *Education Review/ Reseñas Educativas/ Resenhas Educativas* is supported by the edXchange initiative's Scholarly Communications Group at the Mary Lou Fulton Teachers College, Arizona State University. Copyright is retained by the first or sole author, who grants right of first publication to the *Education Review*. Readers are free to copy, display, and distribute this article, as long as the work is attributed to the author(s) and ***Education Review***, it is distributed for non-commercial purposes only, and no alteration or transformation is made in the work. More details of this Creative Commons license are available at <http://creativecommons.org/licenses/by-nc-sa/3.0/>. All other uses must be approved by the author(s) or ***Education Review***. ***Education Review*** is published by the Scholarly Communications Group of the Mary Lou Fulton Teachers College, Arizona State University.

Please contribute reviews at <http://www.edrev.info/contribute.html>.

Connect with *Education Review* on Facebook (<https://www.facebook.com/pages/Education-Review/178358222192644>) and on Twitter @EducReview