Expertise Reversal Effect and Its Implications for Learner-Tailored Instruction

Article *in* Educational Psychology Review · October 2007

DOI: 10.1007/s10648-007-9054-3

Slava Kalyuga
UNSW Sydney

Educ Psychol Rev
DOI 10.1007/s10648-007-9054-3

REVIEW ARTICLE

Expertise Reversal Effect and Its Implications for Learner-Tailored Instruction

Slava Kalyuga

Abstract The interactions between levels of learner prior knowledge and effectiveness of different instructional techniques and procedures have been intensively investigated within a cognitive load framework since mid-90s. This line of research has become known as the expertise reversal effect. Apart from their cognitive load theory-based prediction and explanation, patterns of empirical findings on the effect fit well those in studies of Aptitude Treatment Interactions (ATI) that were originally initiated in mid-60s. This paper reviews recent empirical findings associated with the expertise reversal effect, their interpretation within cognitive load theory, relations to ATI studies, implications for the design of learner-tailored instructional systems, and some recent experimental attempts of implementing these findings into realistic adaptive learning environments.

Keywords Expertise reversal effect · Prior knowledge · Expertise · Cognitive load theory · Learner-tailored instruction

Although advantages of individualized learner-tailored instruction have been recognized for long time and continue to be aspired (e.g., see VanLehn *et al.* 2007 for the most recent manifestation) it still remains a mainly unrealized dream for the majority of educators. Most instructional materials are designed in a fixed, one-for-all fashion, and by default, implicitly if not explicitly, assume novices as intended learners. Unavailability of suitable real-time (online) diagnostic assessment techniques has also impeded the development of learner-tailored environments. Because of the involvement of many complex factors, issues of managing cognitive load by adapting instructions to individual learners, although recognized as important, have been mostly avoided by recent research projects in the field. On the other side, specific developmental projects in adaptive e-learning have been focused mostly on technical issues of tailoring instructional content to learner preferences, interests, choices, history of previous on-line behavior etc. and not based on fundamental cognitive characteristics of learners and evidence-based principles of instructional design.

S. Kalyuga (✉)
School of Education, The University of New South Wales, Sydney, NSW 2052, Australia
e-mail: s.kalyuga@unsw.edu.au

Studies of expert–novice differences in recent decades have clearly demonstrated that learner knowledge base is a single most important cognitive characteristic that influences learning and performance. Recent studies within the cognitive load framework, as well as some older results of research in Aptitude Treatment Interactions (ATI), have demonstrated that designs and techniques that are effective with low-knowledge individuals can lose their effectiveness and even have negative consequences for more proficient learners (see Kalyuga *et al.* 2003; Kalyuga 2005, 2006b; Tobias 1976, 1989 for previous overviews). The reversal in the relative effectiveness of instructional methods as levels of learner knowledge in a domain change has been referred as an expertise reversal effect. The major instructional design implication of these studies is the need to adjust instructional methods and procedures as learners acquire more expertise in a specific domain.

The following sections of the paper will review major characteristics of our cognitive architecture that underlie the expertise reversal effect, main empirical findings associated with the effect, their interpretation within cognitive load theory, relations to ATI studies, implications for the design of learner-tailored instructional systems, and some recent studies in implementing these findings into realistic adaptive learning environments. Some directions for future studies are identified in the conclusion.

Main Features of our Cognitive Architecture

Current theoretical model of human cognition within the cognitive load perspective includes several major characteristics that could be associated with wider principles that may govern all natural information processing systems (see Sweller 2003, 2004; 2007; van Merriënboer and Sweller 2005, for recent reviews of cognitive load theory, descriptions of the above principles, and their implications for the design of instruction). Firstly, our cognitive architecture (a general cognitive system that underlies human performance and learning) is essentially a knowledge-based one. Its operation is founded on a large store of organized knowledge structures in long-term memory with effectively unlimited capacity and duration. Such organized generic knowledge structures (or schemas) are used for mentally categorizing and representing concepts and procedures, and governing our behavior. They effectively determine our capabilities to function successfully in complex environments.

Secondly, our cognitive architecture has a mechanism that limits the scope of immediate changes to the knowledge base. This mechanism is usually associated with the concept of working memory as a conscious processor of information within our focus of attention that is responsible for constructing and updating our mental representations. It is severely limited in capacity and duration when dealing with novel elements of information (Baddeley 1986; Cowan 2001; Miller 1956). Processing limitations of working memory and associated cognitive load represent a major factor influencing the effectiveness of instruction.

Thirdly, our cognitive architecture can make sense of complex situations and coordinate different cognitive activities in conditions of severe working memory limitations. Knowledge structures held in long-term memory allow us to effectively reduce limitations of our cognitive system by encapsulating many elements of information into larger, higher-level units that could be treated as elements in working memory. Similar cognitive load reduction effects could also be achieved by practicing skills until they can operate under automatic rather than controlled processing (Kotovsky *et al.* 1985; Shiffrin and Schneider 1977). A cognitive architecture with limited processing resources would operate most efficiently when basic lower-level mental processes occur automatically preventing the system from an overload by processing demands and leaving cognitive resources for more sophisticated higher-level mental operations involved in thinking, learning, and problem solving. Intensive training on certain routine procedural elements of a task can make them more automatic and free cognitive capacity for enabling more creative cognitive processes of applying knowledge in unfamiliar situations or transferring of training (Cooper and Sweller 1987).

The characteristics of learning and performance alter significantly with the development of learner expertise in a domain. In the absence of relevant knowledge, novices are dealing with many new elements of information that may easily overload working memory. These learners require considerable external support to build new knowledge structures in a relatively efficient manner. In contrast, experts may rely on their available long-term memory knowledge structures for handling situations and tasks within their area of expertise.

High-level professional expertise requires years of extensive learning and practice in a specific domain (Ericsson *et al.* 1993) and involves many essential attributes in addition to the relevant knowledge base. However, one of the most important characteristics of expertise in any domain is the availability of a large number of domain-specific organized knowledge structures (schemas). High-level professional experts are in most cases also experts in solving specific routine tasks in their domains. Task-specific expertise is the ability to perform fluently in a specific class of tasks. A typical indicator of such expertise is performing rapidly advanced stages of solution by skipping some (or all) intermediate steps. Developing task-specific expertise is an important and necessary prerequisite for becoming a higher-level expert in a broader domain. According to this 'narrow' definition of expertise, even preschool children could be experts in solving some specific classes of tasks, for example, paper-folding tasks. The "expert" knowledge base in this case consists of structures and procedures used in performing this specific class of tasks. In this paper, such a 'narrow' view of expertise is used when describing occurrences of the expertise reversal effect in specific task domains.

Executive Function of the Knowledge Base

Within a cognitive architecture that is based on interacting working and long-term memory components, the available knowledge base in long-term memory represents a natural source of internal guidance for cognitive activities. It is assumed that knowledge structures in long-term memory perform an organizing and governing (or executive) role in complex cognitive processes (Sweller 2003). Appropriate knowledge structures are activated, retrieved from long-term memory, and combined to perform a function of managing specific incoming information streams (Kalyuga and Sweller 2005). When performing a complex cognitive task, we construct and continuously update a cognitive representation for the task situation based on our prior schemas for the task and incoming information. This situation model appropriately directs our attention and governs our performance in real time. In the absence of relevant knowledge, we would use random search processes by trying to fit different performance patterns or operations in trial-and-error attempts to handle the task. Alternatively, direct instructional guidance can perform an executive role by providing a partial substitute for the missing knowledge-based executive function for novices by telling them exactly how to handle the situation.

The concept of long-term working memory (Ericsson and Kintsch 1995) that includes structures created by long-term memory schemas associated with currently active components of working memory, may provide a specific mechanism for the executive role of our knowledge base. Thus, during performance of knowledge-based tasks, organized knowledge structures in long-term memory effectively determine the content and characteristics of working memory and govern complex cognitive activities. It is possible to determine the content of an individual's long-term working memory in a specific task situation by analyzing the content of concurrent (think-aloud) verbal reports. An alternative rapid diagnostic approach will be considered later.

In learning, the executive role of long-term memory knowledge is essential for providing cognitively efficient guidance for the construction of new knowledge structures in working memory and their integration (encoding) into available knowledge base in long-term memory. When no specific knowledge suitable for a task is available, we approach the task using mostly random search processes followed by tests of their effectiveness (Newell and Simon 1972). Such search processes require considerable resources of limited working memory and often result in cognitive overload and slow or negligible learning (Sweller 1988).

Our knowledge-based cognitive architecture may tend to minimize cognitive resources involved in performance ('cognitive economy principle') by using available knowledge structures as a more resource-efficient and, therefore, preferable means for governing cognitive activities than relying on alternative search procedures. As a by-product of such a generally efficient cognitive system, the 'cognitive economy' trend may sometimes result in selecting wrong, although well-entrenched, knowledge components for the executive role (for example, using simplistic folk beliefs in place of available scientific knowledge). They would be more 'economical' than extending significant mental resources on searching, reasoning, or trying to accommodate more comprehensive knowledge structures.

The relative share of internal long-term memory structures and external guidance in a learner's executive function for a task depends on her/his level of task-specific expertise. For low-knowledge learners, externally provided guidance may be the only available source of this function. Unless external instructional support substitutes for missing knowledge base, these learners would need to resort to cognitively inefficient search strategies. For experts in the domain, all necessary knowledge structures could be available in long-term memory and there would be no need in an additional instructional support. At intermediate levels of expertise, these two sources of information may be complementary. In an ideal and well-balanced situation, an executive function is based on long-term memory knowledge when dealing with familiar components of incoming information, and on direct guidance when dealing with unfamiliar units of information.

Unbalanced Executive Function and the Expertise Reversal Effect

There could be two major reasons for an unbalanced executive function. Firstly, in a situation where no guidance is provided for dealing with new units of information, learners have to apply general search strategies to cover the gap (for example, novice learners in an unguided discovery learning environment). If challenges of the task exceed the available knowledge structures, the task could cause a cognitive overload. Secondly, if external guidance is provided to learners who have sufficient knowledge base for dealing with the same units of information, learners would have to relate and reconcile the related components of available long-term memory base and externally provided guidance. Such integration processes may impose an additional working memory load and reduce resources available for learning new knowledge.

Presenting knowledgeable learners with detailed external guidance may hinder their learning and performance relative to the levels they could achieve with minimal instructional support. Therefore, as levels of learner expertise in a domain increase, relative effectiveness of learning tasks with different levels of instructional support may reverse. Instructional formats, techniques, and procedures that are optimal for novices may hinder relative performance of learners who are more experienced in a specific task domain by distracting them from fluent execution of appropriate cognitive processes and causing an expertise reversal effect.

Even though knowledge is the most critical characteristic of expertise that influence performance and learning processes, a non-optimal selection of learning tasks and associated levels of instructional guidance, as well as specific representational formats, may not allow experts to take a full advantage of their knowledge base. Mayer (1989) noted that the prior knowledge structures that a learner brings to the learning situation may trigger a cognitive conflict between this knowledge and presented information. Experts' well learned and refined cognitive models may conflict with models presented in instructional materials. Mental effort required to reconcile this conflict could cause a cognitive overload resulting in the expertise reversal effect.

Well ordered and balanced (optimized) executive function assumes that the learning task fits available knowledge-based executive structures and provides challenges just above the level of learner expertise. Unguided effortful search for solutions, as well as paying unnecessary attention to information that could otherwise be processed automatically and effortlessly, would reduce or prevent cognitive resources required for learning meaningful patterns of the task domain. Cognitive load associated with an unbalanced executive function may also de-motivate learners and thus further strengthen the effect (Paas *et al.* 2005).

The expertise reversal effect was predicted within the cognitive load theoretical framework as a form of redundancy effect that could occur when some presented information that was beneficial (and non-redundant) for novice learners became redundant for learners with higher levels of knowledge in a task domain (Kalyuga *et al.* 1998). For example, when related text and pictures are separated in space, their mental integration is expected to increase cognitive load. Physically integrating verbal and pictorial representations may reduce or eliminate this load (split-attention effect). However, for more advanced learners, eliminating non-essential redundant representations was expected to be more effective. For these learners, processing the redundant material may overload working memory relative to information without redundancy.

The effect was then extended to different presentation modalities and levels of instructional guidance, and became dissociated from the traditional redundancy effect (Chandler and Sweller 1991). In the expertise reversal effect, external information becomes redundant relative to a particular learner's internal knowledge structures. Additional cognitive resources are required for cross-referencing overlapping external and internal sources of information rather than only different external sources of information in the redundancy effect. Therefore, although the expertise reversal effect is a form of redundancy in a wider sense (when the learner knowledge base is included in the list of sources of information), it is not an example of a redundancy effect in the narrow sense.

Types of Cognitive Load Contributing to the Expertise Reversal Effect

Cognitive load is the demand for working memory resources of a specific person that are required for achieving goals of a particular cognitive activity or learning task when the individual is fully committed to the task. Actually invested resources obviously depend on motivation and other individual characteristics. Cognitive load always relates to cognitive processes of a specific person. Therefore, it depends not only on objective, depersonalized features of external information presentations or tasks, but also on cognitive characteristics of the learner. For example, the complexity of a task (e.g., the level of interactivity between its elements) is always relative to the learner knowledge base that determines what the elements are in the first place. The subjective nature of cognitive load needs to be emphasized when classifying and describing its sources and categories, especially intrinsic cognitive load.

Intrinsic cognitive load is associated with establishing connections between related elements of information or tasks in working memory and integrating them into available knowledge base. This load is essential for learning and is caused by internal intellectual complexity of the task or material relative to the level of learner task-specific expertise. Intrinsic load is often explicitly or implicitly related only to characteristics of learning materials or tasks (as a number of information elements and their interactivity). However, as soon as we use the term 'cognitive', we refer to human processes and not only to external characteristics of materials. The material only-based definition does not make much cognitive sense without considering cognitive activities of a learner who actually identifies elements of information and establishes patterns of their interactions using effortful conscious processes in working memory. These cognitive activities signify the comprehension of a situation and result in modified or new knowledge structures in long-term memory. Because intrinsic cognitive load is essential for achieving specific learning goals (comprehending a situation, performing a task, constructing new higher-level knowledge structures, achieving flexibility of such structures sufficient for transfer to relatively new task situations, etc.), it is vital to provide all the necessary resources to accommodate this load without exceeding limits of working memory capacity.

In order for the essential cognitive load not to exceed the cognitive capacity of a learner, it needs to be appropriately managed. For example, the initial learning goal could be divided into a series of sub-goals that require less processing resources, with instructional materials and tasks segmented or partitioned into smaller units. Alternatively, some of the essential interactions between elements of information could be excluded from consideration in order to artificially reduce structural complexity of the task on initial stages of learning followed by the fully interactive materials later (an isolated-interactive elements effect; Pollock *et al.* 2002). On the other hand, the management procedures may also involve increasing essential cognitive processing if it is at low levels and much cognitive capacity remains unused, for example, setting more challenging learning goals that require more complex cognitive activities with higher levels of element interactivity. It may also involve preventing uncontrolled reduction of essential cognitive processing, for example, when learners attempt to spontaneously rely on available simplistic knowledge structures (e.g., folk beliefs, scientific misconceptions) in guiding their cognitive activities.

An intentional increase in essential cognitive processing is referred to as an increase in germane cognitive load. It is sometimes difficult to separate intrinsic and germane types of load, since both of them represent useful forms of cognitive load. The concept of germane load was introduced to distinguish useful, learning-relevant demands on working memory from irrelevant and wasteful forms of cognitive processing (Sweller *et al.* 1998). Although intrinsic load is the most important part of learning-relevant cognitive demands, traditionally germane load has been associated with various additional cognitive activities designed to further foster learning or increase levels of learner motivation (e.g., explicitly self-explaining solution steps, imagining procedures described in worked examples, or varying situational features of learning tasks).

In contrast to essential, extraneous (unproductive, irrelevant) cognitive load is associated with a diversion of cognitive resources on activities irrelevant to learning goals because of design-related factors, such as a poor presentation design, inappropriate selection and sequencing of learning tasks, or inadequate instructional support. The expertise reversal effect is associated with two types of situations that cause extraneous cognitive load: 1) insufficient external guidance that does not compensate for limited knowledge base and forces novice learners to search for answers using cognitively inefficient procedures; 2) expert learner knowledge base overlaps with provided external guidance thus forcing learners to waste limited resources on co-referring internal and external representations of the same information. Both these forms of extraneous cognitive load can leave inadequate resources to sustain essential processing. It should be noted that the difference between extraneous and intrinsic cognitive load is relative to levels of learner expertise: some components of cognitive load that are essential for novice learners could become extraneous (irrelevant) for relatively more experienced learners, and vice versa.

It is also possible that the level of intrinsic load that is acceptable for more knowledgeable learners could be overwhelming for novices and exceed their capacity limits. This excessive intrinsic load would cause the disruption of learning processes and effectively become a form of extraneous load. Some previously mentioned techniques and procedures were developed for managing the exceeding levels of intrinsic load for novices. However, such techniques and procedures could become redundant for experts who would unnecessarily divert their resources on performing the required activities. Similarly, instructional methods for enhancing levels of germane load may produce cognitive overload for less experiences learners, thus effectively converting germane load for experts into extraneous load for novice learners. Such situations could also be associated with the expertise reversal effect.

The next sections describe empirical findings related to the expertise reversal effect. Mayer (2001) suggested computing effect size differences by subtracting the effect size for high-knowledge learners from the effect size for low-knowledge learners. Table 1 provides the effect sizes for less and more knowledgeable learners (novices and experts), numbers of participants in each category, and effect size differences for the reviewed studies. In each study, the compared instructional methods are ordered so that effect sizes for novices are positive. Then negative values for experts would indicate actual reversal of the effects (disordinal interactions). In some cases, even though no actual reversals were obtained, there were still notable differences in the magnitudes of the effects.

When sufficient data was not available in a source paper, effect sizes were computed using higher standard deviation values, thus providing conservative estimates. In two papers that did not contain sufficient data for making such estimates, textual description of the results was provided. Most of the reviewed studies used cross-sectional designs with different participants as low- and high-knowledge learners. For this reason, few available longitudinal studies are described in a separate section first.

Longitudinal Studies of Interactions Between Learner Levels of Expertise and Cognitive Load Effects

Kalyuga, Chandler, and Sweller conducted a series of longitudinal studies that were specifically designed to investigate interactions between different cognitive load effects and changing levels of learner expertise in controlled experimental conditions (a detailed overview could be found in Kalyuga 2006b). The general design of those studies included training the same samples of participants from novice to more expert states in specific task areas. Levels of performance and mental effort were measured at different stages along the way to see changes in relative effectiveness and efficiency of different instructional

Table 1 Expertise Reversal Effect: Summary of Results

Source	Experimental conditions	Effect size for novices (*N*)	Effect size for experts (*N*)	Effect size diff.
Longitudinal studies of interactions between learner levels of expertise and cognitive load effects				
Kalyuga *et al.* (1998)	Diagram with embedded text vs. diagram-only:	(17)	(30)	
	Operation and troubleshooting test	1.67	−0.44	2.11
	Fault-finding test	1.89	−0.88	2.77
Kalyuga *et al.* (2000)	Animated diagram with narrated text vs. diagram-only	1.18 (30)	−0.62 (38)	1.80
Kalyuga *et al.* (2001b)	Worked examples vs. problem-solving			
Experiment 1	Writing programs for relay circuits	0.90 (24)	0.23 (24)	0.67
Experiment 2	Writing switching equations for circuits	− 0.26 (24)	−0.75 (24)	0.49
Kalyuga *et al.* (2001a) Experiment 2	Worked examples vs. exploratory learning	0.69 (17)	−0.33 (17)	1.02
Expertise reversal for verbal and pictorial representation formats				
Mayer (2001), medians for a series of studies (1990–2000)	Text and illustrations vs text only			
	Retention questions	n/a	n/a	0.60
	Transfer questions	n/a	n/a	0.80
Lee *et al.* (2006)	Iconic vs symbolic representations in instructional simulations	1.60 (64)	−1.39 (63)	2.99
Yeung *et al.* (1998)	Text with integrated vocabulary vs. text with separate vocabulary			
Experiments 2–3	Primary school vs. university students	0.30 (48)	−0.53 (28)	0.83
Experiments 4–5	Low- vs. high-ability secondary school students	0.13 (56)	−0.53 (57)	0.66
Yeung (1999)	Text with integrated vocabulary vs. text with separate vocabulary	1.08 (84)	−0.79 (17)	1.87
Expertise reversal for instructional guidance and sequencing of learning tasks				
Tuovinen and Sweller (1999)	Worked examples vs. exploratory-based instruction	0.91 (17)	−0.31 (15)	1.22
Kalyuga and Sweller (2004) Experiment 3	Worked examples vs. problem-solving	1.20 (21)	−0.36 (21)	1.56
Reisslein *et al.* (2006)	Example-problem pairs vs. problem-example pairs	0.33 (62)	−0.37 (60)	0.70
	Example-problem pairs vs. faded worked examples	0.28 (63)	−0.26 (61)	0.54
Reisslein (2005)	Slow vs. fast transitioning to problem solving	1.30 (39)	−0.47 (44)	1.77
	Slow vs. immediate transitioning	0.51 (38)	−0.42 (45)	0.93
Seufert (2003)	Directive help for coherence formation vs. no help	1.32 (17)	−0.15 (16)	1.47
	Non-directive help for coherence formation vs. no help	0.95 (15)	−0.55 (19)	1.50
Lambiotte and Dansereau (1992)	Visual aids to lectures:	(37)	(37)	
	Knowledge map vs. list of terms			
	Recall of central ideas	0.61	−0.85	1.46
	Recall of details	0.27	−0.41	0.68
	Outline vs. list of terms			
	Recall of central ideas	0.29	−1.05	1.34
	Recall of details	0.19	−0.53	0.72
Pollock *et al.* (2002)	Isolated-interacting elements vs. interactive-only elements instruction			

Table 1 (continued)

Source	Experimental conditions	Effect size for novices (*N*)	Effect size for experts (*N*)	Effect size diff.
Experiments 1–2		(22)	(25)	
	Low-element interactivity test	0.63	0.06	0.57
	High-element interactivity test	0.86	0.11	0.75
	Test of the practical application	0.71	−0.16	0.87
Experiments 3–4		(18)	(18)	
	Low-element interactivity test	0.64	−0.20	0.84
	High-element interactivity test	1.13	−0.56	1.69
Clarke *et al.* (2005)	Sequential vs. concurrent learning of technology skills and subject domain knowledge	1.47 (8)	−0.30 (12)	1.77
Expertise reversal for dynamic visual representations				
Ollerenshaw *et al.* (1997)	Text with computer simulation labeling parts and operating stages vs. text-only	0.62 (23)	0.00 (18)	0.62
	Text with computer simulation labeling parts and operating stages vs. text with diagram labeling parts	0.92 (23)	0.41 (17)	0.51
Kalyuga (2007)	Static vs. animated diagrams (worked examples)	0.82 (17)	−0.46 (16)	1.28
Schnotz and Rasch (2005)	Static vs. animated pictures:	(20)	(20)	
Study 1	Study time	0.58	−0.58	1.16
	Circumnavigating posttest questions	1.07	0.00	1.07
Study 2	Simulation vs. manipulation pictures:	(13)	(13)	1.60
	Time-difference posttest questions	0.38	−1.22	
Expertise reversal for instructional hypertext and hypermedia				
Shin *et al.* (1994)	Limited-access hypertext (with restricted navigation) vs. full access hypertext	0.53 (52)	−0.13 (58)	0.66
Calisir and Gurel (2003)	Hierarchical hypertext vs. linear text	1.66 (10)	−0.17 (10)	1.83
	Mixed hypertext vs. linear text	1.01 (10)	−0.90 (10)	1.91
Potelle and Rouet (2003)	Content recall questions:			
	Structured hierarchical maps vs. semantic network maps	0.84 (18)	−0.73 (15)	1.57
	Alphabetical lists of topics vs. semantic network maps	0.60 (16)	−0.53 (15)	1.13
Shapiro (1999)	Text with structuring interactive overviews vs. text without structuring interactive overviews	1.17 (46)	0.38 (46)	0.79
Expertise reversal for methods of enhancing germane cognitive load				
Cooper *et al.* (2001)	Studying vs. imagining worked-out procedures			
Experiments 2–3		0.99 (22)	−1.24 (20)	2.23
Experiment 4		0.91 (36)	−0.75 (36)	1.66
Ginns *et al.* (2003)	Studying vs. imagining worked examples	(20)	(26)	
Experiments 1–2	Retention scores	0.45	0.00	0.45
	Transfer scores	0.39	−1.14	1.53
Leahy and Sweller (2005)	Studying vs. imagining instructions:			
Experiment 1	Low element interactivity	0.38 (30)	−0.37 (30)	0.75
	High element interactivity	0.71 (30)	−0.80 (30)	1.51
Experiment 2	Low element interactivity	0.55 (30)	−0.45 (30)	1.00
	High element interactivity	0.97 (30)	−0.79 (30)	1.76

techniques at different levels of learner expertise. Intensive training sessions, identical for all participants, were conducted between the experimental stages to increase learner expertise in corresponding task domains (Fig. 1). The task areas were restricted to relatively narrow classes of tasks to allow noticeable increases in learner expertise within several weeks. On the other hand, they had to be sufficiently expandable to allow a gradual increase in task complexity levels from stage to stage.

When different sources of information that require mental integration for understanding are separated in space or time, the process of integration (including visual search-and-match or cross-referencing) may substantially increase the burden on working memory and inhibit learning. Physically integrated or embedded formats were demonstrated to be an effective alternative to "split-source" instructions (split-attention effect; e.g., Mayer and Gallini 1990; Sweller *et al.* 1990; Tarmizi and Sweller 1988). Since working memory is likely to include partially independent subsystems for processing visual and auditory information (Baddeley 1986; Robinson and Molina 2002), split-attention situations may also be managed by using different modalities. Integration of the verbal auditory and pictorial visual information may not overload working memory if its capacity is effectively expanded by using a dual-mode presentation (modality effect; e.g., Mayer 1997; Mousavi *et al.* 1995; Tindall-Ford *et al.* 1997). However, if sources of information are intelligible in isolation, elimination rather than integration of a redundant source is preferable (redundancy effect; e.g., Chandler and Sweller 1991; Mayer *et al.* 2001).

Whether information is redundant depends on the level of expertise of the learner: what is essential for novices could be redundant for more knowledgeable learners or for the same learners at later stages of instruction. As a consequence, integrated formats that are effective for novices could be ineffective for more expert learners. Similar to visual, auditory explanations may also become redundant when presented to more experienced learners. Kalyuga *et al.* (1998) demonstrated that the relation between the split-attention and redundancy effects reversed as learner gains more expertise. With novice learners, the split-attention effect was obtained: trainees learned best from textual explanations that were embedded into the electrical wiring diagrams. After extensive training in the domain, the effectiveness of the integrated diagram and text condition decreased while the effectiveness of the diagram alone condition increased. After additional intensive training, substantial

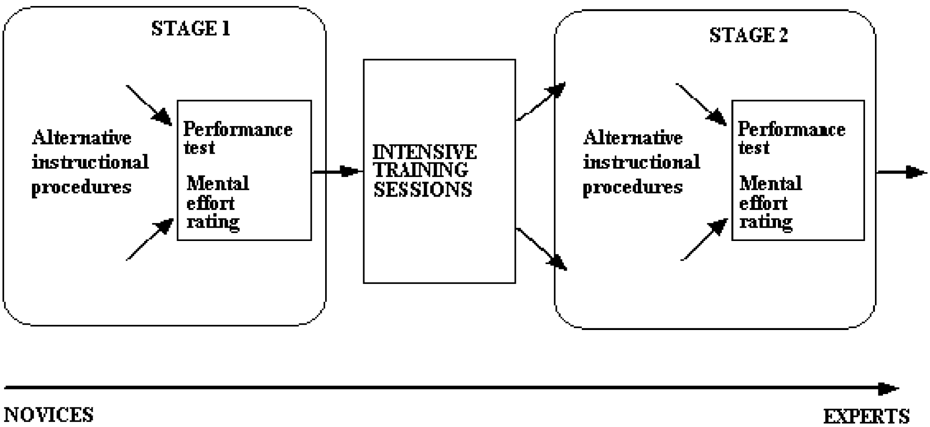

Fig. 1 Experimental sequence for studying interactions between levels of learner expertise and cognitive load effects

differences were observed between the conditions (Fig. 2). Diagram-alone materials were easier to process (according to subjective ratings of learning difficulty) and generated a higher level of performance on the subsequent tests. Textual explanations that were essential for novices became redundant for more knowledgeable learners.

In subsequent longitudinal studies (Kalyuga *et al.* 2000, 2001a, b), more evidence was obtained for an interaction between different instructional methods and levels of learner task-specific expertise. Patterns of results were similar to those shown in Fig. 2. The techniques for reducing extraneous cognitive load that were effective for novice learners (e.g., integrating sources of information or using dual-modality formats in split-attention situations, or using worked examples instead of conventional problem solving) became ineffective and often resulted in negative rather than positive or neutral effects for more knowledgeable learners.

For example, detailed narrated explanations of how to use a specific type of diagrams in mechanical engineering presented concurrently with on-screen animated diagrams that were effective for novices (modality effect), became redundant and reduced relative learning outcomes as learners became more knowledgeable in the task domain (Kalyuga *et al.* 2000). Explanations designed to support construction of knowledge that had already been acquired, needed additional cognitive resources for cross-referencing with available knowledge structures. The relative advantage of the narrated diagrams gradually disappeared while the diagram-alone condition became more effective. After several intensive training sessions, the diagram-only group outperformed the diagram with narrated text group, effectively reversing the results of the first stage. Subjective ratings of learning difficulty supported a cognitive load explanation of the results.

Kalyuga *et al.* (2001b) demonstrated that the superiority of worked examples in programming logic controllers over problem-solving practice (worked example effect; e.g., Cooper and Sweller 1987; Sweller 1988) disappeared as trainees acquired more experience in the task domain. In another experiment with tasks on writing switching equations for relay circuits, there was no difference between conditions initially (the trainees had some familiarity with the task domain), however, after intensive training in the domain, the learning of relatively more complex tasks (with greater numbers of elements in the circuits) was supported better by problem solving practice than by worked examples.

Kalyuga *et al.* (2001a, Experiment 2) compared worked examples with an exploratory-based instruction on writing switching equations for relay circuits. Learners designed different circuits first by using an interactive on-screen template and then wrote equations for those circuits. Although initially the worked examples group outperformed the explor-

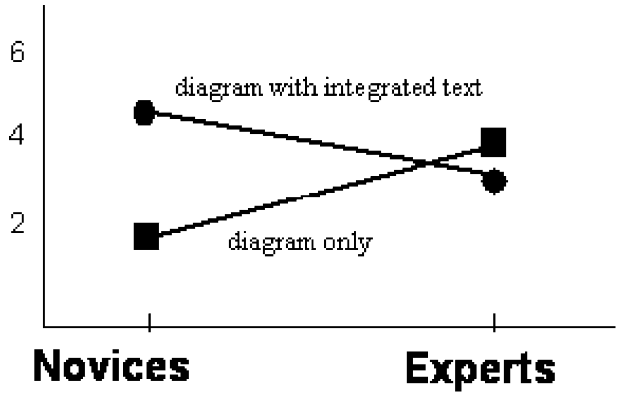

Fig. 2 An interaction between instructional formats and levels of learner expertise according to Kalyuga *et al.* (1998)

atory one, as the level of learner expertise increased after a series of intensive example- and problem-based training sessions, the exploratory group progressed better and eventually outperformed the worked examples group.

The integration of subjectively redundant (relative to the learner knowledge base) instructional support with available knowledge structures may cause an additional cognitive load and interfere with learning compared to minimally guided instruction that relies on learner pre-existing knowledge, especially if more knowledgeable learners cannot avoid processing the redundant components of information. Therefore, as levels of learner prior knowledge in a domain increase, relative effectiveness of different instructional methods may reverse. Methods optimal for low-knowledge learners may hinder learning performance of high-knowledge learners. Using essential and removing redundant information and procedures as learner gains more task-specific expertise, thus reducing or eliminating interfering cognitive processing, is important for optimizing cognitive resources.

Expertise Reversal for Verbal and Pictorial Representation Formats

R. Mayer and his associates conducted a series of studied in learning from text and graphics. The experiments indicated that using graphics usually enhanced learning outcomes for students with low prior knowledge levels, but not those with higher knowledge levels (e.g., Mayer and Gallini 1990; Mayer *et al.* 1995). In a review of those studies, the effect was called an individual differences principle and attributed to the ability of high-knowledge learners to use their knowledge base to compensate for missing instructional guidance (Mayer 2001). Advantages of pictorial representations disappeared with increases in learner levels of expertise.

Lee *et al.* (2006) investigated an interaction between two different modes of visual representations in a gas law simulation for middle-school chemistry students and different levels of learner prior science knowledge. Essential gas characteristics were presented either a symbolic form only (words 'temperature', 'pressure', and 'volume' with corresponding numerical values) or by adding iconic information to the symbolic representations (e.g., burners for temperature, weighs for pressure). The study indicated that whereas low prior knowledge learners benefited more from added iconic representations than from symbolic formats only, high prior knowledge learners benefited more from symbolic only representations than from added iconic ones. Iconic representations were redundant for these learners and interfered with their knowledge-based cognitive processes. It should be noted that the expertise reversal effect was observed only with materials that considered only two parameters at a time and, therefore, had manageable levels of intrinsic cognitive load. For high complexity materials that considered all three parameters concurrently and required excessive levels of intrinsic load for all participants, iconic representations were beneficial for both novices and experts.

A number of studies investigated relative effectiveness of different forms of textual representations. Based on the proportion of correct answers to three types of questions (problem solving, bridging inference, and elaborative inference questions) on the posttest for the maximally and minimally coherent texts as a function of readers' background knowledge, McNamara *et al.* (1996) found that adding additional explanations to an instructional science text in order to increase its coherence were beneficial only for low-knowledge readers. High-knowledge readers benefited more from using the minimally coherent text format. It should be noted that both low- and high-knowledge readers benefited from high-coherence text on reproductive text-based questions. (Since data provided in that paper is insufficient for calculating effect sizes, it is not included in Table 1). Although McNamara's *et al.* (1996) interpretation of their results was based on the experts' more active engagement in the processing of the minimal text format, this study could also be considered as an example of the expertise reversal effect. In similar situations involving extended vs. minimal explanations (e.g., Kalyuga *et al.* 1998; Yeung *et al.* 1998), experts reported lower estimates of cognitive load in the minimal instructional formats in comparison with formats that contained redundant instructional explanations.

Yeung *et al.* (1998) compared two instructional formats of incorporating definitions of unfamiliar words into textual material. One was a traditional glossary placed at the end of the whole text. Another format integrated unfamiliar words' definitions into the text directly above the defined word. According to cognitive load theory, traditional separate glossaries could produce a split-attention effect due to extra effort necessary for locating the required definition in the glossary, comprehending and remembering it while finding the way back to the original word in the text. Secondary school students (5th grade) learned better from the integrated definition format than from the traditional glossary format (as measured by comprehension scores). However, the university students demonstrated better comprehension with the separate glossary format. The increased cognitive load caused by the need for more knowledgeable learners to process redundant for them information was supported by measures of time students spent on referencing definitions. Thus, the use of integrated definitions could have a positive or negative effect on learning depending on levels of learner expertise.

In another set of experiments, Yeung *et al.* (1998; Experiments 4–5) compared the above two instructional formats with 8th grade students. However, in one experiment, low knowledge students (from remedial ESL classes) were involved in the study. In another experiment, higher knowledge-level students were selected. The same reversed pattern of results was obtained: low knowledge students benefited from the integrated instruction, while more experienced learners achieved better results from the traditional separate glossary format. In a replication study with secondary school and university students, Yeung (1999) again demonstrated that novices learned better from the integrated definition format than from the traditional one, and the university students achieved better comprehension scores with the separate glossary format.

Overall, the reviewed studies that demonstrated expertise-related reversals in effectiveness of different verbal and pictorial representation formats indicated that low prior knowledge learners benefited more from integrated multiple (verbal–verbal, verbal–pictorial, or symbolic–iconic) representations than from separated or single representation formats. On the other hand, more knowledgeable learners benefited more from minimal single representations. While integrated instructional formats provided well balanced guidance for novices, additional representations were redundant for more knowledgeable learners and interfered with their learning processes (see also a relevant study of Seufert 2003 described in the next section, for an expertise reversal effect for verbal semantic support when learning from text and pictures).

Expertise Reversal for Instructional Guidance and Sequencing of Learning Tasks

A number of studies investigated the learning effects of different levels of instructional guidance for learners with different levels of expertise in a domain. Tuovinen and Sweller (1999) compared well-guided worked examples with a minimally-guided exploratory-based instruction on how to use a database program. Novice students benefited more from worked examples, with no differences found between conditions for higher-knowledge students.

Kalyuga and Sweller (2004 Experiment 3) studied an interaction between levels of learner expertise in the task domain of calculating distances and projections in coordinate geometry and levels of instructional guidance. Participants (high school students) were divided into two groups of more and less knowledgeable learners based on a median split using scores obtained in a pretest (a rapid diagnostic method was used for evaluating levels of expertise; see a corresponding section below). Posttest results indicated that less knowledgeable high-school students benefited significantly more from well guided worked examples. For more knowledgeable learners, there was a clear indication of problem solving benefits. A significant interaction between knowledge levels and instructional formats demonstrated that the most efficient instructional format depended on the level of learner expertise. As this level increased, performance of the problem solving group improved more than performance of the worked examples group.

Reisslein *et al.* 2006) compared effectiveness of three different sequencing approaches to example-based instructional procedures in the area of serial and parallel electrical circuit analysis for learners (university engineering students) with different levels of prior knowledge in the domain. One approach used traditional example-problem pairs with worked examples followed by isomorphic practice problems. Another approach provided practice problems first with accompanying worked examples for reference if needed (problem-example pairing). The third condition included backward faded worked examples in which increasingly more steps at the end of the solution procedure were omitted. The study demonstrated that novices benefited more from example–problem pairs, whereas experts benefited more from problem–example pairs and faded examples sequences.

In her doctoral dissertation, Reisslein (2005) examined the effect of the pace of transitioning from worked examples to independent problem solving for learners with different levels of prior knowledge in electrical circuit analysis (engineering college freshmen). Under the immediate transitioning condition, learners started practicing problems immediately after the introduction. Under the fast fading condition, worked solution steps were backward faded at a rate of one step with each example. Under the slow fading condition, the rate was halved (one step for every second example). The results on retention posttest indicated significant interactions between levels of learner prior knowledge and the pace of transitioning. More knowledgeable learners performed significantly better in the fast and immediate transitioning groups than in the slow transitioning group, indicating that detailed guidance could be redundant for these learners. On the other hand, learners with low levels of prior knowledge benefited more from slow transitioning condition, thus demonstrating the importance of detailed guidance for novice learners.

Seufert (2003) studied the effect of verbal semantic help for coherence formation in mapping the structure of material when learning from scientific text and pictures (see also Seufert and Brünken 2006). Two kinds of assistance were investigated: directive help (specific direct support) and non-directive help (questions to students providing non-specific hints). Although three levels of learner prior knowledge (low, middle, and high) were considered in this study, the low-level students did not have sufficient prerequisite knowledge (basic concepts of the domain) for learning new material and were not able to use the provided help. Therefore, for the purpose of expert–novice comparisons in this situation, it is more appropriate to consider middle-level participants as novices in the specific task area. The analysis indicated that for these novice learners, both directive and non-directive help conditions were significantly better (with relatively more benefits from the direct help) than the no-help condition based on comprehension posttest results. On the other hand, for experts, the results reversed although there were no significant differences between conditions.

Lambiotte and Dansereau (1992) investigated the effects of three types of visual overhead aids to audiotaped lectures in college-level biology: knowledge maps, topic outlines, and lists of key terms. Tests of free recall of central ideas and details demonstrated that students with low prior knowledge in the domain benefited most from the knowledge maps and least from the lists of key terms. For more knowledgeable learners, the results reversed with the lists of key terms as the most effective learning aid (at both levels of prior knowledge, the effects were stronger for the recall of central ideas than details). Knowledge maps provided required instructional support for novices but were redundant for more knowledgeable learners.

In all of the above studies, novice learners benefited most from well guided low-paced instructional procedures that reduced extraneous cognitive load for these learners, especially when learning structurally complex materials. For more experienced learners, studying and integrating the externally provided detailed guidance with learners' available knowledge structures that provided essentially the same guidance could impose a greater cognitive load than minimally guided forms of instruction. These learners were able to use their relevant knowledge base to guide learning without overloading working memory. Thus, instructional guidance that is essential for novices may have negative consequences for experts by interfering with retrieval and application of their available knowledge structures, especially if these learners cannot ignore or otherwise avoid processing the redundant explanations.

In situations when intrinsic cognitive load exceeds cognitive capacity of novice learners, initially presenting complex material as a set of isolated elements of information that could be processed serially, rather than simultaneously, may eliminate the cognitive overload in working memory. Pollock *et al.* (2002) demonstrated that artificially simplified isolated-elements learning tasks followed by the fully interacting elements instruction benefited low-knowledge learners. However, there were no differences between this method and the traditional approach using complex materials during both stages for learners with higher levels of prior knowledge in the domain.

Clarke *et al.* (2005) investigated interactions between the timing of acquiring specific spreadsheet skills in learning mathematics from spreadsheet applications, and levels of learner expertise in using spreadsheets. The sequential experimental condition provided instructions on spreadsheets prior to applying this knowledge to learning mathematics. In the concurrent condition, instructions on using spreadsheets and mathematical concepts were presented in an integrated format. The results of the experiment indicated that students with lower knowledge of spreadsheets learned mathematics more effectively in the sequential formats in which the relevant spreadsheet skills were acquired prior to attending the mathematical tasks. On the other hand, students who were more experienced in using spreadsheets benefited more from an integrated format in which relatively new spreadsheet skills were acquired during learning corresponding mathematical concepts. Reversed measures of cognitive load (subjective ratings) supported the cognitive load interpretation of the effect.

If instructions on both spreadsheet applications and mathematics are presented concurrently for novices, their working memory may become overloaded and learning inhibited compared to a sequential presentation. Concurrent learning may only be effective for more technologically experienced learners who need to concentrate on learning the relationship between their technology skills and mathematics. Thus, acquiring basic technology skills while learning a specific subject discipline is unlikely to be effective for beginners, and the technology should be learned prior to learning a specific subject area (Clarke *et al.* 2005).

Expertise Reversal for Dynamic Visual Representations

A number of studies investigated interactions between levels of learner prior knowledge and the effectiveness of dynamic visual representations (simulations and animations) vs. traditional static representations. Ollerenshaw *et al.* (1997) observed that low prior knowledge students benefited more from a text with computer-based animated simulation of the pump's operation (with labeled parts and operating stages) than from text-only or text with static diagrams labeling parts. The limited static representations did not provide sufficient guidance to match that of simulated animations. When the same formats were used with high knowledge students, the beneficial effect of the animated format was substantially reduced (although not actually reversed). These learners used their knowledge base to compensate for limited external guidance. The level of provided instructional guidance seemed to be the main factor that caused differences between the investigated instructional formats. To determine how the representational dynamics influenced learning outcomes, levels of external instructional support were equalized in the rest of studies reviewed in this section.

The interaction between levels of learner expertise and effectiveness of animated and static procedural examples was studied by Kalyuga (2007) in the task domain of transforming graphs of linear and quadratic equations in mathematics (e.g., transforming a given graph of the line $y=x^2$ into a graph of the line $y=2(x-1)^2-3$). Participants (university students) were subdivided into groups of high- and low-prior knowledge learners based on results of a pretest. A half of students in each group studied two sequential animated instructional segments on how perform the transformations. Another half studied an equivalent set of static diagrams showing major transformation stages on one screen. The posttest results demonstrated that less knowledgeable learners performed significantly better after studying static examples. Learners with higher levels of prior knowledge showed better results after studying animated instructions. There was a significant interaction between knowledge and instructional formats: as levels of learner expertise increased, the performance of the animated instruction group improved more than performance of the static group.

Schnotz and Rasch (2005) compared effects of animated and static pictures about time phenomena related to the Earth rotation on learners with different levels of learning pre-requisites (a combination of pre-test scores of prior knowledge in the domain and intelligence measures). Two different animated pictures were investigated: a picture that displayed visual simulations of changes over time when circumnavigating the Earth (simulation picture) and a more interactive picture that allowed students to manipulate the display by defining specific day and time for specific cities (manipulation picture). The results of Study 1 that compared animated with static pictures indicated that high learning prerequisite learners spent more time on studying animated than static pictures, whereas low learning prerequisite students spent more time studying static than animated pictures. For circumnavigation posttest questions that required mental simulations, students with low learning prerequisites performed significantly better after learning with static pictures than with animated pictures, while high learning prerequisite students performed equally in both conditions. These students were able to perform mental simulations by themselves and the external support was redundant. The results of Study 2 that compared manipulation and simulation pictures indicated that for time-difference posttest questions, students with high learning prerequisites performed significantly better after learning from manipulation pictures than from simulation pictures, while lower learning prerequisite students performed better after learning from simulation pictures than from manipulation pictures.

According to cognitive load theory, continuous animations could be too cognitively demanding for novice learners because of high levels of transitivity. If static and dynamic visualizations are equivalent in terms of provided supportive information, novice learners would benefit more from studying a set of static diagrams. For more knowledgeable learners, available knowledge structures may help them to handle the transitivity of animated instructions. On the other hand, details displayed in static graphics may need to be integrated and reconciled with knowledge base of these learners imposing additional working memory demands. Similar reversal patterns could be expected with interactive manipulations in comparison with traditional non-interactive pictures: manipulation pictures could impose extraneous load on novice learners but be optimal for more experienced learners. When learners lack sufficient task-specific knowledge, it is important that the appropriate guidance (to serve in the executive role and prevent unproductive search activities) is provided without unnecessary cognitive overload.

Expertise Reversal for Instructional Hypertext and Hypermedia

In their recent overview of cognitive load issues in hypertext learning environments, DeStefano and LeFevre (2007) noted that the general assumption about the role of prior knowledge in learning from hypertext was that high prior knowledge learners could be able to process and make sense of unordered segments of text, and handle interruptions in reading by connecting these segments to existing knowledge structures without overloading working memory. In contrast, low prior knowledge learners may experience cognitive overload in hypertext environments that could inhibit learning outcomes. For example, Shin *et al.* (1994) compared high and low prior knowledge 2nd graders using either full access or limited access (with restricted navigation facilities) versions of hypertext. The results indicated that there was no difference between these two hypertext versions for students with higher levels of prior knowledge in the topic, whereas the more structured limited-access version of hypertext was more beneficial for low prior knowledge learners.

Calisir and Gurel (2003) investigated if linear text and two hypertext structures (hierarchical and mixed) would interact with learner (university students) prior knowledge in the domain of productivity management. While for non-knowledgeable learners, both hypertext structures were significantly better than linear text, results reversed for knowledgeable learners (although differences were not statistically significant). In the linear text condition, knowledgeable learners had higher reading comprehension scores than non-knowledgeable learners. Domain knowledge may have helped these learners to understand and conceptualize the structure of the text. There were no significant differences between knowledgeable and non-knowledgeable learners in the hierarchical and the mixed conditions. In this study, the well structured hierarchical and mixed forms of hypertext enhanced comprehension of less-knowledgeable learners by compensating for the lack of internal conceptual structure of the domain.

Potelle and Rouet (2003) compared three hypertext environments based on different levels of structural organization: a hierarchical map that provided the most explicit structure of the content; a semantic network map; and an alphabetical list of topics without explicit high-level relations. Low-knowledge readers learned global macrostructural concepts in the area of social psychology more from the hierarchical map than from other two structures, while there were no significant differences for high-knowledge readers (although there was a medium to large size effect in favour of the semantic network map on content recall questions). Hierarchical rather than semantic representations facilitated the construction of explicit global-level representations and the integration of individual topics for low-knowledge learners.

Shapiro (1999) studied the relationship between learner prior familiarity with elementary zoology and ecology and interactive overviews as advanced organizers for structuring presented textual descriptions in hypermedia-based learning environments. The study found that the external structuring by interactive overviews produced significant benefits for novices, however made little difference for learners with higher levels of prior familiarity with the domain for whom the explicit external structuring was not necessary.

To conclude this section, even though the expertise reversal was demonstrated with hypertext and hypermedia materials, specific manifestations of the effect depend on factors that require further studies. Cognitive load in learning from textual materials usually depends on the text structure relative to levels of learner expertise: low-knowledge learners benefit from well structured content representations, whereas for high-knowledge learners, different levels of structure may make little difference. When exploring a large number of navigational choices in complex hypertext and hypermedia environments, the cognitive resources required for such search processes (that essentially are extraneous to learning) may become unavailable for constructing relevant knowledge structures. Unsupported hypertext/hypermedia environments could be suitable for experienced learners with sufficient levels of prior knowledge that could guide these learners in their exploration of the environment.

Thus, while for experts in a domain, both hypertext/hypermedia and linear text environments may work well, for novices the answer may depend on the specific nature of the learning material. For example, if well structured hypertext explicitly represents the content better than traditional linear text (e.g., in well structured domains with limited number of links to follow), it still could be beneficial for novice learners; otherwise the linear text would remain a better option. More studies are needed to determine characteristics of hypertext/hypermedia learning materials in specific domains that make them suitable for appropriately balancing the executive function between knowledge-based and instruction-provided guidance.

Expertise Reversal Involving Germane Cognitive Load

Cooper *et al.* (2001) demonstrated that imagining procedures and concepts related to using spreadsheets produced better instructional outcomes than simple studying of worked examples for students who had appropriate knowledge base to construct and run corresponding mental representations. However, the imagining procedure produced a negative effect for low-knowledge students thus reversing the effect. When studying worked examples, novices construct knowledge structures for interacting elements. More knowledgeable learners already have such knowledge structures and studying worked examples is a redundant activity for these learners who may benefit more from additional practice provided by imagining corresponding procedures.

Ginns *et al.* (2003) conducted a similar study with the complexity of learning materials as an additional experimental factor. The study demonstrated that low prior knowledge university students learning structurally complex materials (HTML code) benefited more from studying worked examples than from imagining them. On the other hand, more knowledgeable students dealing with less complex materials (secondary school students studying geometry materials) reached higher levels of transfer test performance when imagining rather than studying examples.

These results were further supported by the study of Leahy and Sweller (2005) with primary school students learning to read a bus timetable (Experiment 1) or temperature graphs (Experiment 2). The same students were used initially as novices, and two weeks later as relative experts in a domain. In Experiment 1, the study condition was more effective for novice learners than the imagination condition, whereas a reversed pattern was observed for more experienced students. Experiment 2 allowed a greater spread between low and high element interactivity materials, and the above pattern of results was replicated with stronger effects.

If specific techniques for engaging learners into additional cognitive activities designed to enhance germane cognitive load (e.g., explicitly self-explaining or imagining content of worked examples) cause overall cognitive load to exceed learner working memory limitations, the germane load could effectively become a form of extraneous load and inhibit learning. This especially applies to novice learners who lack relevant schematic knowledge structures in long-term memory that could effectively increase working memory capacity due to chunking effect. More experienced learners may have sufficient resources for effectively accommodating additional germane load that would enhance their learning outcomes.

Expertise Reversal Effect and Aptitude–Treatment Interactions

The expertise reversal effect could be related to general studies of ATIs that were initiated more than 40 year ago by Cronbach (1967). The concept of aptitude was defined broadly as "a complex of personal characteristics that accounts for an individual's end state after a particular educational treatment, i.e., that determines what he learns, how much he learns, or how rapidly he learns" (p. 23). Relevant aptitudes include knowledge, skills, learning styles, personality characteristics, etc. ATIs occur when different instructional treatments result in differential learning rates and outcomes depending on student aptitudes (e.g., Cronbach and Snow 1977; Lohman 1986; Mayer *et al.* 1975; Shute and Gluck 1996; Snow 1989, 1994; Leutner 1993; Snow and Lohman 1984). Learner prior knowledge is the aptitude of interest in the context of the expertise reversal effect.

Tobias (1976) reviewed a series of studies that consistently demonstrated interactions between prior familiarity with a domain and instructional treatments. In unfamiliar domains, detailed and consistent instructional support (for example, appropriate sequencing of material according to instructional objectives, feedback, etc.) provided to learners in programmed learning environments produced better results than reading less structured materials, although no significant differences were found for familiar materials. Similar results were found in relation to other instructional methods used to improve text comprehension (e.g., reviewing the text, answering adjunct questions; Tobias 1987, 1988).

Thus, an inverse relationship existed between prior achievement in a domain and optimal instructional strategies: the higher the level of prior achievement, the lower the levels of instructional support and structure required for learners and vice versa. Even though the patterns of these results are fully compatible with the expertise reversal effect, it should be noted that in those studies, the instructional support was considered in a rather narrow sense as external attributes of instructional procedures, e.g., assistance in eliciting responses and providing feedback on their accuracy. In a later review, Tobias (1989) suggested that instructional support should be considered in deeper terms of assisting students in using required cognitive processes. This view of instructional support is essential for research in the expertise reversal effect.

It was indicated that prior achievement (or prior knowledge in the current terminology) as determined by detailed pretests scores was an important variable due to its relatively clear definition and meaning. Therefore, ATIs involving prior achievement were easier to investigate and produced more convincing results than other, more vaguely defined, aptitudes. Pre-training learners within a specific task domain was suggested as an appropriate experimental procedure for manipulating this variable (Tobias 1976). This suggestion was, in effect, realized in previously described longitudinal studies of the expertise reversal effect.

In regards to other than prior achievement categories of aptitudes, there was less consistent empirical support found for stable aptitude-treatment interactions (Bracht 1970). The suggested reasons included inadequate aptitude measures that were designed for selection rather than diagnostic purposes (e.g., batteries of aptitude tests based on artificially simplified tasks in laboratory conditions) and inability to apply such measures dynamically, thus ignoring learning and practice effects. In the ATI approach, differences in aptitudes were studied and instructional treatments selected without taking into account differences in associated cognitive processes. Psychometric tools used for measuring aptitudes were unsuitable for analyzing, evaluating, and facilitating ongoing cognitive processes involved in knowledge acquisition (Federico 1980). As a result, traditional ATI research had no significant influence on classroom instruction and was difficult to use for guiding development of practically useful learner-tailored instructional systems (Boutwell and Barton 1974; Federico 1999).

Tobias (1989) suggested that inconsistency and variability of ATIs findings, including those involving learners with different levels of domain-specific knowledge, were due to unsupported implicit assumptions of ATI studies regarding cognitive processes involved. He distinguished between macroprocesses as molar processes under learner control, and lower-level processes that were less controlled by the learners ("molecular" processes, e.g., manipulating information in short-term memory). Contrary to ATI basic assumptions, it appeared that different macroprocesses were neither directly associated with (or caused by) specific instructional methods (unless they were experimentally manipulated), nor necessarily correlated with learner characteristics (Tobias 1989).

Due to relatively recent studies of the role of our knowledge base in cognitive functioning, we know that domain-specific knowledge structures fundamentally influence learning and performance on the basic "molecular" level (which essentially is the level of working memory processes). This makes the prior knowledge base a single most important factor in studying interactions between learner characteristics and instructional methods and signifies a special place for the expertise reversal effect among such interactions.

In addition to these differences in the theoretical approach, the expertise reversal studies used different research methods. Typical ATI studies tested for differences in regression slopes between treatment groups, while studies in the expertise reversal effect usually contrast groups with different levels of expertise or follow learning in longitudinal procedures from novice to more expert states in a specific task domain. These differences in methods reflect previously noted conceptual differences in the expertise reversal research in comparison to the ATI tradition (e.g., considering working memory-based cognitive processing; relating instructional guidance for specific cognitive processes to available long-term memory knowledge; pre-training learners in specific cognitive processes to manipulate levels of their expertise in a domain).

Tailoring Learning Environments to Levels of Learner Expertise

A major instructional implication of the expertise reversal effect is the need to tailor dynamically instructional techniques, procedures, and levels of instructional guidance to current levels of learner expertise as they gradually change during learning. The idea of learner-tailored instruction was clearly articulated still within the original ATI approach. Cronbach (1967) and Glaser (1977) suggested that knowledge of aptitude-treatment interactions and measures of aptitudes could be used for adapting instructional treatments to learner characteristics in order to reach learning goals more efficiently. Cronbach and Snow

(1969) described different ways of dealing with learner individual differences "from procrustean methods that involve little adaptation, through intuitive and little tested rules for adaptation, up to, in principle, tested rules derived from theory" (p. 175). An advanced theory-based approach requires understanding of factors that cause an individual to learn better from one instructional method than from another.

Tennyson (1975) suggested a pre-task adaptation model according to which certain types of students were assigned to specific instructional treatments based on pre-task measures of aptitude taken before the learning session. The level of prior achievement was considered as an important adaptation variable: "Students with high prior familiarity in a given area may be assigned to an instructional treatment, with minimal instructional support, or to a forward-ranching sequence. On the other hand, students with low prior achievement may require maximal instructional support each step of the way. Such adaptation to individual differences would be a notable step towards individualizing the method of instruction rather than merely the instructional rate" (Tobias 1976, p. 72).

Federico (1999) described the pre-task adaptation model as a macro-treatment approach typical of traditional ATI research. In contrast, an alternative micro-treatment approach could be based on within-task measures taken while students are in the instructional situation. These two approaches could be effectively used together by selecting macro-treatments based on initial pre-task measures, and then refining and optimizing instructional procedures using micro-treatments based on continuous monitoring of learning behavior (Federico 1999). Such a combined macro-micro adaptation strategy is similar to the adaptation approach that has been developed within the cognitive load framework based on the expertise reversal effect (Camp *et al.* 2001; Kalyuga 2006a; Kalyuga and Sweller 2004, 2005; Salden *et al.* 2004, 2006; Van Merrienboer *et al.* 2003; Van Merriënboer and Sweller 2005).

According to cognitive load theory, optimizing executive function in learning processes assumes presenting appropriate and necessary instructional guidance at the right time and continuously removing unnecessary redundant information as the level of learner task-specific expertise gradually increases. Detailed direct instructional support as a substitute for missing knowledge structures should be provided (preferably, in integrated verbal-pictorial and dual-modality formats) for novice learners. At intermediate levels of expertise, a dynamically adjusted mix of direct external support for constructing new knowledge and problem solving practice with reduced support for exercising and strengthening previously acquired knowledge could be optimal for learning. At higher levels of expertise, minimally guided problem-solving or exploratory learning tasks based on applying available knowledge structures could provide cognitively optimal instructional methods. Changes in the task-specific knowledge base need to be dynamically tracked and specific instructional techniques and procedures tailored accordingly.

This general adaptive approach based on gradual decreases in the level of instructional guidance with increases in the level of learner expertise could be implemented using available instructional design principles. For example, the principle of scaffolding suggests using worked examples, completion assignments, and conventional problems combined in a completion strategy (Van Merriënboer 1990; Van Merrienboer and Paas 1989; Van Merrienboer *et al.* 2003). Faded worked examples (Renkl 1997; Renkl and Atkinson 2003; Renkl *et al.* 2002) gradually fade worked-out steps with increased levels of learner expertise by replacing these steps with problem solving sub-tasks. As levels of learner task-specific expertise increase, relatively less-guided exploratory, problem-solving, or game-based environments could effectively assist in learning advanced knowledge and skills in specific task domains.

The described adaptive approach allows different levels of learner control, although it has mostly been realized in a system-controlled format: a computer program or instructor dynamically select an instructional method that is most appropriate for the current level of learner expertise. Still on the early stages of traditional ATI research, Merrill (1975) noted that this research assumed the relative stability of aptitudes, treatments, and system- or instructor-controlled decisions on what treatment is best for the learner. He suggested that, since student attributes were dynamic rather than static and continuously changed from moment to moment, learners should be enabled to adapt learning environments by actively selecting treatments most appropriate to their cognitive states. The learner-controlled approach to the individualization of instruction was considered as an alternative to dynamic tailoring of instruction to learner characteristics.

Despite expected advantages of learner control (e.g., positive learner attitudes and a sense of control), research findings have been inconclusive and more often negative rather than positive in relation to learning outcomes (Chung and Reigeluth 1992; Niemec *et al.* 1996; Steinberg 1989, 1977). The effectiveness of this approach depends on student abilities to select appropriate learning strategies. According to cognitive load theory, the level of learner expertise is a defining factor: students could have control over the content and instructional sequences when they have sufficient knowledge in the task domain. Low-knowledge learners, on the other hand, require appropriate assistance.

One form of such assistance is providing advisement to learners for making their own decisions (Tennyson 1980, 1981; Tennyson and Rothen 1979). The advisement strategy combines a degree of learner control with system-controlled and evidence-based task selection procedures. An advanced form of this approach is an adaptive guidance strategy that provides learners with information on the current level of their knowledge, what to study or practice to achieve mastery, how to sequence learning tasks for gradual transition from basic to more complex strategies, and how to allocate cognitive resources (Bell and Kozlowski 2002; Kozlowski *et al.* 2001). As learners acquire basic knowledge, adaptive guidance tailors subsequently suggested more advanced learning tasks.

In summary, the research on cognitively optimized adaptive instruction strategies within a cognitive load framework is still very limited. Optimal adaptive approaches, methodologies, and conditions of their applicability need to be established in controlled experimental studies. In the absence of comprehensive research- and evidence-based recommendations, most of existing adaptive online environments are based on monitoring learner external characteristics (navigational patterns, learning styles, preferences etc.) rather than deep cognitive characteristics, such as available knowledge base and levels of learner expertise. The following two sections review preliminary studies initiated within a cognitive load framework with an ultimate goal of changing this situation.

Rapid Online Evaluation of Levels of Expertise

The ability to diagnose levels of learner expertise rapidly in real time is an important prerequisite to building dynamically tailored learning environments. As mentioned earlier, our knowledge base in long-term memory effectively defines processing capabilities and the current content of working memory during knowledge-based cognitive activities. Therefore, tracing this content may provide indicators of levels of acquisition of corresponding long-term memory knowledge structures and, consequently, levels of expertise in a given class of tasks. Concurrent verbal protocols could obviously be used to obtain such information, although this method is time consuming and not suitable for online use in adaptive instructional systems. Alternatively, such information could also be obtained by observing how learners approach briefly presented tasks. Based on their well structured knowledge base, experts would immediately see a task within their higher-level schemas. Novices may only identify some random lower-level components. Organized knowledge base in long-term memory is the main factor determining such differences. Learners with more extensive and better organized knowledge would be able to retrieve appropriate higher-level solution schemas.

This general idea was implemented in the first-step diagnostic method: learners were presented with selected tasks for a limited time and asked to rapidly indicate their first step towards solution of each task. Well learned (in many cases, automated) higher-level solution procedures would allow more experienced learners to rapidly generate advance steps of the solution and skip some intermediate steps (Blessing and Anderson 1996; Sweller *et al.* 1983). Different first steps could be indicators of different levels of expertise. This technique was validated in a series of studies in algebra, coordinate geometry, and arithmetic word problem areas. Results indicated significant correlations (up to .92) between performance on the rapid tasks and traditional measures of knowledge (Kalyuga 2006d; Kalyuga and Sweller 2004), with test times reductions by factors of up to 4.9.

In an alternative rapid testing method, learners were presented with a series of potentially possible steps at various stages of the solution procedure, and asked to rapidly verify the correctness of these steps. This rapid verification method is easier to implement in online learning environments. It is also potentially usable for relatively poorly defined task areas when solution steps could not be specified precisely (e.g., when the solution procedure requires drawing graphical representations or when there are several possible solution paths). The method was first validated using sentence comprehension tasks and indicated a significant correlation between performance on rapid tasks and traditional measures of reading comprehension, with test time reduced by factor of 3.7 (Kalyuga 2006c). For the rapid test, a sequence of gradually increasing in complexity sentences was developed including simple, composite, and multiple-embedded sentences. Each sentence was displayed for a limited but sufficient for reading time, and four simple statements related to the content of the sentence were presented sequentially on the computer screen for rapid verification.

Using task domains of kinematics (vector addition motion problems) and mathematics (transforming graphs of linear and quadratic functions), students' rapid verification test scores were compared with results of observations of the same students' problem solving steps using video recordings and concurrent verbal reports. Results indicated significant correlations (respectively .71 and .75), with reductions of testing times in rapid online tests by factors of 3.2 and 3.5 (Kalyuga 2007). The above validation studies suggested a sufficiently high degree of concurrent validity for the first-step and rapid verification methods.

Applying Expertise Reversal Effect to the Design of Adaptive Instruction

The suggested diagnostic methods were used in adaptive online tutorials in the domains of linear algebra equations (Kalyuga and Sweller 2004, Experiment 4; Kalyuga and Sweller 2005) and vector addition motion problems in kinematics (Kalyuga 2006a) for high school students. According to the dynamic tailoring approach, the tutorials provided dynamic selection of levels of instructional guidance that were optimal for learners with different levels of expertise based on online measures of these levels. In learner-adapted groups, at the beginning of training sessions, each learner was provided with an appropriate level of instructional guidance according to the outcome of the initial rapid pretest. Depending on the outcomes of the ongoing rapid tests during the session, the learner was allowed to proceed to the next learning stage or was required to repeat the same stage and then take the rapid test again. At each subsequent stage, a lower level of guidance was provided to learners (e.g., worked-out components of solution procedures were gradually omitted and progressively replaced with problem solving steps), and a higher level of the rapid diagnostic tasks was used at the end of the stage. In control non-adapted groups, learners either studied all tasks that were included in the corresponding stages of the training session of their yoked participants, or were required to study the whole set of tasks available in the tutorial.

In the first study (Kalyuga and Sweller 2004), the allocation of learners to appropriate stages of instructional guidance was based on levels of expertise as measured by the rapid online first-step test. Results indicated that learner-adapted condition resulted in significantly better knowledge gains (differences between post-instruction and pre-instruction test scores) than non-adapted condition, with the effect size 0. 46. In another study (Kalyuga and Sweller 2005), the rapid first-step measures of expertise were combined with measures of cognitive load (subjective ratings of task difficulty). Since the expertise is associated with not only relatively higher-level but also lower-effort performance, combining both measures was expected to produce a better indicator of learner expertise in a domain. A combined cognitive efficiency indicator was used for the initial selection of the appropriate levels of instructional guidance, as well as for continuous monitoring of learner progress and tailoring instruction to changing levels of expertise.

The instructional efficiency is usually defined as the difference between standardized scores for performance and mental effort ratings (Paas and van Merriënboer 1993; Paas *et al.* 2003). Because in dynamic adaptive environments, such indicators are required in real time during the session, the efficiency was defined as a ratio of the current level of performance to the current indicator of cognitive load (according to the common notion of efficiency as a result relative to the cost). Critical levels of efficiency were defined for a class of tasks as criteria for achieving proficiency in this task domain. Results indicated that learner-adapted instruction significantly outperformed non-adapted group on both knowledge and efficiency gains, with effect sizes 0.55 and 0.69 respectively.

Kalyuga (2006a) compared non-adapted instruction with two learner-adapted instructional procedures, one based on rapid verification tests and another on the efficiency indicator. A simple threshold-based definition of the efficiency in this study was different from the previous one. A learner was allowed to proceed to the next, more difficult class of tasks if, in a rapid verification test corresponding to the current task level, the learner correctly verified all the suggested steps up to, but not including, the final numerical answer, and rated the task difficulty as below average (e.g., less than 5 on the 9-point rating scale). Both adaptive conditions outperformed the non-adapted group on a number of indicators (cognitive load rating scale, instruction time, and instructional efficiency). However, there were no significant differences between the two adaptation procedures on all dependent variables.

Thus, adapting task selection procedures dynamically to levels of learner expertise enhanced learning outcomes and supported previous results obtained by Camp *et al.* (2001) and Salden *et al.* (2004) in air traffic control training. Despite differences in performance assessment methods, definitions of instructional efficiency, and task selection algorithms, learner-adapted conditions were superior to non-adapted formats in all these studies. Salden *et al.* (2006) also demonstrated that personalized learner-adapted approaches to selection of learning tasks were superior to non-adapted formats (effect size 0.58) with no differences obtained between specific adaptation procedures (task selection using a system-controlled efficiency-based procedure vs. learner-controlled personalized preference procedure).

Most of the above task selection procedures were system controlled. Possible disadvantages of such models could be decreased levels of motivation and the lack of opportunities for the development of self-regulation skills. Alternative task selection models that may eliminate or reduce these disadvantages are shared-responsibility and advisory models (Van Merriënboer *et al.* 2006). Corbalan *et al.* (2006) investigated a shared control model that first selects a subset of tasks based on learner performance scores and cognitive load ratings (a system-controlled component), and then presents this subset to the learner who makes the final decision (a learner-controlled component). This model was compared to a fully system-controlled procedure in a pilot study using a simulation-based learning environment in the domain of dietetics. Results indicated that learners in the shared control condition demonstrated higher posttest performance scores (effect size 0.25) with lower cognitive load (effect size 0.37) than learners in the system control condition.

Shared responsibility models may vary the level of student control as learners develop self-regulation skills sufficient for selecting learning tasks independently. Advisory models could provide learners with additional support in the task selection process. Based on their adaptive guidance approach, Bell and Kozlowski (2002) demonstrated that providing students with adaptive guidance in addition to learner control in a complex learning environment was beneficial for learners. It significantly improved the acquisition of basic knowledge and skills for novice learners and strategic knowledge and transfer capabilities for more advanced students. In their study, guidance was adapted to three levels of performance (low, medium, and high). More refined levels of adaptability to the individual learner progress need to be investigated. Also, the development of learner skills in self-managing levels of cognitive load (as an essential part of self-regulation skills) needs to be investigated in conjunction with adaptive guidance, shared responsibility, and advisory models.

The quality of adaptive environments depends significantly on the accuracy of information about levels of learner knowledge and skills. It is important to have rich and diagnostically-informative learner models that represent true levels of learner expertise in specific task domains. Using traditional (usually multiple-choice) tests and tracing user interactions with the system are usually imprecise and incomplete. Applying modern artificial intelligence approaches and developing sophisticated intelligent tutoring systems using fine-grained production rule-based learner models (e.g., Anderson *et al.* 1992) allowed a significant increase in the precision of adaptive methodologies. However, implementations of such approaches require complex computational modeling procedures and, therefore, have been limited to few well defined and relatively simple for modeling domains (e.g., programming and mathematics). On the other hand, the models that are used in most adaptive hypermedia and web-based environments are based on few discrete coarse-grained levels of user expertise (e.g., high, intermediate, low levels; De Bra and Calvi 1998). Therefore, an important advantage of the suggested embedded rapid diagnosis-based approach to the design of learner-adapted environments is combining sufficiently high levels of diagnostic precision in constructing learner models with simplicity of implementation.

Conclusion

This paper reviewed empirical findings related to the expertise reversal effect and presented a theoretical explanation of the effect within a cognitive load framework. The reviewed studies provided overall support for the described model. They revealed the predicted pattern of results in relation to expert–novice differences. The effect has been consistently replicated in many experiments with a large range of instructional materials (e.g., tasks requiring declarative and/or procedural knowledge in mathematics, science, engineering, programming, ESL, management, social psychology) and participants (from the primary school to university levels) either as a full reversal (with substantial differences for both novices and experts) or, more often, as a partial reversal (with a non-significant difference for either novices or experts, but with a significant interaction). More than 2200 students participated in the reviewed experimental studies, without counting those in previously reviewed studies of Mayer (2001) and in ATI studies with learner prior knowledge as the aptitude of interest.

The conservative estimates for effect size differences ranged from 0.45 to 2.99, with the overall mid-range value of 1.72. A simplistic interpretation of this number is that if there are effects of a similar magnitude on both sides (for novices and experts), the effect size for each side would be around 0.86, a large-size effect by accepted standards. If there is an effect of relatively lower magnitude on one side (the case in most studies), then there would be an accordingly stronger effect on the other side. Surprisingly, the full reversal (a strong form of the effect) was obtained not only and not mostly in strictly controlled longitudinal studies in which the same novice learners were gradually trained to eventually become experts in specific task domains (e.g., Kalyuga *et al.* 1998), but also and mostly in cross-sectional studies (e.g., Calisir and Gurel 2003, Cooper *et al.* 2001, Leahy and Sweller 2005, Lee *et al.* 2006, Yeung 1999). In some of these studies, expert–novice differences were assumed based only on years of schooling rather than established by objective prior knowledge measures.

In cognitive load theory, the expertise reversal effect is associated with imbalances between learner organized knowledge base and provided instructional guidance. Two major indicated types of such imbalances are cased by an insufficient learner knowledge base that is not complemented by appropriate instructional guidance (especially at the initial stages of novice learning) and by overlaps between available knowledge of more advanced learners and provided instructional guidance. The need for higher knowledge learners to integrate and cross-reference redundant instructional guidance with available knowledge structures that relate to the same situations may consume additional cognitive resources. A minimal instructional guidance would allow these learners to take advantage of their knowledge base in the most efficient way. In order to balance the executive function and optimize cognitive load, instructional guidance should be provided at the appropriate time, while unnecessary support removed as learners progress to more advanced levels of proficiency in a specific task domain. Adaptive learning environments that dynamically tailor levels of instructional support to changing individual levels of learner expertise in a domain have the best potential for optimizing cognitive load.

Goals represent an important part of a learner knowledge base and play an important executive role in regulating cognitive processing and directing attention. Balancing external guidance with learner internal goal structures is important in maintaining high levels of motivation. Therefore, the inclusion of affective and motivational factors in research on expertise reversal phenomena remains an essential direction for future research (Paas *et al.* 2005; Tobias 1989). There is a close relationship between our motivational states and the operation of working memory that are linked through attention mechanisms (Eysenck 1982). This close relationship was first noted by Simon (1967): “We can use the term *motivation*... simply to designate that which controls attention at any given time” (p. 34). Exploring instructional implications of these mechanisms may provide an important contribution to research in expertise reversal effect and cognitive load theory.

The expertise reversal effect is a logical extension of the aptitude-treatment interaction approach. Although the need to consider levels of learner prior knowledge was recognized early within that approach, few research studies and instructional design recommendations demonstrated explicitly how to use the ATI approach in practice. Aptitudes and instructional treatments were investigated without taking into account associated cognitive processes, and applied psychometric rather than cognitive diagnostic measurement instruments were not suitable for real-time use in adaptive instructional systems.

Even though a cognitive load approach may effectively handle these shortcomings, a limited number of studies in optimal instructional methods that could be used for balancing executive guidance at different levels of learner expertise is a major limitation of the research on expertise reversal effect. Identifying a broader range of instructional methods and procedures that are optimal for learners with different levels of expertise placed in the environments with different task characteristics and formats remains an essential direction for future research. Also, in previous research, mostly well-defined technical domains have been investigated. Extending findings to relatively poorly defined tasks and domains represents another important research direction.

The recent studies in rapid diagnostic assessment methods may offer appropriate real-time tools for the dynamic optimization of instruction, providing adequate fine-grained measures of levels of expertise with sufficient diagnostic power for learner-tailored instructional procedures. The development of adaptive learning environments in different domains (not only for well-defined tasks in technical areas) would also require rapid diagnostic instruments for measuring levels of learner expertise in poorly defined task areas.

The expertise reversal effect represents an important phenomenon that provides an insight into the operation of our cognitive architecture. It has been observed in many studies within and outside of a cognitive load paradigm, as well as supported by previously conducted studies in aptitude-treatment interactions. In practical terms, it provides a valuable guidance for instructional design, especially for the design of learner-adapted instructional systems.

References

Anderson, J. R., Corbett, A. T., Fincham, J. M., Hoffman, D., & Pelletier, R. (1992). General principles for an intelligent tutoring architecture. In V. Shute and W. Regian (Eds.), *Cognitive approaches to automated instruction* (pp. 81–106). Hillsdale, NJ: Erlbaum.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Bell, B. S., & Kozlowski, S. W. J. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and performance in technology-based training. *Personnel Psychology, 55*, 267–306.

Blessing, S. B., & Anderson, J. R. (1996). How people learn to skip steps. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 576–598.

Boutwell, R., & Barton, G. (1974). Toward an adaptive learner-controlled model of instruction: A place for the new cognitive aptitudes. *Educational Technology, 14*, 13–18.

Bracht, G. (1970). Experimental factors related to aptitude–treatment interactions. *Review of Educational Research, 40*, 627–645.

Calisir, F., & Gurel, Z. (2003). Influence of text structure and prior knowledge of the learner on reading comprehension, browsing and perceived control. *Computers in Human Behavior, 19*, 135–145.

Camp, G., Paas, F., Rikers, R., & van Merriënboer, J. J. G. (2001). Dynamic problem selection in air traffic control training: A comparison between performance, mental effort, and mental efficiency. *Computers in Human Behavior, 17*, 575–595.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293–332.

Chung, J., & Reigeluth, C. M. (1992). Instructional prescriptions for learner control. *Educational Technology, 32*, 14–20.

Clarke, T., Ayres, P., & Sweller, J. (2005). The impact of sequencing and prior knowledge on learning mathematics through spreadsheet applications. *Educational Technology Research and Development, 53*, 15–24.

Cooper, G., & Sweller, J. (1987). The 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 procedures and concepts. *Journal of Experimental Psychology Applied, 7*, 68–82.

Corbalan, G., Kester, L., & van Merriënboer, J. J. G. (2006). Towards a personalized task selection model with shared instructional control. *Instructional Science, 34*, 399–422.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–114.

Cronbach, L. (1967). How can instruction be adapted to individual differences. In R. Gagne (Ed.), *Learning and individual differences* (pp. 23–39). Columbus, OH: Merrill.

Cronbach, L., & Snow, R. (1969). *Individual differences in learning ability as a function of instructional variables (Final Report)*. Stanford, CA: School of Education, Stanford University.

Cronbach, L., & Snow, R. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.

De Bra, P., & Calvi, L. (1998). AHA! An open Adaptive Hypermedia Architecture. *New Review of Hypermedia and Multimedia, 4*, 115–139.

DeStefano, D., & LeFevre, J-A. (2007). Cognitive load in hypertext reading: A review. *Computers in Human Behavior, 23*, 1616–1641.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406.

Eysenck, (1982). *Attention and arousal: Cognition and performance*. Berlin: Springer-Verlag.

Federico, P-A. (1980). Adaptive instruction: Trends and issues. In R. Snow, P-A. Federico and W. Montague (Eds.), *Aptitude, learning, and instruction: Vol. 1, Cognitive process analyses of aptitude* (pp. 1–26). Hillsdale, NJ: Erlbaum.

Federico, P-A. (1999). Hypermedia environments and adaptive instruction. *Computers in Human Behavior, 15*, 653–692.

Ginns, P., Chandler, P., & Sweller, J. (2003). When imagining information is effective. *Contemporary Educational Psychology, 28*, 229–251.

Glaser, R. (1977). *Adaptive instruction: Individual diversity and learning*. New York: Holt, Rinehart, and Winston.

Kalyuga, S. (2005). Prior knowledge principle. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 325–337). New York: Cambridge University Press.

Kalyuga, S. (2006a). Assessment of learners' organized knowledge structures in adaptive learning environments. *Applied Cognitive Psychology, 20*, 333–342.

Kalyuga, S. (2006b). *Instructing and testing advanced learners: A cognitive load approach*. New York: Nova Science.

Kalyuga, S. (2006c). Rapid assessment of learners' proficiency: A cognitive load approach. *Educational Psychology, 26*, 613–627.

Kalyuga, S. (2006d). Rapid cognitive assessment of learners' knowledge structures. *Learning and Instruction, 16*, 1–11.

Kalyuga, S. (2007). When less is more in cognitive diagnosis: A rapid assessment method for adaptive learning environments. *Journal of Educational Psychology*, (in press).

Kalyuga, S. (2007). Relative effectiveness of animated and static diagrams: An effect of learner prior knowledge. *Computers in Human Behavior, 23*.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). Expertise reversal effect. *Educational Psychologist, 38*, 23–31.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors, 40*, 1–17.

Kalyuga, S., Chandler, P., & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology, 92*, 126–136.

Kalyuga, S., Chandler, P., & Sweller, J. (2001a). Learner experience and efficiency of instructional guidance. *Educational Psychology, 21*, 5–23.

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001b). When problem solving is superior to studying worked examples. *Journal of Educational Psychology, 93*, 579–588.

Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Educational Psychology, 96*, 558–568.

Kalyuga, S., & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. *Educational Technology Research and Development, 53*, 83–93.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problem hard? Evidence from Tower of Hanoi. *Cognitive Psychology, 17*, 248–294.

Kozlowski, S. W. J., Toney, R. J., Mullins, M. E., Weissbein, D. A., Brown, K. G., & Bell, B. S. (2001). Developing adaptability: A theory for the design of integrated-embedded training systems. In E. Salas (Ed.), *Advances in human performance and cognitive engineering research* (Vol. 1) (pp. 59–123). Amsterdam: Elsevier Science.

Lambiotte, J. G., & Dansereau, D. F. (1992). Effects of knowledge maps and prior knowledge on recall of science lecture content. *Journal of Experimental Education, 60*, 189–201.

Leahy, W., & Sweller, J. (2005). Interactions among the imagination, expertise reversal, and element interactivity effects. *Journal of Experimental Psychology Applied, 11*, 266–276.

Lee, H., Plass, J. L., & Homer, B. D. (2006). Optimizing cognitive load for learning from computer-based science simulations. *Journal of Educational Psychology, 98*, 902–913.

Leutner, D. (1993). Guided discovery learning with computer based simulation games: Effects of adaptive and non adaptive instructional support. *Learning and Instruction, 3*, 113–132.

Lohman, D. F. (1986). Predicting mathemathanic effects in the teaching of higher-order thinking skills. *Educational Psychologist, 21*, 191–208.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research, 59*, 43–64.

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*, 1–19.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge, MA: Cambridge University Press.

Mayer, R., & Gallini, J. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*, 715–726.

Mayer, R., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*, 187–198.

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development, 43*, 31–43.

Mayer, R., Stiehl, C., & Greeno, J. (1975). Acquisition of understanding and skill in relation to subjects' preparation and meaningfulness of instruction. *Journal of Educational Psychology, 67*, 331–350.

McNamara, D., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*, 1–43.

Merrill, M. D. (1975). Learner control: Beyond aptitude–treatment interactions. *AV Communication Review, 23*(2), 217–226.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

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.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Niemec, P., Sikorski, C., & Walberg, H. (1996). Learner-control effects: A review of reviews and a meta-analysis. *Journal of Educational Computing Research, 15*, 157–174.

Ollerenshaw, A., Aidman, E., & Kidd, G. (1997). Is an illustration always worth ten thousand words? Effects of prior knowledge, learning style, and multimedia illustrations on text comprehension. *International Journal of Instructional Media, 24*, 227–238.

Paas, F., Tuovinen, J., Tabbers, H., & van Gerven, P. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63–71.

Paas, F., Tuovinen, J. E., van Merrienboer, J. J. G., & Darabi, A. A. (2005). A motivational perspective on the relation between mental effort and performance. *Educational Technology Research and Development, 53*, 25–34.

Paas, F., & van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental-effort and performance measures. *Human Factors, 35*, 737–743.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction, 12*, 61–86.

Potelle, H., & Rouet, J. F. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *International Journal of Human-Computer Studies, 58*, 327–345.

Reisslein, J. (2005). *Learner achievement and attitudes under varying paces of transitioning to independent problem solving*. Doctoral dissertation, Arizona State University.

Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analysis. *Learning and Instruction, 16*, 92–103.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21*, 1–29.

Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skills acquisition: A cognitive load perspective. *Educational Psychologist, 38*, 15–22.

Renkl, A., Atkinson, R. K., Maier, U. H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *Journal of Experimental Education, 70*, 293–315.

Robinson, D. H., & Molina, E. (2002). The relative involvement of visual and auditory working memory when studying adjunct displays. *Contemporary Educational Psychology, 27*, 118–131.

Salden, R. J. C. M., Paas, F., Broers, N. J., & van Merriënboer, J. J. G. (2004). Mental effort and performance as determinants for the dynamic selection of learning tasks in air traffic control training. *Instructional Science, 32*, 153–172.

Salden, R. J. C. M., Paas, F., & van Merriënboer, J. J. G. (2006). Personalized adaptive task selection in air traffic control: Effects on training efficiency and transfer. *Learning and Instruction, 16*, 350–362.

Schnotz, W., & Rasch, T. (2005). Enabling, facilitating, and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology Research and Development, 53*, 47–58.

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction, 13*, 227–237.

Seufert, T., & Brünken, R. (2006). Cognitive load and the format of instructional aids for coherence formation. *Applied Cognitive Psychology, 20*, 321–331.

Shapiro, A. M. (1999). The relationship between prior knowledge and interactive overviews during hypermedia-aided learning. *Journal of Educational Computing Research, 20*, 143–167.

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127–190.

Shin, E. C., Schallert, D. L., & Savenye, W. C. (1994). Effects of learner control, advisement, and prior knowledge on young students' learning in a hypertext environment. *Educational Technology Research and Development, 42*, 33–46.

Shute, V. J., & Gluck, K. A. (1996). Individual differences in patterns of spontaneous online tool use. *Journal of Learning Sciences, 5*, 329–355.

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review, 74*, 29–39.

Snow, R. E. (1989). Aptitude–treatment interaction as a framework for research on individual differences in learning. In P. L. Ackerman, R. J. Sternberg and R. Glaser (Eds.), *Learning and individual differences. Advances in theory and research* (pp. 13–59). New York: W. H. Freeman.

Snow, R. (1994). Abilities in academic tasks. In R. Sternberg and R. Wagner (Eds.), *Mind in context: Interactionist perspectives on human intelligence* (pp. 3–37). Cambridge, MA: Cambridge University Press.

Snow, R., & Lohman, D. (1984). Toward a theory of cognitive aptitude for learning from instruction. *Journal of Educational Psychology, 76*, 347–376.

Steinberg, E. R. (1977). Review of student control in computer-assisted instruction. *Journal of Computer-Based Instruction, 3*, 84–90.

Steinberg, E. R. (1989). Cognition and learner control: A literature review, 1977–1988. *Journal of Computer-Based Instruction, 16*, 117–121.

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

Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation*, Vol. 43 (pp. 215–266). San Diego, CA: Academic Press.

Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science, 32*, 9–31.

Sweller, J. (2007). Evolutionary biology and educational psychology. In J. S. Carlson, & J. R. Levin (Eds.), *Psychological perspectives on contemporary educational issues*. Greenwich, CT: Information Age Publishing (in press).

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *Journal of Experimental Psychology General, 119*, 176–192.

Sweller, J., Mawer, R., & Ward, M. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology General, 12*, 639–661.

Sweller, J., van Merrienboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

Tarmizi, R., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*, 424–436.

Tennyson, R. (1975). Adaptive instructional models for concept acquisition. *Educational Technology, 15*, 7–15.

Tennyson, R. D. (1980). Instructional control strategies and content structure as design variables in concept acquisition using computer-based instruction. *Journal of Educational Psychology, 72*, 525–532.

Tennyson, R. D. (1981). Use of adaptive information for advisement in learning concepts and rules using computer assisted instruction. *American Educational Research Journal, 18*, 425–438.

Tennyson, R. D., & Rothen, W. (1979). Management of computer-based instruction: Design of an adaptive control strategy. *Journal of Computer-Based Instruction, 5*, 126–134.

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology Applied, 3*(4), 257–287.

Tobias, S. (1976). Achievement treatment interactions. *Review of Educational Research, 46*, 61–74.

Tobias, S. (1987). Mandatory text review and interaction with student characteristics. *Journal of Educational Psychology, 79*, 154–161.

Tobias, S. (1988). Teaching strategic text review by computer and interaction with student characteristics. *Computers in Human Behavior, 4*, 299–310.

Tobias, S. (1989). Another look at research on the adaptation of instruction to student characteristics. *Educational Psychologist, 24*, 213–227.

Tuovinen, J., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology, 91*, 334–341.

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rose, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science, 31*, 3–62.

Van Merriënboer, J. J. G. (1990). Strategies for programming instruction in high school: Program completion vs. program generation. *Journal of Educational Computing Research, 6*, 265–287.

Van Merrienboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learners mind: Instructional design principles for complex learning. *Educational Psychologist, 38*, 5–13.

Van Merrienboer, J. J. G., & Paas, F. G. W. C. (1989). Automation and schema acquisition in learning elementary computer programming: Implications for the design of practice. *Computers in Human Behavior, 6*, 273–289.

Van Merriënboer, J. J. G., Sluijsmans, D., Corbalan, G., Kalyuga, S., Paas, F., & Tattersall, C. (2006). Performance assessment and learning task selection in environments for complex learning. In D. Clark and J. Elen (Eds.), *Advances in Learning and Instruction* (pp. 201–220). Amsterdam: Elsevier Science.

Van Merriënboer, J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*, 147–177.

Yeung, A. S. (1999). Cognitive load and learner expertise: Split attention and redundancy effects in reading comprehension tasks with vocabulary definitions. *Journal of Experimental Education, 67*, 197–221.

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.