Handbook of Research on Effective Electronic Gaming in Education

Richard E. Ferdig
University of Florida, USA

Volume II
Handbook of research on effective electronic gaming in education / Richard E. Ferdig, editor.

Summary: "This book presents a framework for understanding games for educational purposes while providing a broader sense of current related research. This creative and advanced title is a must-have for those interested in expanding their knowledge of this exciting field of electronic gaming"--Provided by publisher.

Includes bibliographical references.


1. Simulation games in education--Handbooks, manuals, etc. 2. Electronic games--Handbooks, manuals, etc. I. Ferdig, Richard E. (Richard Eugene)

LB1029.S53H36 2008
371.397--dc22

2007052787

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book set is original material. The views expressed in this book are those of the authors, but not necessarily of the publisher.
Chapter XLI
Evaluating and Managing Cognitive Load in Games

Slava Kalyuga
University of New South Wales, Australia

Jan L. Plass
New York University, USA

ABSTRACT

This chapter provides an overview of our cognitive architecture and its implications for the design of game-based learning environments. Design of educational technologies should take into account how the human mind works and what its cognitive limitations are. Processing limitations of working memory, which becomes overloaded if more than a few chunks of information are processed simultaneously, represent a major factor influencing the effectiveness of learning in educational games. The chapter describes different types and sources of cognitive load and the specific demands of games on cognitive resources. It outlines information presentation design methods for dealing with potential cognitive overload, and presents some techniques (subjective rating scales, dual-task techniques, and concurrent verbal protocols) that could be used for evaluating cognitive load in electronic gaming in education.

INTRODUCTION

The field of gaming and play-based virtual environments as a new educational technology and research area is rapidly expanding (e.g., Gee, 2003; Nelson, Ketelhut, Clarke, Bowman, & Dede, 2005; Shaffer, 2006). If we expect this technology to be efficient in helping students to acquire new, complex knowledge and skills, its design should be based on knowledge of our cognitive architecture and its role in learning and problem solving. Processing limitations of working memory represent a major factor influencing the effectiveness of learning and performance, especially for novice learners. For example, committing limited cognitive resources to processing irrelevant, non-essential, distract-
Evaluating and Managing Cognitive Load in Games

Inadequate information; on searching for inadequately located references; or on trying to make essential connection between sources of information that are artificially separated in space or time due to poor interface design could substantially slow down learning and performance.

Considering these limitations is particularly important for educational gaming technologies because games usually require simultaneous performances of several cognitive and motor activities. For example, in the game Peeps, designed as part of the RAPUNSEL project to teach middle-school girls how to program, players have to navigate the 3D virtual environment, search for objects of value, communicate with other players, avoid gobblers who try to steal from them, and collect peaches to maintain their energy level (Plass, 2007a). The educational portion of the game, aimed at learning a Java-like programming language in order to design outfits and dances for their avatar, makes additional requirements on the players’ cognitive resources. Efficient information designs therefore must focus on substantially reducing cognitive stress in order to enhance learning outcomes.

Levels of learner prior knowledge and experience in a domain represent another important related factor that may significantly influence learning from educational games. Performance and learning characteristics of experienced learners differ considerably from those of novices. Well-organized and often fully or partially automated schematic knowledge structures allow more experienced learners to rapidly recognize and categorize familiar patterns of information without overloading working memory, thus avoiding cognitive stress (Sweller, van Merriënboer, & Paas, 1998; van Merriënboer & Sweller, 2005). The information design in educational games should support the rapid acquisition and use of such knowledge structures by reducing or eliminating unnecessary cognitive overload that may otherwise prevent the allocation of sufficient cognitive resources required for efficient learning and performance.

It should be noted that cognitive load—that is, the demand on cognitive resources during problem solving and reasoning—is always associated with conscious cognitive processes that take place in the learner working memory while performing a current cognitive task. Therefore, the issue of cognitive overload is different from (although it may be related to) problems of general information content overload over longer periods of time or perceptual overload that is traditionally considered in interface design and usability evaluation procedures (e.g., Nielsen, 1995). Cognitive load theory is dealing with factors that influence conscious information processing as we perform a specific task in real time on a scale of seconds or minutes rather than hours or days (in other words, we are dealing with micro- rather than macro-level analysis).

Many games have procedures in place that have the potential to overcome high cognitive load for critical tasks, for example, by explicitly providing critical information to solve a task on demand and just in time (Gee, 2003), though the effectiveness of these strategies in reducing cognitive load has not yet been tested empirically.

Evaluation of general usability characteristics of various software applications, including educational games, is traditionally aimed at ensuring that interface components are understandable and recognizable (e.g., have clear meanings and interpretations, employ simple and consistent color-coding schemes, use recognizable and consistent metaphors, use simple and clear language, and provide help if required), and are functionally efficient (e.g., have clear functional roles, provide fast feedback and response times, are easy to recover from errors, and provide clear exit paths) (Nielsen, 2000). Evaluation of cognitive load has not been considered as part of such procedures yet, although there have been some clear indirect indications of possible cognitive overload in the gaming environments. For example, Lim, Nonis, and Hedberg (2006) noted that while being motivating, multi-user virtual gaming environments may also distract from learning because of their high levels of immersiveness and interactivity. In
such environments, users could become heavily involved in interactions that are not related to learning processes, although such activities obviously consume cognitive resources that become unavailable for meaningful learning.

Evaluation of cognitive load characteristics of electronic gaming applications should become an important part of their usability studies. Experience accumulated in this area is very limited (if at all available), since most related research studies have been conducted using multimedia learning environments with relatively low interactivity, such as diagrams with on-screen and/or narrated text, instructional animations, and so forth (see Mayer, 2005, for a recent comprehensive overview of the field). In addition, there have been some preliminary studies of cognitive load issues in the closely associated area of instructional simulations (Lee, Homer, & Plass, 2006). Whereas most gaming environments involve lower-level drill-and-practice activities, such as the honing of skills in the MMPORG World of Warcraft, the goal of educational gaming technology applications is the development and practice of higher-level cognitive skills. In many cases, lower-level drill-and-practice activities and higher-level cognitive processing will compete for learners’ limited cognitive resources, making educational games a prime candidate for research in evaluating and managing learners’ cognitive load with a high ecological validity.

This chapter begins with an overview of a contemporary model of human cognitive architecture, and its role in performance and learning. The chapter then describes different types and sources of cognitive load, design methods, and techniques for dealing with potential cognitive overload. Cognitive load factors that could potentially influence efficiency of interactive gaming applications are analyzed (e.g., levels of element interactivity, their spatial and temporal configurations, redundant representations, representational formats used for input parameters, levels of learner prior experience in a task domain, levels of provided instructional support). General methods for evaluating cognitive load are discussed and the application of concurrent verbal reports for evaluating sources of potential cognitive overload is described.

**COGNITIVE ARCHITECTURE FOR LEARNING AND PERFORMANCE IN EDUCATIONAL GAMES**

Games in general, and educational games in particular, pose unique requirements on the perception and processing of information by gamers/learners. Some of the attractions of the different types of games—for example, the immersive nature of the environment, the discover-type nature of game-play, or the strong emotional impact that games can have—all pose requirements on learners’ cognition. An understanding of the way we process information in learning and performance is critical in order to leverage the game paradigm for educational purposes.

Existing theoretical models of human cognitive architecture and available empirical evidence regarding its functioning in learning and performance indicate several major characteristics that underline operation of this system (Sweller, 2003, 2004; Sweller et al., 1998; van Merriënboer & Sweller, 2005). Firstly, our cognitive architecture includes a large store of organized information with effectively unlimited storage capacity and duration. The concept of long-term memory (LTM) as an organized knowledge base that contains a massive amount of schematic knowledge structures is a specific manifestation of this feature.

Secondly, our cognitive architecture has a functional mechanism that limits the scope of immediate changes to the above information store (that could otherwise be potentially disruptive and damaging for the store). The concept of working memory (WM) and the available associated empirical evidence support this feature (e.g., Baddeley, 1986; Miyake & Shah, 1999). Some models consider WM a separate component of an information processing system, while other models regard WM as an
activated part of LTM. Nevertheless, the essential common attribute of most existing models of WM is its severe limitation in capacity and duration when dealing with novel information.

Thirdly, considering the very restrictive conditions of slow and incremental changes to our knowledge base, most of the information in the long-term knowledge store is borrowed from other sources by being actively reconstructed in WM, reorganized and integrated with available knowledge structures in LTM (rather than being completely rediscovered individually). When suitable sources of knowledge are not available, only partly available, or when the information is truly new, the major mechanism for the generation of new knowledge is a random search for information or solution moves followed by tests of their effectiveness (e.g., see Newell & Simon, 1972, for theories of human general-problem solving mechanisms in unfamiliar situations and corresponding empirical evidence).

Finally, an essential characteristic of our cognitive system is its ability to organize complex situations or tasks, appropriately direct our attention, and coordinate different cognitive activities under conditions of severe WM limitations. It is assumed that information structures in LTM are capable of performing this organizing and governing (executive) role for the cognitive system, and there are effectively no limitations on the amount of the organized information in LTM that can be used for this purpose within WM. The concept of long-term working memory (Ericsson & Kintsch, 1995) provides theoretical and empirical underpinnings for this assumption. In the presence of the relevant organized knowledge base in LTM, WM can effectively handle an unlimited amount of information, organize very complex environments, and govern very rich cognitive activities. In the absence of such knowledge structures, the system must employ search-and-test procedures that require significant WM resources.

Our cognitive system tends to minimize cognitive resources involved in performance of a cognitive task, applying what could be considered a “cognitive economy principle.” Using available knowledge structures to guide cognitive activities is a more resource-efficient and preferable option than relying on alternative search procedures with associated cognitively taxing chains of reasoning. This tendency to minimize cognitive resources may cause the system to select for this guiding role incorrect knowledge structures that may seem suitable for the task but are inappropriate. For example, misconceptions are usually well-entrenched, durable, and relatively simple structures (for example, scientific misconceptions, etc.) that may require less WM resources when governing cognitive processes.

**COGNITIVE LOAD IN LEARNING**

Processing limitations of our cognitive system are a major factor influencing learning and performance. Educational games can, depending on their specific implementation, impose a heavy requirement on the cognitive system due to resources required, for example, for navigational tasks, searching for and processing of hidden cues, or processing of complex narratives and contextual information. Working memory is limited in duration and capacity when dealing with unfamiliar information, and it is easily overloaded if more than a few chunks of new information are processed simultaneously (e.g., Baddeley, 1986; Miller, 1956; Peterson & Peterson, 1959). On a very simple level, we all are experiencing this limitation vividly when trying to dial an unfamiliar phone number that we have just heard (if it contains more than seven or eight digits). Prior knowledge structures held in LTM allow us to effectively reduce these limitations and eliminate WM overload by encapsulating many elements of information into larger, higher-level units that could be treated as elements in WM (Chi, Glaser, & Rees, 1982). To continue the above simple phone number example, if a subset of digits in that number coincides with one’s year of birth or other familiar
number, the task would be noticeably simplified because that familiar combination of digits would represent a single chunk in WM.

Cognitive load could also be reduced by practicing skills until they become automated and do not require conscious controlled processing in WM (Kotovsky, Hayes, & Simon, 1985; Shiffrin & Schneider, 1977). For example, we can carry on a meaningful conversation while involved in a familiar game on a handheld computer when our computer skills are highly automated due to extensive practice. In this case, the basic computer-operating and game rules do not require explicit conscious processing in WM, and cognitive resources (WM capacity) become available for attending and responding to other people. This may not happen when we just start learning how to use the device. All cognitive resources will be devoted to learning. Thus, available LTM structures and levels of their acquisition define the characteristics of WM: its actual content, capacity, and duration.

Accordingly, the characteristics of learning and performance change significantly with the development of learners’ expertise in a specific domain. In the absence of relevant prior knowledge, novices are dealing with many novel elements of information that may easily overload their working memories. These learners require considerable external support to build new knowledge structures in a relatively efficient manner. In contrast, experts may rely on retrieval and application of available long-term memory knowledge structures to handle situations and tasks within their area of expertise. There are no working memory limitations for knowledge-based performance of more proficient learners.

Available knowledge structures not only allow us to chunk incoming information, they also perform executive function when we construct new knowledge and guide our performance. For example, in our everyday life, we easily handle many familiar situations by having acquired knowledge about a variety of types of situations that are recognized, activated, and used to govern our performance (buying groceries, paying bills online, using a DVD player, etc.). Each type is associated with a set of cognitive representations (schemas) that are stored in LTM and provide an executive guidance when activated in a specific situation. Similarly, when involved in an educational game, we construct and continuously update a situation model, based on our prior schemas for the task, recent moves, and incoming information. This situation model directs our attention and governs our performance in real time. In the absence of an appropriate knowledge base in LTM, we would resort to random search processes and trial-and-error attempts to handle the situation. Although such strategies usually allow us to eventually reach the goal in most cases, they are cognitively extremely inefficient because they are associated with high levels of cognitive load (Sweller, 1988).

Alternatively, direct instructions and guidance can perform an executive role by providing a partial substitute for the missing knowledge-based guidance for novices by telling them exactly how to handle the situation or solve a task. In fact, many games provide players with just-in-time critical information on how to perform a task or solve a problem, or allow players to receive this information on demand (Gee, 2003). According to fading effect (Renkl, Atkinson, & Große, 2004), more help should be provided initially when players are less experienced, and this help should be faded out as the players progress and acquire new skills.

However, in a recent study comparing the effectiveness of the exploration of computer simulations for science education to direct instruction, using a worked-out version of the simulation that did not require exploration, Plass et al. (2007b) found that simulation exploration was overall more effective than the direct instruction materials. This research provides preliminary evidence that with the right design, the educational value of simulations and, by extension, the value of educational games may exceed the value of direct instruction.

The relative share of available LTM knowledge structures and optimal direct external guidance for
a task depends on the level of learner expertise. While for novice learners, external guidance may provide the only available executive function, for experts in the domain, all necessary knowledge structures could be available in LTM. At intermediate levels of expertise, these two sources of information need to complement each other. In a situation where no executive guidance is provided for dealing with new elements of incoming information by either of these providers, users must resort to general search strategies, which are very inefficient as learning means. This may happen, for example, when minimally guided exploratory gaming environments are used with learners who have no prior knowledge in the task domain. However, the immersive environments of games have the potential to provide contextual information and situate the learning task in a way that makes it easier for learners to identify related LTM knowledge structures than traditional direct instruction approaches are able to achieve.

On the other hand, there could be an overlap between LTM knowledge structures and external guidance when both are available for dealing with the same situation, game moves, or units of information. In this case, a user would have to relate and cross-reference the overlapping components of the executive guidance. This process of reconciling the related components of available LTM knowledge base and externally provided guidance would likely impose an additional WM load. Consequently, less capacity could be available for new knowledge acquisition and performance improvement, resulting in a phenomenon that has been referred to as the expertise reversal effect (for recent overviews, see Kalyuga, 2005, 2006b, 2007). For example, presenting experienced users with detailed guidance (e.g., “how-to” instructions or detailed game rules explanations designed for novice users) that they do not need any more may hinder their performance relative to other similar experienced users who have not been instructed. Therefore, as levels of learner expertise in a domain increase, relative effectiveness of different design formats may reverse. Presentation formats that are optimal for novices may hinder relative performance of more experienced users. A major design implication of this effect is that information presentation formats and instructional procedure need be tailored to different levels of learner expertise in a specific task domain (Kalyuga, 2006a).

MANAGING COGNITIVE OVERLOAD IN EDUCATIONAL GAMING APPLICATIONS

We have established above that educational games and simulations impose demands on learners’ information processing that go beyond that of many direct instruction approaches. However, there are important cognitive benefits of using simulated environments (including games) as learning tools compared, for example, to experimenting with actual objects in real-life situations. Simulated game environments may help to present only the most essential features of the systems or environments under investigation, and thus allow formulating and testing specific hypotheses and receiving immediate feedback in otherwise very high-load situations. It is also believed that simulated (including game-based) environments enhance learners’ abilities to apply acquired knowledge in complex real-life situations because such environments allow students to learn the specific context in which to use their domain-specific knowledge. For example, Taylor and Chi (2006) compared learning effects from reading a text and using a computer simulation in the domain of project management. Results of pre-test to post-test gains for abstract deep structural and decontextualized knowledge indicated that participants in both conditions improved equally. In contrast, a contextualized case-based assessment of implicit domain knowledge demonstrated significantly improved learning outcomes for the simulation group only. Tennyson and Breuer (2002) noted advantages of using complex task-oriented simulated environments in learning a task as a
complete whole rather than successive parts with a focus on improving and elaborating cognitive problem-solving abilities.

On the other hand, if designed inappropriately, game-based educational environments may contain many sources of high-level cognitive load that prevent effective learning. There are different types and sources of cognitive load in learning. A major type of cognitive load is caused by cognitive activities that are essential for establishing key connections between elements of information, integrating them with available knowledge structures, and building new (or modified) knowledge structures in WM (i.e. cognitive activities associated with comprehension of the situation and knowledge-based response actions). This type of load is referred to as intrinsic cognitive load. It is caused by internal intellectual complexity of the task that is determined by the degree of interactivity between individual elements relative to the specific level of learner expertise in the domain. An element is a highest-level chunk of information for a particular person. The content of various chunks is determined by the schemas the user holds in her or his long-term memory knowledge base.

One of the aspects of the game paradigm that is most appealing to educators is that the educational content can be presented in an integrated, systems-based form rather than in isolation and disconnected from contextual information. However, this integrated representation means that the elements related to a learning task need to be processed simultaneously. Even if the number of elements is relatively small, the material still could be high in element interactivity and may impose a high intrinsic cognitive load. For example, understanding a gaming simulation of a complex environmental system is much more difficult than figuring out the type of each individual element of this system. Even if all elements of the system are well known to a person in isolation (assuming that he or she has pre-acquired schemas for each of those components), when combined in the system they become interconnected and need to be considered simultaneously as a whole in order to understand the simulation. Once the interactions of the components of the system have been learned, lower-order schemas become the elements of a higher-order schema that can further act as a single element reducing the required cognitive effort.

Because intrinsic cognitive load is essential for comprehending a situation or performing a task, it is vital to provide all the necessary resources to accommodate the intrinsic cognitive load without exceeding limits of working memory capacity. In contrast, extraneous load (wasteful, non-constructive load) is traditionally described as a diversion of cognitive resources on activities irrelevant to performance and learning (Sweller et al., 1998). This load is caused by cognitive activities that a user is involved in because of design-related factors (e.g., poor interface design, presentation format, or task sequencing). For example, when related textual, graphical, or audio elements of a gaming application are separated over distance or time, their integration might require intense search processes and recall some elements until other elements are attended and processed. Segments of text need to be held in working memory until corresponding components of a diagram are located, attended, and processed; or images need to be maintained in active state until corresponding fragments of the text are found, read, and processed. Such processes need additional resources and might significantly increase demands on working memory.

Searching for suitable solution steps may also involve keeping a large number of interacting statements in working memory and require significant cognitive resources that become unavailable for other essential cognitive activities. For example, high cognitive demands of familiarization with game rules, evaluating game states, and making specific next-step decisions may leave no cognitive resources available for generalizations and acquisition of meaningful knowledge structures. These cognitive demands are irrelevant to the learning goals and should be considered as an extraneous cognitive load.
In general, extraneous cognitive load could be imposed by one or more of the following sources (brief labels in capital letters will be used later for classifying these sources for evaluation purposes):

1. Separated (in space and/or time) related representations that require users to perform extensive search and match processes (SPATIAL SPLIT-ATTENTION, TEMPORAL SPLIT-ATTENTION);
2. An excessive step-size or rate of information change that introduces too many new elements into working memory and/or introduces them too fast to be successfully incorporated into long-term memory structures (EXCESSIVE INFORMATION);
3. An insufficient externally provided guidance that does not compensate for limited available knowledge, thus forcing users to search for solutions using random procedures (SEARCH); and/or
4. User knowledge base overlaps with provided external guidance thus requiring learners to mentally co-refer different representations of the same information (REDUNDANCY).

The interesting challenge of applying the game paradigm to educational materials is that the previously established definition of extraneous load does not readily apply to games. Although the information design and interaction design of educational games must be informed by the above principles, games often deliberately violate these principles. For example, games routinely separate related objects and information in time and space, and expect the user to find and connect them (item 1, above), use excessive step-size or rate of information change to solve a problem or reach the next level (item 2), or purposefully provide insufficient guidance (item 3). The difference to traditional, direct instruction approaches is that each of these strategies can contribute to achieving an educational goal, in which case they would not be considered wasteful or non-constructive (extraneous cognitive load), but would be strategies to cognitively engage the learner (which is referred to as germane cognitive load).

The intrinsic and extraneous cognitive load results in the total cognitive load imposed on the learner cognitive system. For efficient performance and/or learning, total cognitive load should not exceed limited working memory capacity. When a task does not require high levels of intrinsic cognitive load (e.g., because it is low in element interactivity relative to the current level of learner expertise), the extraneous cognitive load imposed by poor design may not do much harm because total cognitive load may not exceed working memory capacity. In contrast, when the task is characterized by a high degree of element interactivity relative to the person level of expertise, it might require a heavy intrinsic cognitive load. In such situations, an additional extraneous cognitive load caused by an inappropriate design can leave insufficient cognitive resources for efficient performance and/or learning because total cognitive load may exceed the user working memory capacity. Elimination or reduction of extraneous cognitive load by improving interface design, presentation formats, or task procedures may be critical for performance.

In correspondence with the above sources of extraneous cognitive load, the general guidelines for minimizing extraneous cognitive load in gaming applications suggest providing direct guidance and access to required knowledge base, avoiding diversion of cognitive resources on redundant and/or irrelevant cognitive activities, managing step-size and rate of information changes, and eliminating spatial and temporal split of related sources of information. In cases where educational games deliberately deviate from these approaches, designers need to assure that such a deviation is required by a specific learning strategy that contributes to an educational objective.
EMPIRICAL STUDIES OF COGNITIVE LOAD FACTORS IN EDUCATIONAL GAMING APPLICATIONS

As was mentioned in the introduction, direct studies of cognitive load effects in game-based learning environments are extremely rare and mostly limited to the role of instructional guidance as an important factor in reducing high-load situations. Providing sufficient levels of instructional guidance and support for learners is an important means of reducing extraneous cognitive load and improving learning-effective engagement in interactive gaming environments. Just providing learners with higher levels of control that allow them to potentially access additional assistance (e.g., hints) may not be sufficient. For example, it was established that many learners in virtual environments simply did not use the available hint system (e.g., Nelson, 2007; Nelson, Ketelhut, Clarke, & Dieterle, in press).

Moreno and Duran (2004) investigated benefits of guidance in discovery multimedia game-based learning environments in elementary school mathematics. Two representations of the arithmetic procedures were used for addition and subtraction problems: a traditional symbolic representation of the number sentence and a visual representation. The visual representation used a number line and an animated bunny moving along the line according to the number operations performed (facing the left or right sides of the screen if the corresponding numbers have the minus or plus signs). If learners answered a problem correctly, they could see an animated sequence demonstrating major steps in solving the problem. In the guided group, the learners also could hear explanations for each step of the animation. It was assumed that combining symbolic and visual representations could help learners (especially less knowledgeable novice learners) to build connections between formal procedures and their informal intuitive conceptual knowledge (moving along a path). The results demonstrated that learning new mathematical procedures could be overwhelming for novice learners when no guidance is provided. The verbal guidance is an important means to enhance learning in game-based multimedia environments using multiple representations.

Another important result of this study was that students' lower computer proficiency could undermine the potential benefits of learning in game-based environments: high-computer experience learners, especially those with verbal guidance, outperformed low-experience learners in similar conditions (see also Clarke, Ayres, & Sweller, 2005, for a similar conclusion based on studies of learning mathematics using spreadsheet applications). High cognitive demands of familiarization with gaming hardware and corresponding functional procedures may leave no cognitive resources available for acquisition of meaningful domain-specific knowledge structures. A practical implication of this result is that, from a cognitive load point of view, it is important to bring students to a sufficiently high level of computer proficiency prior to involving them in exploring complex interactive environments.

The instructional effectiveness of games could be low (especially for learners with low levels of prior knowledge) if no sufficient instructional support is provided and students are involved in pure discovery learning. A discovery-based computer game may have positive learning effects only when students have sufficient cognitive resources to process multiple representations and sources of information in working memory. Such resources can only be available if the learners have good prior familiarity with the corresponding knowledge domain. Mayer, Mautone, and Prothero (2002) demonstrated that students learned better from a geology game when they received explicit guidance about how to visualize geological structures. Moreno (2004) found that students benefited more from explanatory rather than merely corrective feedback in a multimedia game about environmental problems.
Leutner (1993) investigated learning effects of two forms of instructional guidance—system-initiated adaptive advice and learner-requested non-adaptive pre-tutorial background information—provided in an exploratory computer-based simulated game environment built around the economical situations of small farms for geography high-school classes. Students without any support learned how to play the game but acquired hardly any domain-specific knowledge. On the other hand, with adaptive advice, students were able to acquire a substantial degree of domain knowledge (as measured by an immediate post-test), but limited functional knowledge of how to play the game. In relation to background information, the results indicated that if permanently available, it increased the acquisition of domain knowledge (as measured by a delayed memory retention test). Generally, the available results provide evidence that guided game-based learning environments are more instructionally effective than pure discovery-based games, especially for low-knowledge learners.

There are few studies of cognitive load factors in learning from instructional simulations. Since in most cases educational games are a form of simulated interactive environments, results of these studies could also be applied to educational games. Plass et al. (in press) investigated an interaction between two different modes of visual representations in a gas law simulation for high-school chemistry students and different levels of learner prior knowledge. Essential gas characteristics were presented either in 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, high prior knowledge learners benefited more from symbolic-only representations. Iconic representations were redundant for these learners and interfered with their knowledge-based cognitive processes. Similar to other studies of expertise-related reversals in effectiveness of different verbal and pictorial representation formats (Kalyuga, 2005, 2007), low prior knowledge learners benefited more from integrated multiple (verbal-pictorial or symbolic-iconic) representations than from separated or single representation formats. On the other hand, more knowledgeable learners benefited more from minimal representations. Additional representations were redundant for these learners and interfered with their learning processes.

In another relevant study, Schnotz and Rasch (2005) compared effects of two different formats of animated visualizations of time phenomena related to the Earth’s rotation on learners with different levels of learning prerequisites (a combination of pre-test scores of prior knowledge and intelligence measures). One format displayed an animated visual simulation of changes over time when circumnavigating the Earth (simple simulation). Another format represented an interactive visualization that allowed students to manipulate the display by defining specific day and time for specific cities (interactive simulation). The post-test results indicated that students with high learning prerequisites performed significantly better after learning from the interactive simulation, while lower learning prerequisite students performed better after learning from the simple animated simulation. In this expertise reversal pattern, interactive manipulations could have imposed high levels of extraneous cognitive load on novice learners but were optimal for more experienced learners.

**METHODS FOR EVALUATING COGNITIVE LOAD**

The low number of published research studies on cognitive load in games is in part due to the limited number of methods available to measure cognitive load. There are several indirect methods available that use indicators such as learning outcomes, physiological measures, and behavioral measures in order to determine cognitive load (Brünken,
Plass, & Leutner, 2003). Among the more direct measures that could apply to the measurement of cognitive load in games are subjective and objective approaches to load measurement.

A very rough relative measure of cognitive load could be obtained by using subjective ratings of mental effort, for example, by asking users How easy or difficult was the game to understand? or How hard did you have to think to understand how to play the game? with answers on nine-point Likert scales from extremely easy (1) through neither easy nor difficult (5) to extremely difficult (9). Previous research has indicated that such simple measures could be sufficiently sensitive to variations in cognitive load conditions (Paas, Tuovinen, Tabbers, & van Gerven, 2003). However, as there is no an absolute scale for subjective ratings of mental effort, they are more useful for comparing cognitive load levels involved in alternative applications or interface designs rather than evaluating a single game application. This approach could nevertheless be used for comparing cognitive load imposed by sequential versions of an application in the iterative process of the redesign of components that could contribute to increased cognitive load conditions. The same users could be asked to rate mental effort involved in using the application after each modification stage.

Another method for comparative evaluation of cognitive load that can be used in laboratory settings is the dual-task technique. This method uses performance on simple secondary tasks as indicators of cognitive load associated with performance on main tasks. Various simple responses can be used as secondary tasks, for example, reaction times to some events (e.g., a computer mouse click), counting backwards, or recalling the previous letter seen on the screen of a separate computer while encoding the new letter appearing after a tone sounded. An important requirement is that a secondary task should affect the same working memory processing system (visual and/or auditory) as the primary task; otherwise, it may not be sensitive to changes in actual cognitive load.

Dual-task techniques for measurement of cognitive load in multimedia learning environments were studied by Brünken et al. (2003, 2004) and Brünken, Steinbacher, Plass, and Leutner (2002). In these studies, the secondary task represented a simple visual-monitoring task requiring learners to react (e.g., press a key on the computer keyboard) as soon as possible to a color change of a letter displayed in a small frame above the main task frame. Reaction time in the secondary monitoring task was used as a measure of cognitive load induced by the primary multimedia presentation. The studies demonstrated the applicability of the dual-task approach to measurement of cognitive load experienced by each individual learner.

In order to evaluate cognitive load characteristics of a single gaming application (without comparisons to a variant of the same application), concurrent verbal reports (think-aloud protocols) with audio and video tracking could be used. Although verbal protocols can hypothetically increase cognitive load, there is no evidence that they actually interfere with cognitive processes (Ericsson & Simon, 1984). The generated qualitative verbal data may reflect different types of cognitive load as expressed through the participants’ own language (Kalyuga, Plass, Homer, Milne, & Jordan, 2007; Plass, Toler, & Kalyuga, 2007). In these studies, verbal data from think-aloud interviews was coded using rubrics based on expected learners’ verbal expressions or remarks for different types of cognitive load (see examples below). For each rubric, sample keywords and phrases were set and served as a coding scheme for classifying participants’ remarks into different categories of cognitive load. Verbal data from the protocols can be analyzed by screening digital recordings of each interview. Such records include the synchronized audio and screen captures with screen and audio recording software (e.g., TechSmith’s [2007] Camtasia Studio 5) using the samples of expected responses.

Before commencing the procedure, participants were instructed to think aloud (e.g., “It would really help me to understand what you are thinking while
Evaluating and Managing Cognitive Load in Games

playing the game. I am particularly interested in knowing what you find clear and what you find unclear. If you get quiet, I will ask you to keep talking”). Participants were given only a brief general overview of what the game is about; they were not specifically instructed on how it should be played. Throughout the interview, general probes were used to elicit relevant remarks (e.g., What is your strategy for playing? What are you learning? What is familiar to you? What is unfamiliar? What information are you paying most attention to? What do you ignore?). The probes did not explicitly mention difficulty, effort, and other concepts directly related to the notion of cognitive load (Kalyuga et al., 2007; Plass et al., 2007).

The analysis of verbal protocols for indicators of cognitive load consisted of locating relevant words, remarks, and expressions and relating them to different sources of cognitive load. The following rubrics were used for this purpose, with samples of participants’ remarks:

- Does the application provide sufficient explanations (guidance)? (can’t get an idea, too complex to understand, don’t know what to do, need some hints)
- Are users involved in extensive search processes (e.g., trial-and-error)? (let’s try and see, just enter anything, play with numbers)
- Does the application activate relevant prior knowledge? (don’t know anything about it, never heard about it, it doesn’t ring a bell, thought it was something else)
- Does the application explain things that are already known to the learner? (know this stuff, we did it differently, studied this before)
- Do the unnecessary explanations distract from learning? (it is annoying, need to go through this again, it doesn’t tell me anything new)
- Are too many new elements of information introduced too quickly? (can’t catch it, the information is changing too fast)
- Does the application proceed by too large step-sizes? (plenty of new things, can’t grasp them all, a lot of unknown information)
- Does the application include related verbal and/or pictorial components that need to be studied simultaneously and are located in different parts of the display or not synchronized? (need to jump across the screen, it is over there, this has changed earlier)
- Does the understanding of the interrelated components require extensive co-referencing and temporary holding of much information in memory? (need to go back to the diagram or text, too much to remember, already forgot about that)

The concurrent verbal reporting method for evaluating cognitive load was tested with two versions of an interactive simulation of gas laws for science education similar to those used in the above-mentioned study of Lee et al. (2006). The following are samples of actual learners’ remarks, indicating various forms of extraneous cognitive load (see Table 1).

Based on the total number of extraneous cognitive load-relevant remarks for each condition obtained from 12 participants (university students with very limited knowledge of chemistry), symbolic versions (simulations using text labels for important concepts) were more cognitively demanding (34 remarks) than iconic versions (simulations using icons in addition to text labels for important concepts) (12 remarks). This relative cognitive load standing of conditions coincided with the reversed order of similar conditions based on low-knowledge learner post-test performance obtained in a study with 64 high-school students (Kalyuga et al., 2007), thus providing preliminary evidence for the validity of the suggested evaluation method. Adding dynamic iconic images helped to alleviate cognitive overload by providing a relatively less cognitively demanding interactive learning environment.
### FUTURE TRENDS: TOWARDS ADAPTIVE EDUCATIONAL GAMING

Building adaptive educational gaming environments with optimized cognitive load based on flexible and learner-tailored support is an important future trend with clear educational benefits. There are different general approaches to building complex adaptive learning environments that could also be used in game-based learning. For example, Federico (1999) suggested combining a macro-treatment pre-training adaptation approach based on pretest results with a micro-treatment approach based on within-training measures taken while students are in the instructional situation. Macro-treatments could be selected based on initial pre-task measures, then instructional procedures could be refined and optimized using micro-treatments based on continuous monitoring of learning behavior. An approach proposed by Tennyson and

#### Table 1.

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPATIAL SPLIT-ATTENTION</td>
<td>Watching them all at the same time could be difficult. Confusing, not sure what to do. A little hard to isolate things. I didn’t pay attention to see the actual change on the graph. A lot of things to look at at once. I look at the numbers, but try to look at the graph too. It’d be easier if there was one graph instead of two, so I’d focus on it. A lot is going on at once, numbers are changing. This is really hard. Paying attention to gas particles, and now trying to pay attention to the graph.</td>
</tr>
<tr>
<td>TEMPORAL SPLIT-ATTENTION</td>
<td>I must go back to see previous pressure results. Clicking everything that was before disappears. I forgot what I did at the previous one [step]. Need to refer back to previous step to see the change. Diagram shows past trials, container shows this moment. It would be easier if my old graphs stay on top of the screen. It’s difficult to keep track of previous simulation. I’m manipulating and seeing, but I keep storing it. It’s difficult to look back. I think I missed something, I go back to see pressure-temperature relationship.</td>
</tr>
<tr>
<td>REDUNDANCY</td>
<td>Repetitions, I’ve already realized the relationship from first two [steps]. Extra stuff; flames get in the way (no need to show six flames to show temperature rising).</td>
</tr>
<tr>
<td>EXCESSIVE INFORMATION</td>
<td>Everything’s moving at the same time. It is difficult to figure out what’s happening. It’s difficult because of storing everything in my mind. I have to remember that I’m moving Temperature or Pressure [sliders]. Lost track what I am doing. A lot of things. I checked the solution but did not remember. Looking at text, not seeing visuals (so fast). Too much going on the screen; extra things, like a flashy show.</td>
</tr>
<tr>
<td>SEARCH</td>
<td>I’m just picking numbers.</td>
</tr>
</tbody>
</table>
Breuer (2002) consists of two main components: a curriculum (macro-) component that maintains a student model and an external knowledge base, and an instructional (micro-) component that adapts the instructional strategies according to current learning progress.

An adaptive approach that has been developed within the cognitive load framework is based on the expertise reversal effect (Camp, Paas, Rikers, & van Merriënboer, 2001; Kalyuga, 2006c; Kalyuga & Sweller, 2004, 2005; Salden, Paas, Broers, & van Merriënboer, 2004; Salden, Paas, & van Merriënboer, 2006; van Merriënboer & Sweller, 2005) and assumes the initial selection of an optimal level of learner support based on pre-task measures, and then refining and optimizing instructional procedures using continuous monitoring of learning behavior. According to cognitive load theory, optimizing learning processes requires presenting appropriate and necessary instructional guidance at the right time and continuously removing unnecessary redundant information as the level of learner expertise in a domain gradually increases. Detailed direct instructional support should be provided (preferably in integrated verbal-pictorial and dual-modality formats) for novice learners. Changes in the task-specific knowledge base need to be dynamically tracked and specific instructional procedures tailored accordingly.

The described adaptive approach has mostly been realized in experimental computer-based tutorials in a system-controlled format: a computer program dynamically selects an instructional method that is most appropriate for the current level of learner expertise. The learner-controlled approach could be an alternative to dynamic system-controlled tailoring of instruction to learner characteristics. Despite some expected advantages of learner control (e.g., positive learner attitudes and a sense of control), research findings have been more often negative rather than positive in relation to learning outcomes (Niemec, Sikorski, & Walberg, 1996; Steinberg, 1989). According to cognitive load theory, the level of learner expertise is a defining factor: students could have control over the content and instructional sequences if they have sufficient knowledge in the task domain. Low-knowledge learners, on the other hand, require appropriate assistance. This assistance could be provided as advice to learners to make their own decisions (Tennyson, 1981). 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 & Kozlowski, 2002).

The available research on cognitively optimized adaptive strategies within a cognitive load framework is very limited. Optimal adaptive methodologies and conditions of their applicability need to be established in controlled experimental studies. In the absence of evidence-based recommendations, most existing adaptive online environments are based on monitoring navigational patterns, learning styles, preferences, and other external learner characteristics rather than deep cognitive characteristics, such as available knowledge structures.

CONCLUSION AND IMPLICATIONS

This chapter makes a first contribution to apply cognitive load theory to the design of educational game environments. Research in cognition and instruction has substantially expanded our understanding of mental processes involved in learning, limitations of our cognitive system, and the role of learner prior knowledge. Applying this knowledge to the design of educational games is a necessary condition for their effectiveness.

Games have unique features that place higher demands on learners’ cognitive resources than more traditional direct instruction approaches. Examples for sources of these demands are the need to navigate immersive 3D environments, the
use of discovery-based approaches to the gameplay, the manipulation of learners’ emotions, the need to find hidden cues, or the use of narratives to provide situational context. In addition to the cognitive demands of these features, educational games require the learner to invest cognitive resources in the processing of the content that is to be learned. Therefore, special attention must be devoted to eliminate all sources of unproductive processing of extraneous information. Sources of excessive extraneous cognitive load that may inhibit performance and learning from educational gaming applications are: spatially and/or temporally split elements of information that need to be integrated in order to achieve understanding; an excessive step-size and/or rate of information presentation that introduces too many new elements of information into working memory too quickly to be organized and comprehended; insufficient user support or guidance, especially for low prior knowledge users; and excessive redundant support overlapping with available knowledge of more experienced users.

It is important to recognize that in educational games, design decisions that in traditional direct instruction approaches would have been considered sources of extraneous load may in fact be contributing to the educational objectives of the game, and would therefore be categorized as generating germane load—that is, engaging users in the learning process. For example, games can represent educational content embedded in a rich context, which requires resources to be processed but makes the presented information more meaningful to learners and allows them to connect it to existing knowledge structures. Recent research indicates that in some cases, exploratory environments may indeed be more effective than direct instruction (Plass et al., 2007b). This is an indication that when designed to meet a specific educational purpose, the higher cognitive demand of game-specific features may result in increased mental effort, and as a result in increased learning.

In addition to established techniques of measuring cognitive load (subjective rating scales, dual-task methods), concurrent verbal reports with audio and screen capture of learners’ online behavior may be used for evaluating and comparing levels of extraneous cognitive load in educational games as an important part of their usability studies. Based on such evaluation procedures, educational games could be improved to better match the nature of the human cognitive architecture. For example, direct guidance could be provided to low prior knowledge users at the appropriate time (or on request), unnecessary redundant support could be timely removed as a learner becomes more experienced with the task domain, step-sizes and rates of presentations could be limited to ensure that the learners’ cognitive capacity is not exceeded, split-attention effects could be eliminated or reduced by integrating graphics and text or using auditory modality for presenting verbal elements, and information presentations could be dynamically tailored to changing levels of learner proficiency in the domain. Ultimately, adaptive educational games could expand current fading techniques to allow dynamic tailoring of presentations to changing cognitive characteristics of individual learners to work in harmony with human cognitive architecture. Given the growing research in cognitive load issues in learning, researchers and game designers should be aware of these developments and their implications, as well as methods for evaluating and managing cognitive load in educational games.

**AUTHOR NOTE**

The research presented in this chapter was supported in part by the Institute of Education Sciences (IES), U. S. Department of Education (DoEd) through Grant R305K050140. The content of this publication does not necessarily reflect the views or policies of IES or DoEd, nor does any mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.
References


KEY TERMS

Cognitive Architecture: A general cognitive system that underlies human performance and learning. The understanding of human cognition within a cognitive architecture requires knowledge of corresponding models of memory organization, forms of knowledge representation, mechanisms of problem solving, and the nature of human expertise.

Cognitive Load: Working memory resources required for processing specific information by an individual user. Cognitive load theory distinguishes between the essential (intrinsic) and wasteful (extraneous) forms of cognitive load, and suggests a variety of techniques and procedures (cognitive load effects) for managing essential and reducing extraneous load in learning.

Cognitive Load Theory: An instructional theory describing instructional implications of processing limitations of human cognitive architecture (capacity and duration of working memory) and evolved mechanisms for dealing with these limitations (long-term memory knowledge base and its role in cognition).

Expertise Reversal Effect: Reversal in the relative effectiveness of information presentation formats and procedures as levels of user knowledge in a domain change. For example, extensive external support could be beneficial for novices when compared with the performance of novices who receive a low-support format, but is disadvantageous for more expert users when compared with the performance of experts who receive a low-support format.

Long-Term Memory (LTM): A major part of our cognitive architecture, an organized knowledge base that stores a massive amount of hierarchical knowledge structures.

Sources of Cognitive Load: Features of external information structures or cognitive characteristics of individual users that determine required working memory resources. Intrinsic cognitive load is caused by levels of interactivity between elements of information that need to be processed simultaneously. Extraneous cognitive load is imposed by the design of information presentations (e.g., separated in space- and/or time-related elements; an excessive step-size or rate of introducing new elements of information; limited user knowledge that is not compensated by provided support; user knowledge base that overlaps with provided external guidance).

Working Memory (WM): A major part of our cognitive architecture, a functional mechanism that limits the scope of immediate changes to long-term memory. Depending on a specific model, WM is considered either as a separate component of our cognitive system or as an activated part of LTM. The essential attribute of WM is its severe limitation in capacity and duration when dealing with novel information.