Shared control over task selection

Helping students to select their own learning tasks

The research reported here was carried out at the:

OpenUniversiteitNederland

was funded by:

NWO

The Netherlands Organisation for Scientific Research
Project no. 411-02-107-V

in the context of the research school:

ico

Interuniversity Center for Educational Research

ISBN: 978-90-79447-02-2

Cover design: Marco van Hout and Gemma Corbalan Perez
Printed by Datawyse, Maastricht, The Netherlands

Shared control over task selection
Helping students to select their own learning tasks

Proefschrift

ter verkrijging van de graad van doctor
aan de Open Universiteit Nederland
op gezag van de rector magnificus
prof. dr. ir. F. Mulder
ten overstaan van een door het
College voor promoties ingestelde commissie
in het openbaar te verdedigen

op vrijdag 25 april 2008 te Heerlen
om 15:30 uur precies

door

Gemma Corbalan Perez

geboren op 16 oktober 1975
te Mataró (Barcelona), Spanje

Promotor:
Prof. dr. J.J.G. van Merriënboer, Open Universiteit Nederland

Toegevoegd promotor:
Dr. L. Kester, Open Universiteit Nederland

Overige leden beoordelingscommissie:
Prof. dr. J.H. van Driel, Universiteit Leiden
Prof. dr. B. Gros, Universitat Oberta de Catalunya
Prof. dr. D. Leutner, Universität Duisburg-Essen
Prof. dr. R.L. Martens, Universiteit Leiden / Open Universiteit Nederland
Prof. dr. F. Paas, Open Universiteit Nederland
Dr. H. Tabbers, Erasmus Universiteit Rotterdam

A mis padres
A mis hermanos
A Marco

Acknowledgements

‘Things can change’, is what my family and friends in Barcelona must have thought when I first came to the Netherlands as an Erasmus student in 2000 for a 4-months period and stayed a year longer following a master’s program at the University of Twente. Things changed even more when in December 2003 I returned to the rainy Netherlands and became a PhD student at OTEC.

I would never have been able to finish my dissertation without the help, support, guidance, and efforts of my supervisors Jeroen van Merriënboer and Liesbeth Kester. I thank them for their generous time and commitment whenever it was needed and for keeping their doors and inboxes always opened for me. Jeroen, thank you for the inspiring discussions, for guiding and teaching me throughout this research project, and for providing me with an excellent atmosphere for doing research. Liesbeth, thank you for your inspiring ideas, enthusiasm, and continuous support, (sometimes beyond your obligations, I will never forget our trip to Bakel!). Also thanks for all the laughter about daily issues. Het was gezellig!. It has been a true pleasure to work with both of you. Also, thank you for not sharing my worries about the planning and fully trusting that the ‘end’ will come to a ‘good end’.

I am also very grateful for having an exceptional assessment committee. Special thanks to Fred Paas for the continuous opportunities during the last months.

I would also like to thank Ingrid, Audrey, Nicole, and Mieke, for their great support and assistance while completing my dissertation. I am also thankful to the coordinators and teachers who made the difficult task of finding participants for my experiments possible: Irma Rabelink (ROC Eindhoven), Joep Gulikers, Hans Collijn, Marinus Thoonsen, and Marlene Smeets (Arcus College Heerlen). Thanks to Monito, especially Bas Jansen, for the programming of all the learning environments for my experiments and the hundreds of emails we exchanged. Also thanks to Hub Kurvers and Wim Slot for their technical assistance, and to Hanneke Kester, Lily Fredrix, Joop de Kraker, and Daniel Burgos for the content-expert advice for the experiments.

Thanks to the group of PhD students for the ‘gezellige’ atmosphere during the last years, research discussions, social events, making the conference trips worth double, improving my Dutch, and assisting me during my experiments: Wendy, Fleurie, Amber, Sandra, Femke, Greet, Marjo, Danny, and Helen. Wendy for sharing many personal and professional discussions and Fleurie for sharing the best day of her life with me. Sandra and Femke, mijn twee lieve kamergenootjes en paranimfen!, bedankt voor een supergezellige kamer, vriendschap, al onze project-gerelateerde en niet-projectgerelateerde discussies en voor onze ‘forever 25 years’. Greet, mijn lieve buurvrouw, je bent meer dan een collega, je bent ook een vriendin. These thanks are certainly extended to the former PhD students: Liesbeth B., Tamara, Marieke, Pieter, Karen, Judith, Silvia, and Ron; and other colleagues for the nice atmosphere, without being exhaustive: Ellen, Iwan, Adriana (gracias por no hacerme olvidar mi idioma materno), Olga, Dominique (bedankt voor de goede tip!). Liesbeth B., me ha encantado tenerte como kamergenootje y amiga! Bedankt voor de gezellige atmosfeer en discussies in onze mooie kamer, in conferenties, en daarbuiten. Ellen, bedankt voor onze leerzame Nederlands-Spaanse avonden en alle ‘buiten-OTEC’ gelegenheden. Ya sabes que te aprecio no sólo como compañera, sino también como amiga.

During these years I had the great opportunity to share some months with our Spanish visitors at the OUNL: Prado, JuanMa, Davinia, Pablo, and Luis. Thanks not only for the ‘very authentic’ moments, but also for the ‘very authentic’ Spanish products, including the jamón ibérico!

My dearest friends Elena and Amaya are in part responsible of how ‘things have changed’. After years of sincere friendship, the three of us decided to go for the Erasmus experience together and discovered not only that, as compared to the Netherlands, ‘in Spain, when rain, rain’ but also many positive aspects of the Netherlands and the ‘Dutchies’. Amayita, guapa, mira lo que has conseguido. Me sigues debiendo una, pero yo también te la debo a tí. Elena, brufi, as always we keep following the same path, me alegro mucho de que hayais venido a vivir a Holanda y tenerte de vez en cuando para mí. Also special words go to Ana. Anocha, me alegro mucho de haberte conocido en Enschede y mantenerte como amiga, me alegra mucho que coseches tantos éxitos y espero poder combinarlo para vernos más. I miss the three of you.

Also special thanks goes to Daniel and Rocio. Barcelona is not so bad for two Madrileños when you get used to see the beach and the Sagrada Familia every day!. Thank you Daniel for being a great colleague, team-mate (as long as pure chocolate was supplied) and great friend. Rocio, thank you for your friendship and for sharing the stress of the last months completing our dissertations. Also special thanks to my friends Marivi, Natalia, Juana Mari y Rosalía, just for being there and forgiving me for being far away from them.

Other friends I have the pleasure to share some nice moments with in Maastricht: Orsi, Aurelie (two new, proud mothers), Sabrina, Maud, and Sjoerd. Thanks for being interested in how things developed and understanding the constraints to meet, especially during the last months.

The last part of my acknowledgements includes the most special thanks in the personal sphere.

First, I would like to thank my mother Marti, my father Pedro, my brothers David and JuanJo, and my sister Marta for their continuous love and support, their daily phone calls and patience when often - specially during the final months of the dissertation - I wasn’t around because my project needed me. However, I can’t say that my project needed me more than they did, especially in some occasions, and therefore this dissertation is especially dedicated to them. Gracias por vuestro apoyo y vuestra comprensión, sobre todo en estos últimos meses!.

And finally, I would like to thank Marco... Chiqui, thank you for all your support, for our long phone-calls when we still lived more than 1500 kilometres away from each other, for being how you are, and for loving me so much. I can't wait for our wedding the 31st of Augustus in Barcelona!

Gemma Corbalan Perez
Maastricht, 2008

Contents

1

General Introduction

This chapter briefly introduces the different components that play a role in personalized, dynamic selection of learning tasks. It discusses the task features that can be dynamically selected throughout the training as well as variables that form the basis for task selection. In addition, it describes approaches to the selection of tasks using program control, learner control, and so-called shared control. Finally, it presents an overview of the studies that were carried out to investigate under which conditions learner control over task selection is most effective.

In the field of instructional design, there is a tendency to focus on authentic learning tasks based on real-life situations to help learners develop transferable skills (Reigeluth, 1999; van Merriënboer & Kirschner, 2001). However, especially for less experienced learners in a domain, complex skill acquisition by performing authentic tasks can easily cause cognitive overload because of the limited processing capacity of working memory (Baddeley, 1992; Sweller, 1988). To enable the use of authentic and complex tasks in education, and yet prevent overloading learners' cognitive system, there is a need to flexibly personalize the training program to each individual learner. This shift to personalized instruction is in line with current educational approaches which focus on flexible curricula. These curricula move from 'same for all' education which presents all learners with the same sequence of learning tasks, towards 'just for me' education which offers each learner a unique task sequence dynamically adapted to individual needs and preferences.

Task sequence can be personalized either by a computer program or by the learner herself. At one extreme, *program controlled approaches* provide each individual learner a sequence of suitable tasks, which can be determined at each moment in time. This requires systematic and real-time assessments of the learner's progress (e.g., Camp, Paas, Rikers, & van Merriënboer, 2001; Jochems, van Merriënboer, & Koper, 2005; Kalyuga & Sweller, 2005; Renkl & Atkinson, 2003; Salden, Paas, & van Merriënboer, 2006; van Merriënboer & Kirschner, 2007; van Merriënboer, Sluijsmans, Corbalan, Kalyuga, Paas, & Tattersall, 2006). At the other extreme, *learner controlled approaches* provide learners the freedom to choose their own learning path according to their own needs and/or interests (Merrill, 1980; van Merriënboer, Schuurman, de Croock, & Paas, 2002). Learner control may be beneficial for learning even when choices are trivial (Cordova & Lepper, 1996; Katz & Assor, 2007; Kinzie, 1990; Lepper, 1985; Morrison, Ross, O'Dell, & Schultz, 1988). However, in general, studies report both beneficial and detrimental effects of learner control on learning (e.g., Katz & Assor, 2007; Williams, 1996). This seems to indicate that learner control functions differently depending on what (e.g., pace, display, task features) is selected by whom (i.e., domain novices or more experienced learners).

The aim of the studies presented in this dissertation is to investigate under which conditions learner control over task selection is most effective. In the next sections, task features that can be dynamically selected throughout the training as well as variables that form the basis for task selection are examined in more detail. Next, program control, learner control, and, finally, so-called 'shared control' are described as alternative approaches to the selection of learning tasks. Finally, an outline of the remaining chapters of this dissertation is provided.

Task Features in Dynamic Task Selection

Task features may comprise: the level of *difficulty* (from simple to complex, or from easy to difficult) (Chapters 2 and 3); the level of *support* (in a completion or fading strategy embedded support diminishes from 'full support' provided by worked examples that must be studied by the learner, to 'no support' provided by conventional problems that must be solved by the learner) (Chapters 2 and 3); the *surface* or irrelevant task features (Chapters 3, 4 and 5); and the *structural* or relevant task features (Chapters 5 and 6). For each individual learner, it should be possible to select at any given point in time a task or a range of tasks containing optimal task features. Learning tasks that are suitable for presentation as the learner progresses throughout the training are stored in a substantial learning-task database which contains enough variability to fit the learners' needs and interests.

Variables used as a Basis for Dynamic Task Selection

In models for dynamic task selection, decisions are based on a continuously updated learner portfolio which contains *learner variables*, such as the combination of performance and invested mental effort to attain that performance, and *task variables*, such as the surface features and structural features selected in previous tasks, that help to select the best (subset of) future task(s) for learning.

With regard to *learner variables*, *performance* is commonly used as input for dynamic task selection and can be defined as the effectiveness in accomplishing a particular task (Camp et al., 2001). Cognitive load theorists have proposed that performance measures alone are not a sufficient basis for task selection and should be enhanced by considering the imposed cognitive load to attain this performance. Cognitive load is commonly measured by the *mental effort* invested in task performance. Invested mental effort thus refers to the cognitive capacity that is allocated by the learners to accommodate the demands imposed by the task (Paas, Tuovinen, Tabbers, & van Gerven, 2003). A combined measure of performance and invested mental effort forms a more comprehensive representation of the learner's level of expertise than performance alone and can be used for dynamic task selection (e.g., Camp et al., 2001; Kalyuga & Sweller, 2005; Paas et al., 2003; Salden et al., 2006). In the pilot study presented in Chapter 2 and in the study described in Chapter 3, task difficulty and support of each newly selected task are dynamically tailored to the learner variables performance and invested mental effort.

The learner portfolio also contains information on previously presented *task variables* to decide which task features should be selected next. In this dissertation the focus is on *surface* features and *structural* features. Surface features (e.g., species and trait in inheritance tasks because Mendel's laws are the same for animals, plants, and humans) refer to task aspects that are not relevant to how the problem is solved (i.e., solution steps) and are generally salient for learners. Structural features (e.g., the solution steps in inheritance tasks or the underlying mathematical procedure in statistical problems) refer to task aspects that are relevant to solution steps and are generally not salient for especially novices in a particular domain (Chen & Mo, 2004; Cummins, 1992; Gick & Holyoak, 1987; Quilici & Mayer, 1996, 2002). In the studies presented in Chapters 4 and 5, surface features are taken as a basis for task selection. Additionally, Chapters 5 and 6 include task selection rules which also take the structural features of previously performed tasks into account.

Program Control and Learner Control over Task Selection

The issue of the locus of instructional control, that is, either external (program control) or internal (learner control), has been a primary concern in the upsurge of computer-assisted instruction (Lawless & Brown, 1997; Tennyson & Buttery, 1980). *Program control* ensures a suitable sequence of learning tasks according to the learners' needs (e.g., Camp et al., 2001; Kalyuga & Sweller, 2005; Salden et al., 2006; van Merriënboer & Kirschner, 2007). For example, a task that is too difficult or presents not enough support could hamper learning because it may easily overload learners. Task selection rules might avoid cognitive overload by adapting task difficulty and support levels according to learners' current capabilities. However, a high level of program control may negatively affect learners' motivation. *Learner control* - in contrast to program control - enables learners to make instructional decisions according to their current knowledge, interests, and preferences (Merrill, 1980, 1994; van Merriënboer et al., 2002). This is believed to positively influence learning and motivation (Schnackenberg & Sullivan, 2000; Williams, 1996), provided that learners perceive the control given and use this control to select personally relevant tasks (Katz & Assor, 2007).

Finally, the program and the learner may *share control* over task selection. In this two-step, dynamic process, a computer program first selects 'on the spot' a subset of learning tasks with suitable characteristics (e.g., difficulty, support, surface and structural features) as the learner progresses throughout the training (program control). This makes instruction more effective and efficient. Second, the learner selects from this subset one task to work on (learner control). This makes the instructional experience more motivating and relevant for learners which ultimately will yield more effective and efficient learning. However, the effectiveness of shared control largely depends on which task features are selected by the program and which task features are selected by the learners themselves. For example, learners can be given control over task features they are able to perceive and to handle, but not over task aspects that are beyond their knowledge and skills or which are not salient for them.

Structure of this Dissertation

Chapter 2 introduces a task selection model with shared control, in which first a program makes a pre-selection of optimal tasks, from which the learner then makes the final selection. It also reports the results of a pilot study testing this model. The model forms the basis of the study portrayed in Chapter 3 and is used as the starting point of the studies described in Chapters 4 to 6.

Chapter 3 builds on the model proposed in Chapter 2 and reports on the effects of adaptation (present vs. absent) and control over task selection (program vs. shared) on transfer test performance, efficiency, and task involvement. A program dynamically adapted the level of difficulty and available support of learning tasks in the dietetics domain to the learners' competence and reported cognitive load. The program presented either (a) one task - program control - or (b) three tasks - shared control - differing on surface features, from which the learner selected one.

The experiments described in Chapters 4, 5, and 6 (domain: genetics) tested the effects of program control and learner control over the selection of learning tasks with different surface features (Chapters 4 and 5) and different structural features (Chapters 5 and 6) on transfer test performance, efficiency, and motivation. Chapter 4 reports an experimental study in which participants' perception of control was manipulated by task selection rules that limited learners' choices to three program pre-selected tasks containing either similar *or* dissimilar surface features as compared to each prior task. Learners were expected to benefit from choosing pre-selected tasks with surface features dissimilar from the surface features of previously performed tasks.

Chapter 5 explores the effects of program control and learner control over the selection of tasks with different surface features and different structural features on transfer test performance and efficiency. In the program control conditions, one task with dissimilar surface features or dissimilar structural features from the preceding task was presented each time. In the learner control conditions, learners were presented with a subset of four tasks which differed in the level of dissimilarity (low to high) of either surface features or structural features from the previous task, whereafter learners made the final selection. It was expected that it is better to give learners the freedom to select tasks with different surface features, because the saliency of those features enables them to select a varied set of personally relevant tasks. No beneficial effects on learning were expected from learner controlled selection of tasks with different but non-salient structural features.

The study described in Chapter 6 investigates the effects of feedback (present vs. absent) and control over the selection of tasks with different structural features (program control vs. learner control) on transfer test performance, efficiency, and motivation. Since feedback helps learners to recognize the structural features, learners who had control and were given feedback were expected to be better able to perceive the control given to them, and thus to choose more personally relevant tasks, which facilitates learning and motivation. In addition, the sole provision of feedback was also expected to yield higher transfer test performance, efficiency, and motivation.

Finally, Chapter 7 presents an overview and a general discussion of the results of the studies presented in Chapters 3 to 6. It closes with a description of limitations of the presented studies and practical implications for future research.

2

Towards a Personalized Task Selection Model with Shared Instructional Control[1]

Modern education emphasizes the need to flexibly personalize learning tasks to individual learners. This chapter discusses a personalized task-selection model with shared instructional control based on two current tendencies for the dynamic sequencing of learning tasks: (1) personalization by an instructional agent which makes sequencing decisions on the basis of a learner's expertise, and (2) personalization by the learner who is given control over – final – task selection. The model combines both trends in a model with *shared* instructional control. From all available learning tasks, an instructional agent selects a subset of tasks based on the learner's performance scores and invested mental effort (i.e., program-control). Subsequently, this subset is presented to the learner who makes the final decision (i.e., learner control). A computer-assisted instructional program has been developed to put the model into practice and preliminary results are discussed. The model can be used to increase the efficiency and effectiveness of instruction and to make it more appealing by providing the learner an optimal level of control over task selection.

---

[1] This chapter is based on: 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.

Rapid technological developments in modern society increase students' need to acquire complex cognitive skills and to transfer those skills from formal educational settings to work situations (Reigeluth, 1999; van Merriënboer & Kirschner, 2001). There is a tendency in the field of instructional design to focus on authentic learning tasks based on real-life situations to help learners develop transferable skills. But especially for novice learners, the acquisition of complex skills by performing authentic learning tasks is heavily constrained by the limited processing capacity of working memory because it easily causes cognitive overload. According to cognitive load theory (Sweller, 1988; Sweller, van Merriënboer, & Paas, 1998; van Merriënboer & Sweller, 2005), cognitive load is a construct representing the load that performing a particular task imposes on the human cognitive system. Learning is encouraged if the cognitive system is not overloaded and if available cognitive resources are actually allocated to learning processes rather than extraneous processes that do not directly contribute to learning.

In order to enable the use of authentic tasks in education, and simultaneously prevent learners from being overwhelmed by them, there is a need to personalize the nature and sequence of learning tasks for each individual learner. Personalization of learning materials to individual learners is believed to facilitate learning, because (a) the level of difficulty and available support of each task prevents cognitive overload, and (b) other task features are varied in such a way that learning is promoted. This chapter describes a model of personalized task selection which enables the development of personalized education. The model combines two approaches to personalization in order to cope with increasingly complex learning situations: program-controlled instruction, in which an instructional agent (e.g., computer, teacher) makes decisions on learning tasks, and learner-controlled instruction, in which the learner him or herself makes such decisions.

First, according to models of program-controlled instruction, learning is influenced by both characteristics of the learner, such as expertise, abilities, and attitudes (Zimmerman, 2002), and characteristics of the tasks that are presented in the learning environment, such as task complexity, amount of learner support, and other task features (Lawless & Brown, 1997). The use of electronic learning environments allows for personalization by dynamically changing the instruction (i.e., learning tasks) in response to *input from the learner* (e.g., performance scores and invested mental effort to attain them) or based on previously selected *task features* (e.g., surface or irrelevant task features, and structural or relevant task features). Several studies indicate that personalization in response to input from the learner leads to a more efficient training program and higher transfer test performance than the use of a fixed sequence of ready-made learning tasks (e.g., Camp, Paas, Rikers, & van Merriënboer, 2001; Kalyuga & Sweller, 2005; Salden, Paas, Broers, & van Merriënboer, 2004; van Merriënboer, Schuurman, de Croock, & Paas, 2002).

Second, according to models of learner-controlled instruction, complex learning should not only aim at developing complex skills but also at promoting self-regulated learners who are able to effectively select their own learning tasks. Pure program-controlled instruction cannot enhance the acquisition of self-regulation skills because an instructional agent rather than the learner makes the selection. Giving learners – some – control over particular aspects of their learning environment creates the necessary preconditions for practicing self-regulation skills and is a first step towards teaching those skills (Kinzie, 1990). In addition, it is expected to make learning more appealing with beneficial effects on learner motivation, which in turn may increase learning outcomes (Wolters, 2003; Zimmerman, 2002). Giving learners control over task selection assumes that learners are able to select the most suitable learning task according to their current state of knowledge, interests, and preferences (van Merriënboer et al., 2002). Novices in a particular learning domain who are provided with learner control, however, may not be able to pick up essential information for learning so that there is a decrease in learning and learning goals are not reached (Merrill, 2002). More experienced learners have the necessary knowledge to make the right selections, and giving them learner control prevents that they receive information they already know.

Concluding, there should be a gradual transition from program-control to learner-control if learners acquire more expertise. But, even experienced learners may not profit from full learner control if they are overwhelmed by the amount of choice, for instance, if they have to select from hundreds of tasks (Schwartz, 2004). Thus, for a novice learner it is best to select one learning task from a small set of available tasks; for a more experienced learner it is best to select one learning task from a larger set of tasks, but even a highly experienced learner should not select one learning task from a very large set of tasks. The personalized task-selection model with shared instructional control presented in this chapter combines the advantages of both program-controlled and learner-controlled instruction. When a learner works on learning tasks, an instructional agent continuously assesses performance and invested mental effort to select an optimal *subset* of following learning tasks. This subset is then presented to the learner who makes the final selection. Thus, the instructional agent and the learner share control over the whole process of learning task selection. Shared control is expected to be more effective (i.e., higher learner performance) and more efficient (i.e., higher performance combined with lower mental effort invested and/or less instructional time) than both complete program and learner control. In addition, it is believed to be more appealing and to promote the development of self-regulation skills.

The structure of this chapter is as follows. First, a description of the model and its main components is given. Second, the model is translated into practice and a specific learning environment, which is developed on the basis of the model, is described. Finally, the results obtained in a pilot study that examined the effects of the learning environment are described and the implications of the results are discussed.

The Model

The personalized task-selection model with shared instructional control presented in this chapter aims at providing each individual learner, after finishing one or more learning tasks, the best next task given his or her level of expertise, thus yielding a personalized sequence of learning tasks in an environment for complex learning. Figure 2.1 depicts the personalized task-selection model with shared instructional control.

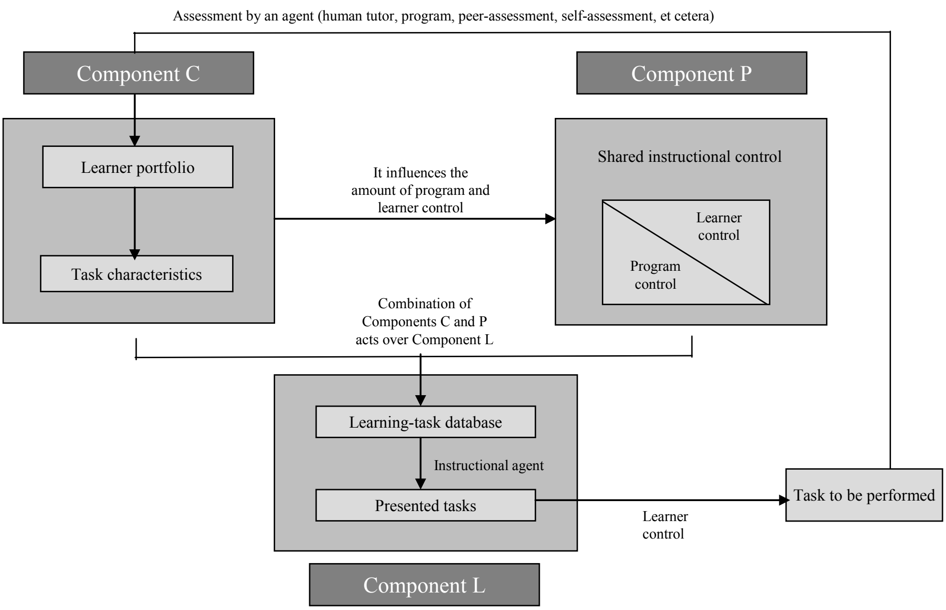


Figure 2. 1. The personalized model for dynamic task selection.

Figure 2.1 gives a diagram of the model including three components: (a) characteristics (component C), (b) personalization (component P), and (c) learning-task database (component L).

Component C includes task characteristics and learner characteristics, which are documented in a learner portfolio. Task characteristics include the level of complexity (i.e., from simple to complex, or from easy to difficult), embedded learner support (i.e., from full support to no support), and other task features (e.g., the surface features, for instance, the context in which the task is performed, display or presentation mode, et cetera). The learner portfolio contains, for instance, information about the learner in terms of task performance and invested mental effort on already completed learning tasks, combined in a measure of expertise.

Component P refers to the personalization mechanism. It combines two apparent opposite tendencies of personalization: program-controlled instruction and learner-controlled instruction. Program-controlled instruction includes task-selection rules used by an instructional agent to base decisions on. Learner-controlled instruction lets the learner select the learning tasks from a smaller or larger subset of – pre-selected – tasks. Both forms complement each other in such a manner that when the program has a relatively large control over task selection, the learner receives a relatively small amount of control, and vice versa. In that way, control is shared by the program and the learner. The relative amount of program and learner control over task selection varies based on the contents of the learner portfolio.

Finally, Component L includes the learning-task database with tasks with diverse levels of complexity, embedded support, and other task features. The instructional agent pre-selects a subset of learning tasks from this database and presents only this subset to the learner, who should make the final selection. Theoretically, the size of this subset may range from 1 (i.e., no learner control) to all tasks in the database (i.e., full learner control).

Component C: Characteristics

Component C refers to learning task characteristics and the learner portfolio. If the characteristics of the task presented to an individual learner are not appropriate this may either hamper learning because of cognitive overload, if the task is too difficult or does not include enough support, or hamper learning because the task features do not stimulate the learner to construct new knowledge. The learner portfolio should contain the information that helps to select the best task(s) for learning. Relevant task characteristics and the learner portfolio are discussed in the next subsections.

Learning task characteristics. The four-component instructional design model (4C/ID-model; van Merriënboer, 1997; van Merriënboer, Clark, & de Croock, 2002) provides guidelines that consider the limitations of working memory for the design of educational programs, or integrated curricula, to teach complex cognitive skills. In an integrated curriculum, learning tasks are ordered in an easy-to-difficult sequence, learner support decreases from high to no support, and learning tasks vary from each other on external characteristics that also differ in the real world (e.g., context, display and presentation mode, input for the task, et cetera). Our personalized task-selection model with shared instructional control takes the 4C/ID-model as a starting point and distinguishes the same three aspects of learning tasks for personalization: (a) level of complexity, (b) embedded learner support, and (c) other task features. Learner support decreases and task complexity increases as learner's level of expertise increases. Moreover, varying the task features increases the variability of practice, which promotes learners to construct new knowledge through abstraction and generalization (van Merriënboer, 1997).

First, with regard to the *level of complexity* it is argued that, due to limited processing capacity, providing very difficult learning tasks right from the start of a curriculum may have negative influences on learning, performance, and motivation (van Merriënboer, Kirschner, & Kester, 2003). In terms of cognitive load theory the difficulty of a task yields intrinsic cognitive load, which is a direct result of the complex nature of the learning material. That is, intrinsic cognitive load is higher when the elements of the learning material are highly interconnected (i.e., high element interactivity) and lower when they are less interconnected (i.e., low element interactivity; for a review, see Sweller, 1994). A personalized instructional design should provide each learner with tasks that are neither too difficult nor too easy. By gradually increasing the difficulty of tasks in such a way that the individual learner's growing expertise is taken into account, the intrinsic cognitive load yielded by these tasks is properly managed.

Second, providing learners with tasks with an appropriate amount of embedded support is essential for learning because this prevents cognitive overload as well. This embedded support is related to the sequence of operators (i.e., the solution steps) necessary to reach the goal state. Different types of embedded support take different aspects of the problem solving process into account. Product support provides the learners with solution steps; process support presents the rationale behind the problem-solving process itself (van Merriënboer, 1997).

Worked-out examples are learning tasks that provide maximum product-oriented support. They confront the learners not only with a description of a given state (i.e., problem state) and the criterion for an acceptable goal state or solution, but also with a description of the solution steps necessary to reach the solution. The learners are asked to carefully study the solution or 'best example'. Worked-out examples prevent extraneous cognitive load, that is, load irrelevant for learning because learners do not have to bother about the tentative application of mental operations but can focus all their attention on learning the relevant solution steps. Completion problems provide learners with a partial rather than a complete solution. Like worked-out examples, they help learners focus on the problem state and relevant solution steps. But in addition, learners have to carefully study the partial solution because otherwise they will not be able to find the remaining solution steps. This directs learners' attention to processes relevant for learning, hereby enhancing so-called germane cognitive load (i.e., load relevant for learning) and facilitating problem solving and transfer performance (Renkl & Atkinson, 2003; van Merriënboer et al., 2002). Conventional tasks, finally, provide learners with a given state and a criterion for an acceptable solution only. No solution is provided. For experienced learners, conventional tasks enhance the generation of creative solutions because the tasks are authentic and the learners already possess the knowledge that is necessary to approach the task. But novice learners lack relevant knowledge and use cognitively demanding strategies to reach the solution, increasing extraneous cognitive load and hampering learning. Thus, a smooth transition from worked-out examples, through completion tasks, to conventional tasks takes learners growth in expertise into account.

In this respect, the ‘expertise reversal effect’ (Kalyuga, Ayres, Chandler, & Sweller, 2003) states that successful instructional techniques for novice learners (e.g., presenting tasks with high embedded support) often lose their value, or even produce opposite effects, when used with more experienced learners. For instance, presenting a diagram with integrated textual explanations may be an effective technique for novice learners, who need the explanations to understand the diagram, but an ineffective technique for experienced learners, who already possess the knowledge necessary to understand the diagram. When they are nonetheless forced to process the textual information that is redundant with their current knowledge, this hinders learning because they must unnecessarily invest cognitive resources (i.e., extraneous cognitive load) to find out that the information is redundant.

Process-oriented support provides learners with heuristics for problem solving that may help them to reach an acceptable solution. Three types of process-oriented support are modeling examples, process worksheets, and performance constraints (van Merriënboer, 1997). *Modeling examples* present the learner with an expert or professional who is demonstrating the problem solving process and explaining the reasoning behind it, thus verbalizing why s/he is doing what s/he is doing. *Process worksheets*, in a paper-based format or as an on-line tool, provide a description of helpful phases and rules-of-thumb to guide learners through the problem-solving process. Finally, *performance constraints* force learners to use expert’s approaches to problem solving while they are performing the learning tasks.

Third, learning tasks should differ from each other on *task features* that also differ in the real world. Examples are surface features, such as the context or setting in which the task is performed, the structural features, such as the solution steps in inheritance tasks, the way the task is displayed, the interface that is used, and the mode of presentation of the task (e.g., video, audio, graphics, or text). By designing tasks with varying task features, variability of practice is increased. Variability of practice enhances germane cognitive load because it promotes learners to develop generalized and abstract knowledge: it increases the probability that similar features are identified and that relevant features are distinguished from irrelevant ones (van Merriënboer, 1997). Moreover, variability of practice is an effective strategy for gaining and keeping a learner’s attention (Paas, Tuovinen, van Merriënboer, & Darabi, 2005). If learners can choose between tasks with different features, they may select a task according to their interests which makes instruction more personally relevant and therefore positively affects learning.

To conclude, the personalized task-selection model with shared instructional control presented in this chapter aims at personalizing the level of task difficulty, the level of embedded learner support, and the other task features. Personalization takes place in such a way that intrinsic load is optimized, extraneous load decreased, and germane load increased – all within the limits of totally available cognitive resources.

Learner portfolio. Besides learning task characteristics, component C also takes learner characteristics, kept in a learner portfolio, into account for task selection purposes. Individual differences between learners strongly influence how instruction must be designed in order to be most effective (Shute & Towle, 2003). According to the personalized task-selection model presented in this chapter, the characteristics of each learner are represented in the portfolio at each specific moment in time, and so provide the basis for the dynamic selection of one or more appropriate learning tasks. Three aspects of the learner taken into account for task-selection purposes are: (a) task performance, (b) invested mental effort, and (c) level of expertise.

The first aspect, *performance*, is commonly used as input for dynamic task selection and can be defined as the effectiveness in accomplishing a particular task (Camp et al., 2001). Continuous assessment is needed for the dynamic personalization of learning tasks. Assessment can either be made by a human or an electronic tutor, a peer (peer-assessment), or even the learner himself (self-assessment). The Protocol Portfolio Scoring (PPS; Straetmans, Sluijsmans, Bolhuis, & van Merriënboer, 2003) argues that to gather reliable and valid information about a learner's performance, complex behaviours must be assessed on multiple aspects, which require different embedded assessment methods. According to the PPS, performance on tasks with support (e.g., worked-out examples, completion tasks) is mainly used to decide on the appropriate level of support and the desirable task features of subsequent tasks. Performance on conventional tasks without support is mainly used to decide whether the learner is ready to advance to a higher difficulty level or whether the learner has completed the whole educational program.

Three often used measures of performance are time-on-task, process quality, and product quality. Assessment of time-on-task is an important measure in learning tasks where speed plays a crucial role for acceptable performance (e.g., physical activities, driving a car, controlling air traffic, et cetera). Time-on-task may indicate how easy or difficult a learning task was for a particular learner because it typically increases with complexity. Accordingly, if time on task is very high for a learner it may also indicate high cognitive load or even overload. Thus, a learner who needs a considerable investment of time to perform a specific learning task should not be given the same subsequent learning task as another learner who performed the same task in considerably less time.

The quality of the process indicates learners' understanding of the principles underlying their responses or solutions. Process data include verbal protocols, video and audio tapes, and retrospective reports. Observation instruments provide information on the accuracy of learners' actions, errors, misconceptions, and so forth. The quality of the available knowledge can also be measured by a variety of traditional assessment tools, such as checklists, multiple-choice questions, open questions and so forth.

The quality of the finished product indicates if the learner already masters required knowledge and skills. Detailed standards of performance in terms of what learners should be able to do, according to particular values and exhibiting particular attitudes, help to assess the quality of the product. Standards must be observable and measurable. For instance, one standard of performance in a course on dieting could be: "learners have to be able to find a balance between the energy intake and the energy expenditure of a person who is on a diet".

The second aspect, *invested mental effort,* refers to " . . . the cognitive capacity that is actually allocated by the learner to accommodate the demands imposed by the task" (Paas, Tuovinen, Tabbers, & van Gerven, 2003, p. 64). Paas et al. (2003) describe two classes of methods for measuring mental effort: subjective techniques (e.g., rating scales, self reports) and objective techniques (i.e., dual-task reaction times, heart-rate variability, pupillary dilation). Cognitive load is commonly measured by the mental effort invested in task performance. According to Paas et al. (2003) a meaningful interpretation of cognitive load measures can only be given in the context of its associated performance and vice versa.

This leads to the third aspect, *learner expertise,* which refers in our model to the combination of performance and mental effort. The 4C/ID-model states that performance measures alone are not a sufficient basis for task selection and should be enhanced by considering the investment of mental effort to attain this performance. Thus, if two learners reach the same performance but one learner invests a lot of mental effort and the other does not, the first learner is best presented with an easier new learning task than the second learner.

Recent methods for the assessment of expertise focus on the combination of measures of performance and cognitive load (van Merriënboer & Sweller, 2005). Dynamic task-selection approaches require the just-in-time gathering of learner's data. Thus, the level of expertise is continuously assessed, updated in the portfolio after each learning task, and used to determine the optimum level of learner support, task complexity, and other task features of the next learning task. It takes the expertise reversal effect into account by selecting only the task(s) adapted to the learner's current expertise.

Component P: Personalization

The model presented in this chapter combines program-controlled and learner-controlled approaches to reach a situation in which the instructional agent and the learner share responsibility over task selection. To which extent one approach or the other approach is emphasized depends on the task characteristics and the learner portfolio, described in component C. This subsection describes both approaches as well as their combination.

First, *program-controlled instruction* has become much easier to realize with the upsurge of computer-assisted instruction, because it is no longer the teacher who has to make decisions for each individual learner. Program-controlled approaches to task selection aim at selecting just-in-time the most suitable learning task from an existing database with tasks, on the basis of information representing what the learner is already able to do or not yet able to do, and/or what the learner already knows or does not yet know. Intelligent Tutoring Systems (ITS), which are adaptive problem-solving environments, have been recognized by artificial intelligence researchers as rich environments that capture some benefits of human tutors. In a typical ITS, a distinction is made between a domain model representing the domain that must be learned, a student model representing what the learner is already able to do or not yet able to do as an "overlay" of the domain model, and an instructional agent which makes decisions on the selection of learning tasks, feedback, presentation modes, and so forth (Corbett, Koedinger, & Hadley, 2001). Originally, ITSs focused on part-task approaches teaching the domain piece by piece. Only with the introduction of dynamic whole-task approaches to teaching over the last decade, it became possible to apply techniques for task selection to learning environments for complex cognitive skills (for a review, see Salden, Paas, & van Merriënboer, 2006).

An example of an electronic learning environment that allows for personalization is the Completion Assignment Constructor (CASCO) described by van Merriënboer and Luursema (1996). It is an ITS for teaching introductory computer programming. CASCO's decisions are based on a student model and made by an instructional agent using rules for the construction and selection of learning tasks. Fuzzy-logic rules are used to construct completion tasks from worked-out examples (ranging from tasks for which the student must add a few lines to a computer program, to tasks for which the student must write nearly the whole program) and to prioritize those learning tasks from 'most suitable' to 'least suitable' for presentation to an individual learner. Van Merriënboer et al. (2002) conducted a study with CASCO in which learners received tasks with no support, support, and *personalized support*. For a transfer test that was performed after the learning tasks, the proportion of correctly used programming concepts was higher for the personalized support group than for the no-support and support groups.

Other studies have found positive results of personalizing the level of difficulty of learning tasks on the basis of performance and mental effort scores (combined in a measure of expertise). In the domain of Air Traffic Control, Camp et al. (2001) and Salden et al. (2004) compared the effectiveness of a fixed easy-to-difficult sequence of learning tasks with dynamic task selection aiming at *personalized difficulty*: the higher the level of expertise of the learner, the more difficult the next learning task. In both studies, personalized task difficulty yielded more efficient transfer test performance than the use of a fixed sequence of easy-to-difficult tasks. Finally, Kalyuga and Sweller (2005) conducted a study in the domain of algebra in which both the level of difficulty and the given support for the next task were adapted to learner expertise. In their study, learners in the group with *personalized support and difficulty* showed higher gains in algebraic skills from pretest to posttest and higher gains in cognitive efficiency than learners in a yoked control group.

Second, besides program-controlled approaches to task selection, there is an increasing tendency to make education student-centered. *Learner-controlled instruction* assumes that learners are able to monitor their own learning and that this will accommodate individual differences. From a student-centered perspective, personalization seeks to take into account the special needs of individual learners (Brna & Cox, 1998). Furthermore, technological advances make it possible to implement types of computer-assisted instruction such as simulations and microworlds, which provide a lot of freedom to learners. These applications allow learners to control diverse aspects of the presented learning tasks (Bell & Kozlowski, 2002).

The provision of learner control has an evident influence on learner's performance (Gray, 1987; Kinzie, Sullivan, & Berdel, 1988), since success in performance depends, amongst other factors, on the learner's level of expertise in the domain and on the learner's self-regulation skills to make proper use of the control provided and to select appropriate aspects of the learning tasks. Novices in a domain or students with poor self-regulation skills lack the ability to make productive use of learner control. They may not select the necessary aspects of learning tasks because they lack adequate knowledge to make educational decisions relevant for learning. Niemiec, Sikorski, and Walberg (1996) conclude that as the level of expertise increases, it is appropriate to decrease program control and increase learner control. In an old study carried out by Fry (1972), low-expertise learners learned significantly less than high-expertise learners in a learner-controlled condition. Gay (1986) also found that in a learner-controlled condition, learners with low prior understanding achieved significantly lower post-test results than learners with high prior understanding.

Personalizing the level of learner control according to the level of expertise of the learner might help learners to practice their self-regulation skills and further develop those skills. According to Merrill (1994), by providing control, learners are brought in a position to "learn better how to learn". Thus, gradually increasing the level of learner control might promote self-regulation skills, that is, the learners' ability to properly select learning tasks according to their abilities, interest and needs. In addition, both actual learner control and even perceived control may positively influence learner motivation. A study carried out by Lahey, Hurlock, and McCann (1973) shows that the perception of learner control has a favorable influence on learners' attitudes. Fry's research (1972) concluded that learners with higher degrees of learner control learned the least but had the most favorable attitudes toward the method of instruction. In a study carried out by Kinzie and Sullivan (1989), 79% of the students in the learner-controlled condition chose to return to the same condition rather than to a program-controlled condition. In contrast, only 19% of the students in the program-controlled condition chose to return to the program-controlled condition. Lahey et al. (1973) also found that learners preferred learner-controlled instruction over program-controlled instruction, although no differences in performance were found. Thus, giving control to learners not always leads to a higher performance but evidently affects motivation in a positive way (Judd, 1972; Lahey, 1976).

Third, the *personalized task-selection model with shared instructional control* presented in this chapter combines program-controlled and learner-controlled approaches to the selection of learning tasks. It aims to cope with the emerging need to personalize authentic and complex learning tasks to each individual learner, preventing cognitive overload by dynamically adapting the level of task difficulty and embedded support. According to this model, an instructional agent first personalizes the characteristics (i.e., difficulty, support, other features) of the presented learning tasks based on learner's expertise (as in Camp et al., 2001; Kalyuga & Sweller, 2005; Salden et al., 2006), and second encourages learners to make a final selection from the subset of tasks pre-selected by the instructional agent. Thus, possible negative effects of too much learner control are reduced by limiting the amount of tasks to choose from, which will still give learners a sense of control without overwhelming them and hampering learning (Kinzie & Sullivan, 1989).

The extent to which either the program or the learner have control over task characteristics (i.e., level of task difficult, embedded support, and task features) mainly depends on the level of expertise of the learner. In other words, personalization (component P) is strongly influenced by the task characteristics and the learner portfolio (component C). On the continuum from program-controlled instruction to learner-controlled instruction, novices who lack the necessary level of knowledge to select optimal learning tasks are placed on the program-controlled side, while more experienced learners are placed on the learner-controlled side. So, as the level of expertise increases, program control will decrease and learner control will increase. However, even highly experienced learners should not always be given full learner control. If the amount of tasks in the learning-task database is very large (e.g., hundreds of tasks), even expert learners may become overwhelmed and demotivated by an excessive amount of freedom (Schwartz, 2004). The learning-task database (component L) is further explained in the next section.

Component L: Learning tasks

Learning tasks differ from each other with regard to difficulty, embedded support and other task features, making it necessary to develop an extensive range of tasks that include and combine those characteristics. The tasks are stored in a learning-task database, which contains tasks with all possible combinations of levels of support and complexity as well as enough variability over other task features to allow for generalization and abstraction by the learner. Learning tasks that are suitable for presentation according to the instructional agent are selected in real-time from this database. Table 2.1 shows an example of how task complexity, embedded support, and task features might be interrelated to each other in the learning-task database.

Table 2.1.
Learning-task Database with the Combination of Different Levels of Complexity and Different Levels of Support

	WOE1 [a]	WOE2 [b]	COMP1 [c]	COMP2 [d]	CONV [e]
Complexity 1	• [f] • •	• • •	• • •	• • •	• • •
Complexity 2	• • •	• • •	• • •	• • •	• • •
Complexity 3	• • •	• • •	• • •	• • •	• • •

[a] WOE1= Worked-out example with process and full product support
[b] WOE2= Worked-out example with full product support
[c] COMP1= Completion task with high product support
[d] COMP2= Completion task with low product support
[e] CONV= Conventional task without support
[f] = Learning task. Each cell contains several (3 in the table) learning tasks with different task features that belong to one complexity level and one level of learner support

Table 2.1 shows three levels of increasing complexity or difficulty (from top to bottom). Within each of the three complexity levels, tasks represent five decreasing levels of learner support (i.e., from worked-out examples to conventional tasks). The faster the increase of expertise of the learner, the faster he or she advances to higher complexity levels, and the lower the support he or she will receive in the next learning task. In addition, for each specific complexity and support level, different tasks are available with different task features (in Table 2.1, for instance, there are three tasks for each combination of difficulty and support). After a subset of tasks has been pre-selected from the learning-task database, the learner makes a final selection of one task to work on.

The model into practice

This section presents an application developed according to our model (see Figure 2.2).

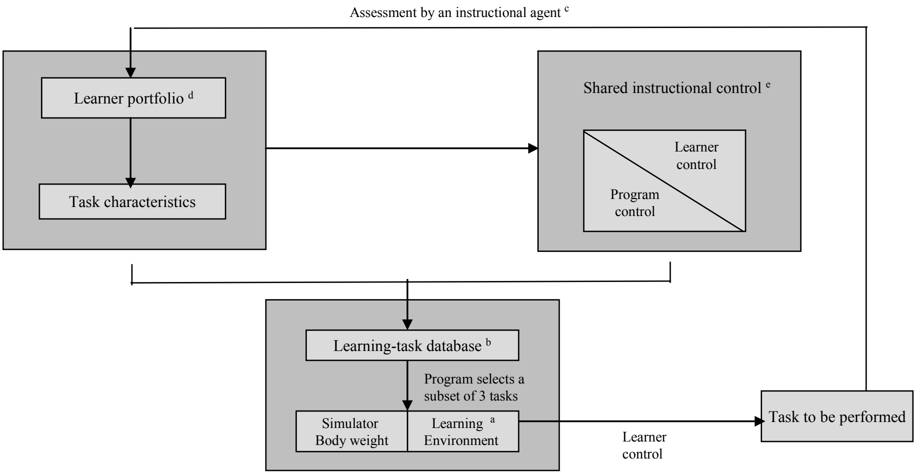


Figure 2.2. Learning environment developed on the basis of the personalized model. The basic components are: (a) an electronic learning environment, (b) a learning-task database, (c) an assessment tool, (d) a learner portfolio, and (e) an instructional agent to realize personalization.

An *electronic learning environment* has been developed enabling learners to work on learning tasks in the domain of dietetics. Learning tasks differ in their level of complexity, level of support, and other task features (i.e., surface features), as described in Component C. The learning environment is a Web application written in the scripting language PHP. Learning tasks are presented in the learning environment, and after each learning task multiple-choice questions and a mental effort rating scale are presented to measure performance and mental effort and to calculate an expertise score. A MySQL database connected to the learning environment contains all learning tasks, multiple-choice questions, a personalization table, and various kinds of logging information.

A simulator called Body-Weight[2] is used by the learners to perform all necessary operations needed to answer the multiple-choice questions presented during the learning tasks. Its main purpose is to help learners understand how body weight is influenced by food intake, physical activity, and several other factors. In the simulator, the students can practice the learning tasks presented in the electronic

[2] The Body Weight simulator was developed based on the calculation model from Dr. K. Westerterp, Rijksuniversiteit Limburg, the Netherlands.

learning environment; it is the interface that confronts them with authentic real-life situations. The learning tasks, in combination with the simulator, allow learners to study the effects on body weight and body composition (e.g., total fat percentage) of alterations in energy intake, physical activity, and other factors such as gender and smoking.

The application contains a *learning-task database* (Component L) with tasks that are representative for the domain of dietetics. The task characteristics as described by Component C, that is, level of complexity, embedded support, and surface task features are combined in this database. Five levels of complexity are distinguished in a simple-to-complex order. Within each complexity level, five levels of embedded support are differentiated. Moreover, for each level of support within each complexity level, three tasks are available with differing surface features that also vary in the real world: different persons with different gender, weight, age, energy intake, specific diet, physical exercise, cultural backgrounds, different habits, and so forth.

An *assessment tool* assesses performance and invested mental effort after each learning task to update the learner portfolio (Component C) and to provide the information for the selection of the next learning task. In order to assess performance, learners must answer multiple-choice questions, based on predefined performance objectives, after each task. Moreover, the assessment tool monitors learner's actions on the simulator in order to assess the accuracy of the problem-solving process. Such measures make possible the assessment of the learner's performance in real-time after each task is performed, an important requisite for a program that dynamically adapts learning tasks. The results of both the multiple-choice questions and the accuracy of the process are combined in one measure of performance. A 7-point subjective rating scale is used to assess the learner's investment of mental effort in the learning task, with values ranging from extremely low (1) to extremely high (7). The scale is adapted from the original mental effort scale developed by Paas (1993). Performance and mental effort scores are updated in the learner portfolio after each learning task.

The *learner portfolio* contains the information collected by the assessment tool. The portfolio enables the selection of a subset of suitable learning tasks from which the learner can make a final selection. It is updated after each learning task by combining new performance and invested mental effort scores with previous scores (i.e., the learner's history). It is important to take the learning history into account because this flattens out undesirable effects due to unstable measures, for instance, a high performance measure that is not due to acquired knowledge and skills but to good luck. The new scores in the learner portfolio are used by the instructional agent to personalize the selection of the next learning task(s) in such a way that it best suits the individual learner at a particular moment in time.

Finally, the *personalization* (Component P) that is realized by the instructional agent is twofold: (a) the updated learner portfolio is used to select a subset of suitable learning tasks from the learning-task database, and (b) the selected tasks are presented to the learner who makes the final selection. The program employs a simple task-selection algorithm that decides on the “jump size” from one complexity level to another complexity level, and from one level of support to another level of support, on the basis of a combination of performance and mental effort scores. The algorithm is presented in Table 2.2.

Table 2.2
Task-selection Algorithm Indicating Jump Sizes Between Learning Tasks

Mental effort	Performance						
	1	2	3	4	5	6	7
1	0	0	0	+3	+4	+5	+6
2	0	0	0	+2	+3	+4	+5
3	-2	-1	0	+1	+2	+3	+4
4	-3	-2	-1	0	+1	+2	+3
5	-4	-3	-2	-1	0	+1	+2
6	-5	-4	-3	-2	0	0	0
7	-6	-5	-4	-3	0	0	0

As can be seen from Table 2.2, the jump size is determined from the combination of performance and mental effort scores. Basically, mental effort scores (ME) are subtracted from performance (P) scores to compute the jump size. Thus, the higher the performance and the lower the invested mental effort, the larger the positive jump size. For instance, a score of 5 on performance and 2 on mental effort yields a jump size of +3 (5 – 2 = 3), meaning that the level of support decreases 3 levels, or, if there are less than three support levels available at the current complexity level, the learner will move to the lowest support level because the learner is only allowed to progress to a next complexity level after successful completion of a conventional, non-supported task. Accordingly, the lower the performance and the higher the invested mental effort, the larger the negative jump size. For instance, a score of 2 on performance and 5 on mental effort yields a backward jump of 3 steps (2 – 5 = -3), meaning that either the level of support increases 3 levels, or, if there are less than three support levels available at the current complexity level, the learner will move back to the highest level of support at the current complexity level. The algorithm applies some additional rules to make the instruction encouraging and motivating for the learners. If the performance score is 5 or higher or the mental effort score is 2 or lower, the learner will not jump backwards. And if the performance score is 3 or lower or the mental effort score is 6 or higher, the learner will not jump forward.

Discussion and conclusions

This chapter discussed a personalized task-selection model with shared instructional control, combining program-controlled and learner-controlled approaches to the selection of learning tasks. It has been found that personalization of learning tasks yields more efficient and more effective learning than a fixed sequence of learning tasks that is identical for all learners. In addition, learner control is believed to make learning more appealing and to encourage the development of learners' self-regulation skills. The proposed model combines the strong points of both approaches and was therefore expected to make learning more effective (i.e., higher transfer test performance), more efficient (i.e., a more favorable ratio between performance and time on task or mental effort), and more appealing (i.e., the learner will show more interest in the learning tasks). A higher motivation will also positively influence the amount of mental effort learners invest in learning and their willingness to become engaged in additional instructional activities. If learners attribute the 'investment of more mental effort' to 'better outcomes' they will perceive a higher self-efficacy while implementing the required actions to perform such tasks, which in turn will positively affect motivation (Bandura, 1997; Keller, 1983b; Kinzie, Sullivan, & Berdel, 1988).

A pilot study has been conducted in order to examine whether personalized selection of learning tasks with shared instructional control led to better results than personalized instruction with full program control. Twenty-five nursing students (6 males and 19 females) from a school for senior vocational education in Eindhoven (the Netherlands) participated in this pilot study. Their average age was 18.2 years (SD = 3.66). Twelve and thirteen learners were randomly allocated to the program-controlled condition and the shared-controlled condition, respectively. Results on both performance scores and invested mental effort were in the expected direction (see Table 2.3). Participants in the shared-controlled condition achieved a higher mean performance score (i.e., Cohen's d = 0.25 which indicates a small effect size) and a lower mean mental effort score (d = 0.37 which indicates a small effect size) than participants in the program-controlled condition. Instructional efficiency (see Paas & van Merriënboer, 1993; Paas et al., 2003) was computed on the basis of performance and invested mental effort. When the performance is higher than might be expected on the basis of invested mental effort (P > ME) the instruction is relatively efficient, and when the performance is lower than might be expected on the basis of invested mental effort (P < ME) the instruction is relatively inefficient. Results on instructional efficiency are in the expected direction, with a higher mean efficiency score for the shared-controlled condition (d = 0.36 which indicates a small effect size; see Table 2.3).

Table 2.3
Overview of Test Results and Interest

	Method of personalized training			
	Shared instructional control $n = 13$		Full program control $n = 12$	
	M	*SD*	*M*	*SD*
Mental effort	3.70	1.08	4.11	1.15
Performance	59.43	13.70	55.52	17.87
Mental Efficiency	0.21	1.12	-0.23	1.34
Interest in training*	2.77	0.95	2.05	0.88

*$p < .10$

Motivation was measured with items from the interest/enjoyment subscale of the Intrinsic Motivation Inventory (IMI; Deci, Eghrari, Patrick, & Leone, 1994), which were translated into Dutch by Martens and Kirschner (2004). On a questionnaire with 7-point scales filled out after the course, learners in the shared instructional control condition reported a slightly significant higher interest in the learning tasks than learners in the full program-control condition ($p < .10$; $d = 0.79$ which indicates a medium effect size; see Table 2.3). The results of the pilot study provide some preliminary evidence that the personalized model is a promising approach to increase the flexibility and quality of educational programs. Moreover, they support the idea that the provision of (limited) learner control may enhance learners' motivation.

Some final comments have to be made. With regard to the assessment of complex performance, the PPS requires the measurement of many qualitatively different aspects with different measurement instruments. Our application only employed multiple-choice questions and accuracy scores to assess performance, and a rating scale to assess the mental effort invested to reach this performance. Future studies need to assess qualitatively different aspects (e.g., problem-solving aspects, routine aspects, attitudinal aspects) and use different measurement instruments to develop a learner portfolio with valid and reliable information to base the selection of learning tasks on.

With regard to the level of learner control, future studies may progressively increase learner control as the level of learners' expertise increases. For instance, first giving novices only control over selecting task features with a given support and complexity (as in the presented pilot study), then giving them, as expertise increases, control over task features and the amount of learner support, and finally giving them control over task features, support, as well as task complexity. This way, the complex relationships between task features, support, and complexity are considered together with the level of expertise. If too much control is given to novices in the learning domain, or if too little control is given to experienced learners, this may hamper their learning, performance and motivation. However, the question still remains how to combine learner and task characteristics most effectively, and therefore how to balance program and learner control (component P in Figure 2.1) for each learner. Giving novices full program control may demotivate them, and giving experienced learners full learner control may overwhelm them if the amount of tasks to choose from is very large. As Schwartz (2004) argues, " . . . at this point, choice no longer liberates, but debilitates" (p. 2).

Another relevant issue is raised by Bell and Kozlowski (2002), who found in their study positive effects on the nature of the learners' study and practice, self-regulation, knowledge acquired, and performance when students with learner control received some form of *advice*. Specifically, meta-cognitive advisory models explicitly help students to apply cognitive strategies for assessing their own performance, and to select new learning tasks based on this assessment, which may enhance their cognitive strategies for regulating their own learning (Kicken, Brand-Gruwel, & van Merriënboer, 2005). Thus, future studies may integrate shared instructional control with advisory models that help learners to select appropriate tasks.

To conclude, the first results of an application built according to our personalized model with shared instructional control are promising. However, this is only a first step towards more flexible, demand-driven educational programs. Future studies also need to personalize the amount of learner control to the learner's level of expertise, and include advisory models to assist learners with their decisions until they eventually become self-regulated high-ability learners.

3

Selecting Learning Tasks: Effects of Adaptation and Shared Control on Learning Efficiency and Task Involvement[3]

Complex skill acquisition by performing authentic learning tasks is constrained by limited working memory capacity (Baddeley, 1992). To prevent cognitive overload, task difficulty and support of each newly selected learning task can be adapted to the learner's competence level and perceived task load, either by some external agent, the learner herself, or both. Health sciences students (*N* = 55) participated in a study using a 2 x 2 factorial design with the factors adaptation (present or absent) and control over task selection (program control or shared control). As hypothesized, adaptation led to more efficient learning; that is, higher learning outcomes combined with less effort invested in performing the learning tasks. Shared control over task selection led to higher task involvement; that is, higher learning outcomes combined with more effort directly invested in learning. Adaptation also produced greater task involvement.

---

[3] This chapter is based on: Corbalan, G., Kester, L., & van Merriënboer, J.J.G. (in press). Selecting learning tasks: Effects of adaptation and shared control on learning efficiency and task involvement. *Contemporary Educational Psychology*.

In addition to incorporating authentic learning tasks, which are based on real-life tasks, modern educational approaches often aim at adapting a sequence of learning tasks to the needs of each individual learner (Kalyuga & Sweller, 2005; Renkl & Atkinson, 2003). Rather than one curriculum for all learners, such approaches allow each learner to have her own curriculum. This study addresses the questions how adaptation of task selection can be realized, and which agent, that is, human tutor, computer program or learner should be responsible for it.

Especially for novice learners, the acquisition of complex skills by performing authentic or real-life tasks is heavily constrained by the limited processing capacity of working memory because such tasks easily cause cognitive overload (Baddeley, 1992; Sweller, 1988). Within the framework of cognitive load theory, three types of cognitive load are identified: intrinsic, extraneous and germane load (Sweller, 1988; Sweller, van Merriënboer, & Paas, 1998; van Merriënboer & Sweller, 2005). Intrinsic load is inherent to a learning task and depends on the number of interacting elements that have to be related, controlled, and kept active in working memory during task performance. For example, learning vocabulary and speaking in a foreign language cause low and high intrinsic load, respectively. Extraneous and germane load are the result of the instructional design. Extraneous load is ineffective load due to poorly designed instructional material, resulting, for example, from the need to combine information from different sources to complete a learning task. Germane load occurs when load is imposed by processes that are directly beneficial for learning. For instance, a high variability in a set of learning tasks may stimulate learners to construct more integrated cognitive schemata. In the current study, task load is seen as a combination of intrinsic and extraneous load, that is, all load caused by *performing* the task but not directly caused by learning processes. In contrast, germane load is caused by *learning* to perform the task.

Cognitive load can be measured in several ways. Paas, Tuovinen, Tabbers, and van Gerven (2003) describe three different techniques including subjective measures, secondary-task methods, and psychophysiological measures. With regard to subjective measures, cognitive load researchers have commonly measured cognitive load as perceived mental effort (Paas et al., 2003). Self-ratings of invested mental effort have been widely used and quickly accepted amongst cognitive load researchers because they are unintrusive, reliable, relatively easy to use and analyze, and provide a good indication of the overall cognitive load a task imposed (Paas & van Merriënboer, 1994; Paas et al., 2003). Possible limitations such as socially desirable ratings and the tendency to keep answering in the same manner (Richman, Kiesler, Weisband, & Drasgow, 1999) are usually overcome through randomized or counterbalanced design.

In order to enable the use of authentic tasks in education, and yet prevent overloading the learners' cognitive system, task characteristics should be adapted to the individual needs of learners (Corno & Snow, 1986; Kalyuga & Sweller, 2005; Park & Lee, 2003; Salden, Paas, Broers, & van Merriënboer, 2004; van Merriënboer & Luursema, 1996; van Merriënboer, Schuurman, de Croock, & Paas, 2002). Two approaches to individualization that aim to cope with increasingly complex learning situations are program control, in which a computer program is responsible for the process of selecting new learning tasks, and learner control, in which the learner controls the selection of tasks (Corbalan, Kester, & van Merriënboer, 2006).

The issue of the locus of instructional control, whether it is external (i.e., program control) or internal (i.e., learner control), has been a primary concern in the upsurge of computer-assisted instruction (Lawless & Brown, 1997; Tennyson & Buttery, 1980). First, according to models of *program-controlled instruction*, learning is influenced by the characteristics of the learners (Zimmerman, 2002) and the learning tasks (Lawless & Brown, 1997). Intelligent Tutoring Systems (ITS) are adaptive problem-solving environments which make a distinction between: (a) a domain model representing the domain that must be learned, (b) a student model representing what the learner is already able to do or not yet able to do, and (c) a computer program which makes instructional decisions in response to input from the learner (Corbett, Koedinger, & Hadley, 2001; Park & Lee, 2003). Using ITS may reduce both training time and costs, which is particularly interesting for domains in which these aspects are of great importance, as in aviation and industry (Camp, Paas, Rikers, & van Merriënboer, 2001). In addition, technology-based instruction has the potential to become an important resource to improve learning in K-12 classrooms. However, still few ITSs have become successfully implemented products to enhance learning, especially in K-12 settings (Wong, Chan, Chou, Heh, & Tung, 2003), probably because they are difficult and costly to develop, although attempts are being made to address these challenges (e.g., Beal, 2004).

Second, besides program-controlled instruction, there is an increasing emphasis on providing learners with control over their own learning path. *Learner-controlled instruction* assumes that learners are able to monitor their own learning processes and that this will accommodate individual differences. Giving learners some control over instructional aspects creates the necessary preconditions for practicing self-regulation skills and is a first step towards teaching those skills (Kinzie, 1990). Technological advances make it possible to implement types of computer-assisted instruction such as simulations and microworlds, which are multimedia learning tools that provide users with dynamic elements that are under their control. Whereas simulations are more aligned with traditional instructional uses of educational software and allow learners to run experiments, microworlds take a more constructivist approach and allow learners, for instance, to design their own experiments (Rieber, 2005). Both simulations and microworlds provide learners with considerable freedom in choosing aspects of learning such as the content, the sequence, and the pace of the instruction (Bell & Kozlowski, 2002).

This study focuses on adaptive task selection with shared control, in which a computer program and the learner share control over the selection of learning tasks. In this two-step process, the computer program first selects a subset of learning tasks with characteristics (difficulty, available support) that are adapted to the needs of the individual learner (program control). Second, the learner selects from this subset one task to work on (learner control). We hypothesize that adaptive task selection with shared control will have a positive influence on both learning efficiency, due to the task adaptation made by the computer program, and task-involvement, due to the fact that the learner feels to be in control over her own learning. In this introduction, the advantages and disadvantages of program control and learner control are described first. Then, adaptive task selection with shared control is discussed.

Adaptive Task Selection with Program Control

Good instruction accommodates relevant individual differences among learners (Shute & Towle, 2003). A predefined and fixed sequence of learning tasks cannot take the different levels of competence, misconceptions, interests, and learning styles of a heterogeneous group of learners into account. But with adaptive task selection with program control, the program may dynamically adapt the task characteristics of each newly selected learning task to the characteristics of the individual learner.

With regard to learner characteristics, measuring a learner's competence and cognitive load is essential for the adaptive selection of new learning tasks. First, competence refers to the combination of knowledge, skills, and attitudes that allows for the performance of real-life tasks (Baartman, Bastiaens, Kirschner, & van der Leuten, 2007; van Merriënboer, 1997). Eraut (1994) stresses that skills cannot be separated from knowledge, as this would exclude the practical know-how to perform real-life operations. Hence, the assessment of competence requires a combination of assessment methods (Baartman et al., 2007). Second, assessing the cognitive load imposed by the performance of a certain task may provide additional insight into the learner's needs. For example, if after one task two learners shows the same competence level but one learner reports higher task load than the other, the first learner is best supported by presenting her with a new learning task which is easier or provides more support than is required by the second learner. The combination of performance and task load has been proposed by Paas and van Merriënboer (1993) as a reliable estimate of the relative *efficiency* of learning. According to this approach, efficiency is high if performance is higher than might be expected on the basis of the invested mental effort required to perform the task. Conversely, efficiency is low if performance is lower than might be expected on the basis of the invested mental effort to perform the task.

With regard to task characteristics, in a well-designed curriculum learning tasks are ordered from easy to difficult, and learner support decreases as the learners' competence increases (van Merriënboer, 1997; van Merriënboer, Clark, & de Croock, 2002; van Merriënboer & Kirschner, 2007). Accordingly, in this study the *difficulty* and *support* of selected learning tasks are adapted to the characteristics of each individual learner (i.e., level of competence and perceived task load). In terms of cognitive load theory, the difficulty of a task determines the intrinsic cognitive load, which is a direct result of the complex nature of the learning material. Tasks should be selected that are neither too difficult nor too easy for the learner. Two studies in the Air Traffic Control domain which adapted the level of difficulty (Camp et al., 2001; Salden et al., 2004) showed that adaptive task selection based on competence and cognitive load yielded better learning outcomes than a fixed task sequence.

Furthermore, the amount of embedded support may determine the extraneous load. When novices start working on a range of complex tasks, it is essential to provide them with support, which gradually diminishes in a process of 'scaffolding' as their competence increases. The 'completion strategy' (van Merriënboer, 1997, van Merriënboer & Kirschner, 2007) is a powerful approach to scaffolding. In this approach, tasks with a particular level of difficulty are organized from worked-out examples, via completion tasks, to conventional tasks. First, worked-out examples confront learners not only with a description of a given state and the criterion for an acceptable goal state, but also with a description of all solution steps. Then, completion tasks provide learners not with all solution steps but with a partial solution that must be completed by them. Finally, conventional tasks provide learners with a given state and a criterion for an acceptable goal state only: learners must independently generate the whole solution. Experienced learners have relevant knowledge that enables them to approach a conventional task. However, when novice learners in a domain are confronted with conventional tasks, they use cognitively demanding strategies such as means-ends analysis and working backward to reach a solution, increasing extraneous cognitive load because those strategies are not efficient ways to learn (Renkl, Stark, Gruber, & Mandl, 1998; Sweller, 1988).

The 'expertise reversal effect' (Kalyuga, Ayres, Chandler, & Sweller, 2003) states that successful instructional techniques for novice learners (e.g., presenting tasks with high embedded support) often lose their value when used with more experienced learners. For instance, presenting a diagram with integrated textual explanations may be an effective technique for novice learners who need the explanations to understand the diagram. However, experienced learners who already possess the knowledge necessary to understand the diagram must invest cognitive resources unnecessarily before they can determine that the information is redundant. Such redundant cognitive processing constitutes extraneous cognitive load, which may hinder learning. In a study carried out by Kalyuga and Sweller (2005), both the level of difficulty and the level of support were adapted to learner's competence and cognitive load ratings. Learners in the adaptive group showed higher gains in algebraic skills from pretest to posttest, and higher gains in efficiency compared to learners in a control group. The research reported in this article also adapts the level of difficulty and the level of support to the individual learner's competence level and perceived task load. However, we will use another measure of competence and another selection table for choosing tasks, and we provide learners with some control over the – final – selection of learning tasks.

Although adaptive task selection with program control may have positive effects on learning efficiency (i.e., higher learning outcomes combined with lower task load, that is, less effort invested in performing the tasks), it also has clear limitations. Program control over task selection leaves learners with no freedom of choice, which may negatively affect their motivation, specifically their task involvement and interest. One way to overcome this problem is to give learners some control over the selection of learning tasks, which has positive effects on motivation (Kinzie & Sullivan, 1989; Ross, Morrison, & O'Dell, 1989; Schnackenberg & Sullivan, 2000).

Task Selection with Learner Control

Recent instructional theories advocate on-demand methods of education in which learners are given freedom over their own learning path. This is in line with the presented study, in which learner control explicitly refers to control over *task selection*. Merrill (1994) suggested that by providing control, learners will acquire more effective ways of learning and become better equipped to adapt to diverse situations. Learner control is typically perceived as something which will enhance motivation, and consequently may increase learning outcomes (Reeves, 1993). Motivated learners engage in learning activities and allocate cognitive resources to learning because they derive satisfaction from performing the task (Deci, Vallerand, Pelletier, & Ryan, 1991). The invested mental effort and its associated learning outcomes have been recognized as an indicator of the learners' *involvement* in a task (Paas, Tuovinen, van Merriënboer, & Darabi, 2005). Accordingly, learner control, amongst other elements, is considered as a determinant of intrinsic motivation to learn (Deci & Ryan, 1985). When intrinsically motivated, learners engage in activities out of *interest* (Deci et al., 1991).

First, with regard to the learner's task involvement, the effort invested in learning processes is a direct indicator of motivation (Keller, 1983b). Thus, learner control, amongst other factors, may increase learners' task involvement. Learners involved in learning are more inclined to be engaged in learning processes such as exploration, abstraction, and generalization (van Merriënboer, 1997), and will invest more mental effort in the construction of cognitive schemata (Keller, 1983b; Paas et al., 2005). This enhanced engagement may positively influence learning outcomes (Greene & Miller, 1996). If learners attribute 'achieving better outcomes' to 'higher mental effort invested', they will perceive a higher self-efficacy in implementing the required actions to perform such tasks, affecting motivation positively (Bandura, 1997; Keller, 1983b; Kinzie, Sullivan, & Berdel, 1988; Zimmerman, 2000).

Until now, cognitive load theorists have typically focused on comparing instructional formats in terms of their efficiency. However, the importance of motivation for learning has not been sufficiently explored. According to Paas et al. (2005), mental effort and performance have both cognitive and motivational components. Consistent with the efficiency approach, Paas et al. (2005) proposed a complementary approach to calculate the relative task involvement in instructional conditions. According to this approach, the higher the learner's task involvement, the higher the mental effort directly invested in learning (i.e., germane load), which is likely to enhance learning outcomes.

Second, the learner's level of interest is another important motivational factor (Keller, 1983b). When the learning environment gives learners the opportunity to explore, interest is more likely to be retained. Educational researchers have identified two types of interest, namely, personal and situational interest (Alexander & Jetton, 1996; Hidi, 2001, 2006; Hidi & Renninger, 2006). Personal interest develops slowly over time, is internally oriented, and of enduring personal value. Situational interest is transitory, external, and environmentally triggered. Personal interest most likely results from repeated experiences and develops over time, whereas situational interest can be increased by, for instance, emphasizing learners' choices (Schraw, Flowerday, & Lehman, 2001) as is intended in the current study. Accordingly, learner control is expected to make learning more interesting (Fry, 1972; Kinzie & Sullivan, 1989; Lahey, Hurlock, & McCann, 1973), with potential benefits for learning outcomes (Wolters, 2003; Zimmerman, 2002). A study in the math domain carried out by Cordova and Lepper (1996) revealed that participants in the choice conditions (i.e., participants who were given control over irrelevant aspects of the tasks) reported liking the program significantly more and scored significantly higher than those in the no-choice conditions.

Despite the apparent beneficial effects of learner control on learning, novices generally lack the necessary knowledge to make effective educational decisions and may omit essential aspects of learning (Merrill, 2002). It is also apparent that learner control may introduce potential problems with cognitive load. Even highly experienced task performers with full control may become overwhelmed by a high amount of choice, hampering their learning (Iyengar & Lepper, 2000; Schwartz, 2004). In addition, Niemiec, Sikorski, and Walberg (1996) argue that as the level of experience increases, it is appropriate to decrease program control and increase learner control. In this respect, several studies (Fry, 1972; Gay, 1986) showed that with learner control, learners with low prior knowledge learned significantly less than learners with high prior knowledge. Hence, giving (perceived) control to learners does not always lead to higher performance, but can positively affect motivation (Fry 1972; Judd, 1972; Lahey, 1976; Lahey et al., 1973). Accordingly, we discuss an alternative approach that combines the benefits of program control and learner control over task selection, namely, adaptive task selection with shared control.

Adaptive Task Selection with Shared Control

Both program control and learner control over task selection may have beneficial effects on learning. However, a high level of program control may negatively affect learners' task involvement and interest. A high level of learner control may overwhelm even expert learners if the number of tasks to choose from is (too) large and learners may omit essential aspects of learning. The present study combines program and learner control into a task-selection approach with shared control. In this two-step approach, the *program* first uses a measure of the individual learner's competence level and task load ratings to select from an existing database with learning tasks a subset of tasks with an optimal level of *difficulty* and an optimal level of *support*. All the tasks in the selected subset have the same difficulty and the same level of support. They only differ in *surface* features, that is, aspects of the task that are not related to goal attainment, such as the color of the eyes of a person in dietetic problems. In the second step, the learner makes the final selection of one task from the pre-selected subset of tasks. Thus, the learner may select one task with the surface features she prefers.

As an illustration of this two-step process, take a learner who has solved a dietetic task with a particular level of difficulty (e.g., difficulty level 4 of a range of 5 levels) and a particular level of support (e.g., highest support level 1 of a range of 4 levels). The learner performs reasonably well (e.g., 7 out of 10 points) and reports a moderately low task load (e.g., 3 out of 7 points). According to predefined task-selection rules, the program now first presents the learner three tasks which have the same difficulty (e.g., again at difficulty level 4) and the same level of support (e.g., now *without* support, that is, support level 4 rather than support level 1 as for the previous task). The three pre-selected tasks in the subset have varying surface task features, such as the subject's age and habits. In the second step of the task-selection process, the learner then selects the task with her preferred surface features from the pre-selected subset.

We hypothesize that adapting the difficulty and the embedded support of the learning tasks to the level of competence and invested load of the learner will make learning more effective (i.e., higher learning outcomes) and more efficient (i.e., higher learning outcomes combined with less effort invested in performing the learning tasks). Moreover, it is hypothesized that the greater perceived control offered by shared control will have positive effects on learners' motivation, depicted by increased task involvement, that is, higher germane load combined with higher learning outcomes. This in turn is also expected to positively affect learner's interest in the learning tasks and the training program as a whole.

Method

Participants

Fifty-five first year students (13 males and 42 females; mean age = 22.40 years, *SD* = 7.27) in Dutch Vocational Education and Training (VET) in the Health Sciences domain participated in the experiment. Learners had no prior knowledge in dietetics which was the learning domain used in the experiment. All participants were entered into a lottery making them eligible to win one of 18 music compact disks. Participants were randomly assigned to one of the four experimental conditions in a 2 x 2 factorial design: adaptation with shared control (*n* = 15); non-adaptation with shared control (*n* = 12); adaptation with program control (n = 15), and non-adaptation with program control (n = 13).

Materials and Measurements

Training Phase

Learning-task database. The learning-task database contained tasks of five difficulty levels in the dietetics domain. Each difficulty level comprised five levels of embedded support. Each level of support contained three different tasks with varying irrelevant surface features.

There were five *levels of difficulty* (1 to 5), defined in cooperation with three domain experts. Each subsequent, more difficult level included a new element or a combination of new elements increasing the difficulty of the task. In level 1, participants used their *own* data to calculate changes in their body weight over time, taking energy intake and energy expenditure as input variables. This makes the task personally relevant and helps participants to get acquainted with the material. In level 2, participants must identify the changes in body weight of a *specific person* over time, based on energy intake and energy expenditure. Subject characteristics were predetermined in the learning task. This required learners to transfer the learned procedures to an unfamiliar situation. In level 3, participants must investigate the differences in body weight and the factors influencing *fat percentage* (a new element) between a man and a woman of the same age, height, weight and pattern of activities. In level 4, participants must simulate three *different strategies* for the treatment of obesity which require more solution steps, and conclude which strategy is most appropriate. In addition, participants must infer what happens to body weight when a person on a diet returns to her original habits (a new factor - body adaptation to the new situation after following a diet - plays a role). In level 5, participants must study the effects of smoking on the body weight of a given person. This implies taking another additional input variable, the increase of the *basal metabolism rate,* into account. Moreover, learners must simulate the effects on body weight when the same person stops smoking, which decreases the basal metabolism rate. This requires learners to simulate the same person in the simulator taking body changes as well as the decrease of basal metabolism rate into account.

There were also five *support levels,* differing with regard to the amount of embedded support, and diminishing in a process of ’scaffolding’ according to the completion strategy described earlier (van Merriënboer, 1997). The five levels, ordered from high to low support, are: (1) worked-out examples that included both full product support (i.e., all the solution steps and the ‘expert’ solution are given) and process support (i.e., the ‘why’ or the rationale behind the solution steps is given), (2) worked-out examples or learning tasks that provided full product support but no process support, (3) completion problems with high support or learning tasks that provided many but not all solution steps, (4) completion problems with low support or learning tasks that provide a few solution steps, and (5) conventional problems or learning tasks that did not provide any support.

Within each level of difficulty (except for difficulty level 1, in which learners used their own data), three different tasks with different *surface features* that did not influence the difficulty or support levels (i.e., different persons with different characteristics, such as age, habits, appearance and background) were included. Figure 3.1 shows a learning-task database that combines different levels of difficulty, five levels of embedded support, and three task features per support and difficulty level.

	Task support levels				
	WOE1 a	WOE2b	COMP1c	COMP2 d	CONVe
Difficulty 1	Task 1 f Task 2 Task 3	Task 4 Task 5 Task 6	Task 7 Task 8 Task 9	Task 10 Task 11 Task 12	Task 13 Task 14 Task 15
Difficulty 2	Task 16 Task 17 Task 18	Task 19 Task 20 Task 21	Task 22 Task 23 Task 24	Task 25 Task 26 Task 27	Task 28 Task 29 Task 30
Difficulty n	Task n Task n Task n	Task n Task n Task n	Task n Task n Task n	Task n Task n Task n	Task n Task n Task n

Mr. Brown
-English painter
-36 years old
-84 kilos
-Swims 3 hours per week

Mrs. Van Hout
-Dutch teacher
-51 years old
-67 kilos
-Plays golf 2 hours every Sunday

a WOE1= Worked-out example with full product and process support
b WOE2= Worked-out example with full product support
c COMP1= Completion task with high product support
d COMP2= Completion task with low product support
e CONV= Conventional task without support
f = Learning task. Each cell contains several (3 in the Table) learning tasks with different task features that belong to one difficulty level and one level of learner support

Figure 3.1. Learning-task database with the combination of different levels of difficulty, different levels of support, and different task features.

Electronic learning environment. The learning environment was a Web application written in the popular web scripting language PHP and especially developed for the current study. A MySQL database connected to the learning environment contained all learning tasks, registered competence and cognitive load measures, a selection table for making a pre-selection of tasks, and various kinds of logging information. In the electronic learning environment participants were presented with (a) a Web application in which the learning tasks in the domain of dietetics were presented (see Figure 3.2a), and (b) a simulator called “Body Weight” (see Figure 3.2b), which allowed learners to retrieve and process the necessary data to perform the presented learning tasks. The learning environment was a Web application connected to the learning-task database and contained the following instruments to gather information on learner behaviour: (a) practice multiple-choice questions, (b) performance measures of operating the simulator (i.e., whether learners use the relevant windows in the simulator to reach the solution, such as the ‘eating meter’ or the ‘physical activity meter’ to calculate the amount of calories gathered by energy intake or burned by energy expenditure, respectively), and (c) cognitive load measures for task load and germane load. In the Body Weight simulator, participants could look up the energy in kilojoules of a specific drink or type of food, estimate the energy expenditure of a person, or simulate changes in a person’s body weight and body composition using energy intake, energy expenditure, and other parameters as input variables.

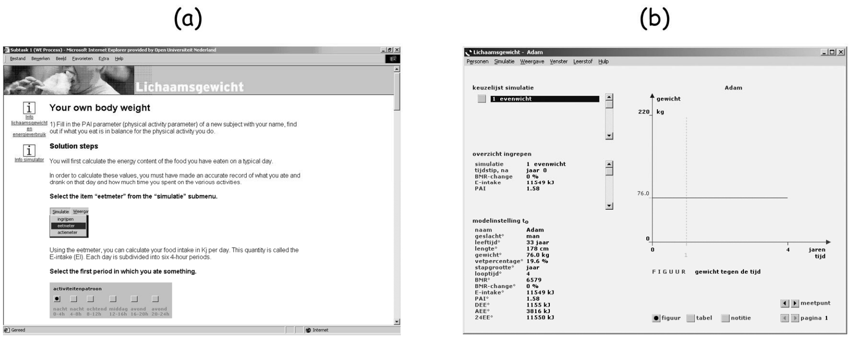


Figure 3.2. The learning environment, showing part of a learning task (a) and the simulator used to perform the task (b).

In the adaptive/program-control condition, the level of difficulty and the level of support of selected tasks were based on the learner’s competence and task-load scores, and one task with randomly selected surface features was presented to the learner. In the adaptive/shared-control condition, the level of difficulty and the level of support of selected tasks were again based on the learner’s competence and task-load scores, but now three tasks with different surface features were presented to the learner, so that the learner could make a final selection from these three tasks. In the non-adaptive/program-control condition, each learner was paired (i.e., *yoked*) to one learner in the adaptive/program-control condition and received exactly the same sequence of tasks as his or her yoked counterpart. In the non-adaptive/shared-control condition, each learner was paired to a learner in the adaptive/shared-control condition and received the same subset of tasks as his or her yoked counterpart.

Competence (C). After each learning task, participants received six multiple-choice questions (with three answering options) to measure acquired knowledge. Each correct answer scored 100 points/6 questions = 16.67 points. Furthermore, an assessment tool monitored the relevant windows opened in the simulator to assess the accuracy on actual performance. This was calculated by counting the number of opened windows proportional to the number of windows that had to be opened to correctly complete the learning task. Scores could range from 0% of correct windows opened (0 points) to 100% of correct windows opened (100 points). For instance, if four windows had to be opened to correctly complete the task and the learner opened only two of them, the score would be 100/4 = 25 * 2 = 50% = 50 points. Competence was measured with the formula ((60 * score on multiple-choice questions) + (40 * score on correct windows opened)/100), leading to a minimum score of 0 and a maximum score of 100. This measure allows for real-time assessment of a learner's competence, weighting knowledge measures and actual performance. The weight of the knowledge measure is somewhat higher than the weight of the actual performance, because knowledge is seen as a prerequisite for the ability to open the correct windows.

Task load (L). After each learning task, task load was measured with a one-item 7-point rating scale as the 'effort required to perform the task', ranging from a very small amount of effort (1) to a very high amount of effort (7). The internal consistency of the test was .93 (Cronbach's alpha). The task load is used to make task-selection decisions in the adaptive conditions and to compute the learning efficiency as described in a later section.

Selection table. In the two adaptive conditions, the MySQL database connected to the learning environment contained a selection table (see Table 3.1). The selection table indicated the 'jump size' or progression from one level of support to another level of support, and from one difficulty level to another difficulty level. Competence and task load scores are used as a learner variable for dynamic task selection. This approach has also been successfully used in other studies (Camp et al., 2001; Salden et al., 2004). To correct for extreme values, the mean of the competence measure on the last learning task and the previous learning task was computed with a higher weight for the last learning task (70%) than for the previous learning task (30%), leading to a minimum score of 0 and a maximum of 100.

Table 3.1
Selection Table Indicating Jump Sizes Between Learning Tasks

Task Load	Competence 1	2	3	4	5	6	7
1	0	0 [c]	0 [c]	+3	+4	+5	+6
2	0 [b]	0	0 [c]	+2	+3	+4	+5
3	-2	-1	0	+1	+2	+3	+4
4	-3	-2	-1	0	+1	+2	+3
5	-4	-3	-2	-1	0	+1	+2
6	-5	-4	-3	-2	0 [a]	0	0 [d]
7	-6	-5	-4	-3	0 [a]	0 [a]	0

[a]Adjusted jump size = 0 because the computed jump size is negative and the competence score is 5 or higher (rule a)
[b]Adjusted jump size = 0 because the computed jump size is negative and the task-load score is 2 or lower (rule b)
[c]Adjusted jump size = 0 because the computed jump size is positive and the competence score is 3 or lower (rule c)
[d]Adjusted jump size = 0 because the computed jump size is positive and the task-load score is 6 or higher (rule d)

To compute the jump size (J), task load scores (TL) are subtracted from competence scores (C). The higher the competence score and the lower the task load, the larger the positive jump size. For instance, a score of 5 on competence and 2 on task load yields a jump size of +3 (i.e., 5 – 2 = 3), meaning that the level of support *decreases* three levels (e.g., from a worked-out example with product support to a conventional problem). If there are less than three support levels available at the current difficulty level, the learner will move to the lowest support level (i.e., a conventional task) because the learner is only allowed to progress to the next difficulty level after successful completion of a conventional task (i.e., a task without embedded support). That is, only once the learner has successfully solved a conventional task at a particular difficulty level, s/he is considered to master the required competence level and is allowed to proceed to the next, higher difficulty level. Accordingly, the lower the competence level and the higher the task load, the larger the negative jump size. For instance, a score of 2 on competence and 5 on task load yields a backward jump of 3 steps (i.e., 2 – 5 = -3), meaning that the level of support *increases* three levels. But, if there are less than three support levels available at the current difficulty level, the learner will move back to the highest level of support (i.e., a worked example with process and product support) at the current difficulty level.

The selection table also applies some additional rules: if the computed jump size is negative *and* the competence score is 5 or higher (rule a), or if the computed jump size is negative *and* the task-load score is 2 or lower (rule b), the learner will not jump backwards (i.e., the adjusted jump size = 0) because an easier task or a task with more embedded support may not be challenging enough. Additionally, if the computed jump size is positive *and* the competence score is 3 or lower (rule c), or if the computed jump size is positive *and* the task-load score is 6 or higher (rule d), the learner will not jump forward (i.e., the adjusted jump size = 0) because a more complex task or a task with less embedded support may overwhelm the learner. In Table 3.1, these additional rules yield a jump size of 0.

Germane load. After each task, germane load was measured with a one-item 7-point rating scale as the 'effort invested in gaining understanding of the relationships dealt with in the simulator and the task', ranging from minimum effort (1) to maximum effort (7). The reliability of the germane load measures reported during training was .95 (Cronbach's alpha). Germane load directly reflects the effort a participant has invested in learning and is used to compute task involvement (see below).

Training time. The database connected to the learning environment tracked the time (in minutes) participants spent during training.

Test Phase

Learning outcomes. Learning outcomes were measured with a conceptual knowledge test consisting of 20 multiple-choice questions, administered to the participants after the training. All questions had three alternative answers that were presented in a random order. The test assessed participants' understanding of the dietetics domain (i.e., reasoning with effects of alterations in energy intake, physical activity, and other factors such as gender and smoking on body weight and body composition). An example item is:

Anouk has started smoking. Will her Basal Metabolism Rate (BMR) be affected?

(a) Yes, her BMR will increase.
(b) Yes, her BMR will decrease.
(c) No, her BMR will remain the same.

Three items were not included in the analysis, because one item had an item difficulty value (*p*) of 1 and two items had a negative item-test correlation. Item difficulty (*p)* is defined as the proportion of participants who answer an item correctly (Crocker & Algina, 1986). A *p*-value of 1 means that 100% of the participants answered this item correctly. This means that the correct answer was probably too obvious. The maximum test score was thus 17 points. The internal consistency of the test was .63.

Learning efficiency. The Paas and van Merriënboer procedure (1993; Marcus, Cooper, & Sweller, 1996; Paas et al., 2003) was used to calculate the efficiency of the instructional conditions. First, learning outcomes (i.e., the score on the conceptual knowledge test) and task-load scores for each participant are transformed into

z-scores, using the grand mean across conditions. Then, the mean z-scores for every condition are represented in a Cartesian coordinate system with task load z-scores on the horizontal axis and learning outcomes z-scores on the vertical axis (see Figure 3.2). The line $LO = TL$ through the origin indicates a neutral efficiency. The efficiency, E, is calculated as the perpendicular distance from a data point in the coordinate system to the line $LO = TL$ (Paas & van Merriënboer, 1993). The formula for calculating this distance is:

$$Learning\ Efficiency = \frac{Z_{LearningOutcomes} - Z_{TaskLoad}}{\sqrt{2}}$$

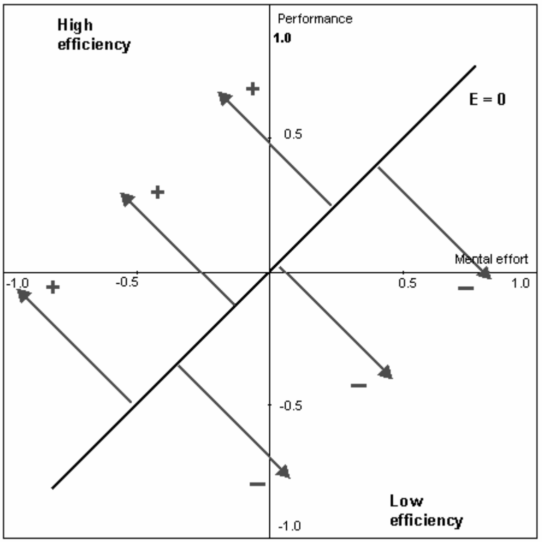


Figure 3.2. Efficiency measure in a Cartesian coordinate system.

Task involvement. The computation of task involvement (Paas et al., 2005) was analogous to the computation of learning efficiency. Now, learning outcomes and germane load (GL) scores are transformed into z-scores using the grand mean across conditions. The task involvement is calculated as the perpendicular distance from a data point in the coordinate system to the line $LO = -GL$. The formula for calculating this distance is:

$$Task\ Involvement = \frac{Z_{LearningOutcomes} + Z_{GermaneLoad}}{\sqrt{2}}$$

Interest scale. After each task in the practice session, learners completed a 7-point rating scale that measured *interest-in-task* with the statement ‘I found the computer lesson interesting', ranging from strongly disagree (1) to strongly agree (7). In addition, in the test phase participants answered a questionnaire that measured their *interest-in-training*. The questionnaire contained 7 items from the interest/enjoyment subscale of the Intrinsic Motivation Inventory (IMI; Deci, Eghrari, Patrick, & Leone, 1994) (e.g., "*I would describe the computer lesson as very interesting*", "*While I was carrying out the computer lesson, I was thinking about how much I enjoyed it*"*)*, which was translated from English into Dutch by Martens and Kirschner (2004). The interest questionnaire had a reliability of .92 (Cronbach's alpha).

Procedure

Introduction. One week before the computer session, all participants participated in an oral introductory session in which both the learning environment and the functioning of the simulator "Body Weight" were presented and explained in a Microsoft® PowerPoint® presentation. In addition, participants were given a short introduction to the dietetics domain. During this introduction the participants could ask questions and the experimenter made sure that the whole procedure was clear to all participants before the actual experiment started.

Training phase. During the training phase participants worked in the learning environment on the learning tasks, using the body weight simulator. Participants were not informed on how the tasks were selected or pre-selected (for the program-control and the shared-control conditions, respectively). The first learning task at the first level of difficulty was used as a practice task, in which all participants could practice with their own data. Competence and task-load scores for the second task (i.e., a conventional task at difficulty level 1) were assessed and used as the first input for task selection. After each task, competence measures were taken and participants indicated on 7-point rating scales the amount of task load and germane load they perceived while working on the learning task. It was emphasized that they were not allowed to skip any part of the answer of the competence and cognitive load questions. If they did, the program prompted them to answer the questions before they could continue. During the training phase the time spent by the participants was logged.

Test phase. One week after the computer session, participants were presented with the paper-and-pencil conceptual knowledge test to measure learning outcomes and the interest questionnaire to assess their interest in the training phase. During the test phase participants were allowed to work at their own pace.

Results

A significant main effect of adaptation was found on training time (i.e., the total amount of time spent on all learning tasks), F (1, 51) = 39.59, p < .001, MSE =

619.42, η^2_p = .437. Participants in the adaptive conditions spent more time on training (*M* = 129.68, *SD* = 23.29) than participants in the non-adaptive conditions (*M* = 87.30, *SD* = 25.95). No effects on training time were found for control or the interaction between adaptation and control. Therefore, ANCOVA's with total training time as a covariate are used in the subsequent analyses and estimated marginal means are presented. For all statistical tests a significance level of .05 was maintained. Table 3.2 provides an overview of the training results.

Table 3.2
Overview of Results from the Training Phase

	Condition							
	Adaptation program control *n* = 15		Adaptation shared control *n* = 15		Non-adaptation program control *n* = 13		Non-adaptation shared control *n* = 12	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Training Time (min.)	126.70	29.27	132.65	15.76	88.55	28.01	85.95	24.69
Practice Performance (max. = 100)	71.90	14.86	75.22	16.36	50.08	19.27	45.77	18.99
Task Load (max. = 7)	3.01	1.20	3.14	1.10	3.61	0.86	4.09	0.97
Germane Load (max. = 7)	4.40	0.56	4.63	0.94	3.73	0.83	4.35	0.50

Note: Estimated marginal means are presented with total training time as a covariate.

Training Phase

Competence scores. A significant main effect of adaptation was found on the competence scores, F (1, 50) = 16.51, *MSE* = 305.25, $p < .001$, η^2_p = .248. Participants in the adaptive conditions achieved higher competence scores (*M* = 73.56, *SD* = 15.46) than participants in the non-adaptive conditions (*M* = 47.97, *SD* = 18.87). No effects on the competence scores were found for control or the interaction between adaptation and control.

Task load. Similarly, a significant main effect of adaptation was found on task load during training, $F(1, 50) = 4.42$, *MSE* = 1.04, $p < .05$, η^2_p = .081. Participants in the adaptive conditions experienced a lower task load (*M* = 3.07, *SD* = 1.14) than participants in the non-adaptive conditions (*M* = 3.85, *SD* = .93). No effects on task load were found for control or the interaction between adaptation and control.

Germane load. A significant main effect of control on germane load during training was found, $F(1, 50) = 4.46$, *MSE* = 0.55, $p < .05$, η^2_p = .082. Participants in the shared-control conditions reported higher mental effort in learning (*M* = 4.49, *SD* = .77) than participants in the program-control conditions (*M* = 4.07, *SD* = .75).

No effects on germane load were found for adaptation or the interaction between adaptation and control.

Test Phase

Not all participants filled out the conceptual knowledge test. Only the data of participants who completed the conceptual knowledge test and the interest questionnaire (n = 50) were used in the analysis. The number of participants that dropped out was evenly distributed over the conditions ($X^2 = .38$, p = .95). Table 3.3 provides an overview of results from the test phase.

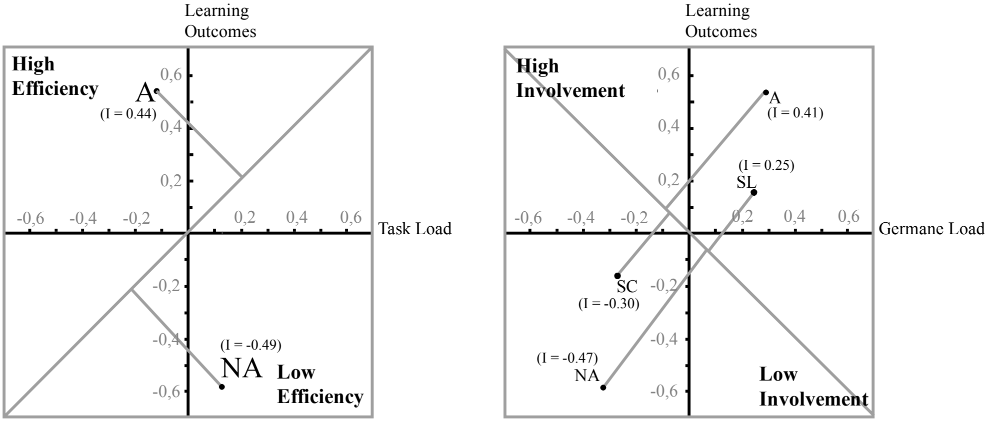
Table 3.3
Overview of Results from the Test Phase

	Condition							
	Adaptation program control *n* = 12		Adaptation shared control *n* = 14		Non-adaptation program control *n* = 12		Non-adaptation shared control *n* = 12	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Learning Outcomes (max. = 17)	12.88	2.57	14.21	1.21	11.95	2.67	12.01	1.91
Learning Efficiency	0.29	1.36	0.60	0.66	-0.32	1.05	-0.67	0.76
Task Involvement	0.15	0.81	0.68	0.94	-0.75	0.87	-0.19	0.75

Note: Estimated marginal means are presented with total training time as a covariate.

Learning outcomes. A significant main effect of adaptation was found on learning outcomes, $F(1, 45) = 4.28$, $MSE = 4.06$, $p < .05$, $\eta^2_p = .087$. Participants in the adaptive conditions scored higher ($M = 13.55$, $SD = 2.07$) than participants in the non-adaptive conditions ($M = 11.98$, $SD = 2.27$). No significant effects on the test scores were found for control or the interaction between adaptation and control.

Learning efficiency. A significant main effect of adaptation was found on learning efficiency, $F(1, 45) = 6.25$, $MSE = 0.98$, $p < .025$, $\eta^2_p = .122$. As hypothesized, participants in the adaptive conditions showed higher efficiency scores ($M = .44$, $SD = 1.03$) than participants in the non-adaptive conditions ($M = -.49$, $SD = .91$). No effects on learning efficiency were found for control or the interaction between adaptation and control. Figure 3.3a depicts a graphical representation of the efficiency based on the standardized scores for learning outcomes and task load.



(a) Learning Efficiency (b) Task Involvement

Figure 3.3. Graphical representation of significant effects on learning efficiency (a) and task involvement (b).
Note. A = Adaptation, NA = Non-adaptation, SL = Shared control, SC = Program control.

Task involvement. Similarly, a significant main effect of control was found on task involvement, $F(1, 45) = 5.37$, $MSE = 0.70$, $p < .025$, $\eta^2_p = .107$. As hypothesized, participants in the shared-control conditions showed higher task involvement ($M = .25$, $SD = 1.05$) than participants in the program-control conditions ($M = -.30$, $SD = 1.02$). Moreover, a significant main effect of adaptation was found on task involvement, $F\ (1, 45) = 7.81$, $p < .025$, $\eta^2_p = .148$. Participants in the adaptive conditions showed higher task involvement ($M = .41$, $SD = 0.92$) than participants in the non-adaptive conditions ($M = -.47$, $SD = .84$). No effects on task involvement were found for the interaction between adaptation and control. Figure 3.3b depicts a graphical representation of task involvement based on the standardized scores for learning outcomes and germane load.

Interest. Table 3.4 presents the mean scores for interest-in-task (measured for each learning task during practice) and the interest-in-training (measured with the interest questionnaire in the test phase). No significant differences between conditions were found.

Table 3.4
Mean Interest-in Task and Interest-in-Training (maximum score of 7)

	Condition											
	Adaptation with program control			Adaptation with shared control			Non-adaptation with program control			Non-adaptation with shared control		
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Interest-in-Task	3.69	1.13	15	3.72	1.30	15	4.0 1	0.59	13	3.98	1.37	12
Interest-in-training	3.58	1.20	12	3.38	1.33	14	3.75	1.31	12	3.91	1.58	12

Note: Estimated marginal means are presented with total training time as a covariate.

Discussion

The first hypothesis of this study that adapting the difficulty and support of the learning tasks to the learners competence scores and perceived task load would make learning more effective and efficient was clearly confirmed by the findings. The learning outcomes of participants who received adaptive training were higher, and they experienced a lower task load during practice than participants who received non-adaptive training. In addition, competence scores of participants in the adaptive conditions were also superior to competence scores of their yoked counterparts. Adaptive training may have reduced the task load during practice to an acceptable level, and therefore, participants may have used their freed-up cognitive resources for learning.

Some comments should be made with regard to the higher training time for participants in the adaptive conditions. These participants may have noticed the relationship between their performance and the difficulty and/or embedded support of the subsequent tasks, whereas participants in the non-adaptive conditions probably lacked this association, which might have negatively influenced their time investment. Since total training time could have influenced the results, all reported analyses included time as a covariate. In addition, the fact that learner control did not yield higher learning outcomes supports the idea that learners with lower levels of competence in a domain lack the ability to make productive use of learner control. In this study, learners cannot be considered to have a substantive level of competence, for which longer exposure to the learning materials (e.g., weeks) than provided in this study would be needed.

The second hypothesis that shared control has positive effects on learner motivation was partially confirmed by the data. Participants in the shared-control conditions showed higher task involvement. In other words, the choice provided positively influenced the amount of effort invested in learning, combined with higher learning outcomes. An explanation is that these participants perceived that their effort was well invested and were thus motivated to invest germane load. Furthermore, task variability can be seen as a strategy to gain the learner's attention (Keller, 1983b). Hence, the relative variability provided by the three tasks presented in the shared control conditions may have further contributed to the positive effect on germane load. However, the absence of a significant effect of shared control on learning outcomes may indicate that the variability of the characteristics of the presented tasks may not have been large enough. A higher degree of variability might have yielded a significant effect on learning outcomes in favor of the shared control conditions.

Another interesting finding pertained to the positive effect of adaptation on task involvement. Providing learners with an appropriate amount of embedded support may have a positive influence on their task involvement, because it prevents the cognitive load of a learning task from becoming too high to perform the task. If this load is too high the learners will lose motivation to continue working on the task (de Croock & van Merriënboer, 2003). In addition, learners provided with an optimum level of task difficulty might be willing to invest effort in learning, which in combination with higher learning outcomes indicates higher task involvement. To sum up, our main results are clearly in favor of adaptive instruction with shared control as expected.

Whereas participants in the shared control conditions showed a higher task involvement, they did not report a higher interest in the learning tasks or in the training. A possible explanation is again related to the limited amount of learner control available. Providing learners with a wider range of tasks to choose from could have revealed differences in interest amongst the experimental conditions. Another feasible reason may be that interest is evoked when learners are given more opportunities for exploration within the learning environment. Participants in the shared control condition were presented with three tasks to choose from, but once the task was selected, the actual performance of the task involved precisely the same activities as the pre-selected task in the program control conditions. Other studies (e.g., Overskeid & Svartdal, 1996; Reeve, Hamm & Nix, 2003; Schraw, Flowerday, & Reisetter, 1998) reported that when provision of choice is the *only* aspect involved to enhance motivation, this may not positively affect interest in the learning tasks. In contrast, a study by Cordova and Lepper (1996) included other aspects (such as internal locus of control and volition) and found that participants reported liking the training more. Hence, the provision of learner control over task selection may be considered as only one aspect to enhance interest, which needs to be combined with other aspects to become effective. For example, other factors such as the pace of instruction or the learner's background knowledge may influence interest. In addition, whereas shared control did not arouse learners' interest and learners did not report liking or enjoying the instruction more, positive results on task involvement indicate that learners still persisted in investing effort to learn from the tasks. Furthermore, that shared control was beneficial for task involvement but not for interest seems to support Paas et al.'s (2005) argument that combining cognitive load and performance measures offers a supplementary approach to inventories that collect motivational data and, in addition, yields information that is not directly reflected in performance-based data.

Our results are consistent with cognitive load theory, which states that an optimal instructional design should decrease extraneous and intrinsic cognitive load and encourage learners to use their freed-up cognitive resources for learning (that is, increase germane cognitive load). From a cognitive load perspective, providing learners with tasks that differ on a number of relevant dimensions from previous learning tasks may increase germane load and improve the construction of cognitive schemata. In our study, extraneous and intrinsic load were successfully reduced by adapting the level of difficulty and support to a learner's competence and task-load scores, and task involvement was induced by providing shared control, recognizing the important role of motivation in designing instruction. These findings are consistent with the results of several studies in other domains that have tailored the difficulty level (Camp et al., 2001; Salden et al., 2004) and both the difficulty level and the level of support (Kalyuga & Sweller, 2005) based on performance scores and cognitive load ratings in the domains of Air Traffic Control and algebra. Hence, initial instruction of a complex skill in educational settings can be facilitated by designing and adapting instruction according to cognitive load theory. Future studies may test the applicability of the adaptive approach in other domains, especially in less structured areas, such as language monitoring comprehension in online reading tutors (Kalyuga & Sweller, 2005).

Assessment of complex performance must include several qualitatively distinct aspects (e.g., breadth and depth of an integrated and organized knowledge base, the possession and implementation of flexible problem-solving strategies, learners' self-monitoring skills, or categorical diagnosis of problems) to obtain valid and reliable information. In our study, learners competence scores were only based on answers to multiple-choice questions and performance measures of operating the simulator. The use of more advanced process-tracking methods, such as concurrent verbal protocols, retrospective reporting, and eye tracking would provide more sensitive indicators of a learner's competence and her understanding of the rationale behind the steps performed, and will thus further refine the basis for adaptive task selection, which in turn may provide superior learning results. Furthermore, our study used task load and competence measures for task selection purposes. In future research, germane load might be considered as an additional factor for task selection. A high germane load indicates that the learner is investing a substantial part of her available cognitive resources in learning. A selection table that incorporates this type of load should therefore aim to keep it as high as possible. When competence is (relatively) high, a learner who reports a high germane load should not receive a less difficult task or a task with more support, but an equally difficult or even more difficult task with equal or less support. This guarantees that every subsequent task is challenging for the learner. Such a refined selection table including germane load might be expected to be superior to the selection table used in the current study, in which we tried to keep the subsequent tasks challenging by adding a rule that prescribed not to select a less difficult task or a task with more support when the competence-score was 5 or higher and the task load was 2 or lower.

Concerning the measurement of cognitive load, theorists are faced with the challenge to distinguish the different types of cognitive load through self-reporting instruments. In this respect, Opfermann, Gerjets, and Scheiter (2005) found preliminary differential effects in a study in which cognitive load was measured with six items that assessed the different types of cognitive load on a 9-point Likert scale. Our findings may also indicate that learners seem to be able to distinguish between task load, which may be seen as a combination of intrinsic and extraneous load, and germane load.

With regard to the learning outcomes, two remarks should be made. First, the multiple-choice questions might have been relatively easy for the participants who scored moderately high in all conditions. More complex test questions could have increased differences between the experimental conditions. Second, learning outcomes were only measured with the conceptual knowledge test. In future research, transfer tasks should also be used to measure participants' learning outcomes.

To conclude, the results of this study indicate that adapting the difficulty and support of selected tasks to the learner's level of competence and task load and providing learners with some control over the process of task selection is advisable. Adaptive task selection yielded more effective and efficient learning. In addition, shared control enhanced learners' motivation. Further research is needed to determine ways to control extraneous and intrinsic cognitive load and to optimize germane load, for example, by providing learning tasks that differ on a number of relevant dimensions from previously presented learning tasks to ensure a high variability which helps learners to construct new schemata, with positive effects on learning.

4

Combining Shared Control with Variability over Surface Features: Effects on Transfer Test Performance and Task Involvement[4]

Positive effects of learner control decrease when learners do not perceive the control given to them, make suboptimal choices, or are cognitively overloaded by the amount of choice. This study proposes shared control (i.e., learners choose from a pre-selection of suitable tasks) over highly variable tasks to tackle these problems. Ninety-four students participated in a 2 x 2 factorial experiment with the factors control (program, shared) and variability of surface features (low, high). Results show superior effects on transfer test performance and task involvement of shared control when learners can choose from pre-selected tasks with surface features that are different from the surface features of previous tasks.

[4] This chapter is based on: Corbalan, G., Kester, L., & van Merriënboer, J.J.G. (in press). Combining shared control with variability over surface features: Effects on transfer test performance and task involvement. *Computers in Human Behavior*.

Recent instructional theories advocate on-demand methods of education in which learners are given freedom to choose their own learning path (Bell & Kozlowski, 2002; Schnackenberg & Sullivan, 2000; Williams, 1996). In contrast to program controlled instruction, learner controlled instruction allows learners to make their own decisions on specific elements of instruction. These include, for example, the instructional components (e.g., learning tasks, information elements), the component characteristics (e.g., task contexts, modality of presented information), as well as task sequence and instructional pace. Instructional locus of control can be thought of as a continuum ranging from full program control to full learner control, involving several forms of shared control (Hannafin, 1984). Learner control is believed to positively influence learning and motivation (Kinzie & Sullivan, 1989; Ross, Morrison, & O'Dell, 1989; Schnackenberg & Sullivan, 2000; Williams, 1996). It permits learners to adapt particular characteristics of the learning material to their individual preferences and needs (Kinzie, 1990; Merrill, 1994), and has been theorized to be a useful alternative to the classical aptitude-treatment interaction approach in that learners can become system independent (Federico, 1980).

Although the beneficial effects of learner control are supported from a theoretical perspective, empirical research shows both beneficial and detrimental effects on learning. These inconsistent results suggest that learner control can be either motivating or demotivating (Katz & Assor, 2007). According to Skinner (1996), there is little consensus on the kinds of control that are beneficial or harmful for learning and on how these may interact with certain learner and situational characteristics. In any case, the unconditional use of learner control is not supported (Freitag & Sullivan, 1995; Lin & Hsieh, 2001; Skinner, 1996; Williams, 1996). Potential threats of learner control include, amongst others: (1) a lack of perception of control when learners do not see the choices provided as sufficiently different from each other; (2) making suboptimal choices because learners are not aware what is best for their learning, and (3) a high cognitive load on learners' processing resources influenced by the amount of choice available. Well-designed education should prevent these potential pitfalls and ensure the necessary conditions to optimize learner control.

First, with regard to *perception of control*, some authors (Cordova & Lepper, 1996; Katz & Assor, 2007; Kinzie, 1990; Langer, 1975; Lepper, 1985; Taylor & Brown, 1988) argue that the positive effects of learner control remain apparent even when the choices provided are irrelevant for learning and the control is merely an illusion. In a study carried out by Cordova and Lepper (1996), participants who were given control over aspects of the task that were not relevant for learning (e.g., the names of the characters of a computer game designed to teach arithmetical and problem-solving skills) scored significantly higher on a posttest than participants who did not have control at all. In a recent study, Hasler, Kersten, and Sweller (2007) found that participants in learner-paced conditions achieved higher learning efficiency despite the fact that the control options were rarely used. The mere availability of control had an added value on learning, since the instructional content was identical in both program paced and learner paced conditions. In addition, when learners do not perceive that they are in control of something this decreases performance and increases frustration, especially when they are still forced to make a selection (Burger & Cooper, 1979). Accordingly, perception of control refers to the individual belief of how much control is available (Skinner, 1996). In line with this idea, many theorists (Averill, 1973; Burger, 1989; Skinner, 1996) argue that learner control is only a powerful predictor of functioning and positively related to performance if it is actually perceived as control by the learners (Savage, Perlmunter, & Monty, 1979).

Furthermore, when learners perceive they are in control of something, they will most likely be engaged in learning activities, and will actually allocate their cognitive resources to learning because of the satisfaction derived from just performing the task (Deci, Vallerand, Pelletier, & Ryan, 1991; Fisher & Ford, 1998; Keller, 1983; Salomon, 1983, 1985). When learners are willing to invest mental effort in learning, this may in turn positively influence performance (Paas & van Gog, 2006; Volet, 1997). In this respect, perceived learner control is believed to increase learners' involvement. An involved learner will most likely invest more mental effort in performing the learning tasks, which might result in higher learning outcomes (Kinzie, 1990; Paas, Tuovinen, van Merriënboer, & Darabi, 2005). Accordingly, lack of perceived control decreases learners' involvement in learning, threatening learning outcomes.

Second, novice learners commonly make *suboptimal choices* due to a lack of sufficient or adequate domain knowledge. If learners select what they like rather than what they need, learner control may even have negative effects on learning. For instance, learners may reduce instructional time by skipping or omitting considerable amounts of instructional materials essential for good performance (Merrill, 2002; Ross & Morrison, 1989; Snow, 1980; Williams, 1996). Hence, learners with insufficient or inadequate knowledge should not be allowed to make instructional choices unless they are taught or supported to make the right selections first. With shared control, for example, the learner may make the final selection of one learning task from a subset of suitable tasks, which are pre-selected from all available tasks by an instructional agent (e.g., a teacher or computer system).

In addition, learners may opt to select learning tasks with highly similar surface features (i.e., irrelevant aspects of the task that are not directly related to goal attainment) because those tasks are more familiar to them. In a study by Ross, McCormick, and Krisak (1986), college students majoring in nursing or education were allowed to select from several alternative themes (e.g., sports, medical, educational, abstract) to learn statistics. They found that nursing students tended to select the medical theme while education students tended to select the educational theme. Hence, when learners exercise contextual control, such as over surface features, this may make instruction more personally relevant to them without affecting the basic lesson content (Hannafin, 1984; Kinzie, 1990). However, learner control over context may ultimately be detrimental for learning since learners may tend to continuously select similar contexts. Increased exposure to a variety of contexts is believed to promote transfer (Ross & Morrison, 1989). Variability over surface features, which enables learners to distinguish task-relevant from task-irrelevant information, encourages the adaptation to new tasks. Providing learners with learning tasks representing a high variability of surface features has been shown to facilitate the construction of more general problem-solving rules. Thus, it helps learners to abstract away from the contexts and to focus on those components that are shared by the learning tasks, enhancing transfer of learning to new, unfamiliar situations (Chen & Mo, 2004; Holyoak & Koh, 1987; Quilici & Mayer, 1996). This abstraction process requires the mindful engagement of the learners, increasing their "germane" cognitive load (i.e., load caused by cognitive processes that directly contribute to learning; Clark, Nguyen, & Sweller, 2005; Paas & van Merriënboer, 1994; Sweller, van Merriënboer, & Paas, 1998; van Merriënboer, Kester, & Paas, 2006; van Merriënboer et al., 2002).

Surface features (e.g., species in inheritance tasks because the rules for the transmission of hereditary characteristics from parent organisms to their offspring are the same for animals, plants, and humans) are often more perceptible for learners than structural features (e.g., parents' genes composition) (Cummins, 1992; Gick & Holyoak, 1987; Quilici & Mayer, 1996, 2002). Learners who are presented a subset of learning tasks that *differ* in surface features from prior tasks will most likely recognize those differences, perceive the choices as being valuable and use this information to make tasks more relevant for them when provided with control (Katz & Assor, 2007; Kinzie, 1990). Hence, learner control over task selection may positively influence their perception of control - as compared to learners who are presented with one 'program selected' task - and thus their involvement and transfer performance. Learners who are presented with a subset of tasks with surface features *similar* to that of prior tasks, in contrast, will recognize the *lack* of differences. This may even lead to frustration, especially when they are still forced to make a selection (Burger & Cooper, 1979) which may negatively influence their perception of control and thus their involvement in learning and performance on transfer tasks. Hence, to enhance the perception of control the choices provided to learners should be attractive ones (Kinzie, 1990). In this study, the provision of a subset of variable tasks is presented to make choices more attractive (Keller, 1983). In addition and at least equally important, learners who are given a subset of learning tasks differing in surface features from prior tasks are prevented from making suboptimal choices because all the tasks they can choose from differ from the prior one with regard to their surface features. As argued above, this ensures a certain amount of variability which is expected to facilitate transfer of learning.

Third, learner control may introduce potential problems with excessive *cognitive load* (Scheiter & Gerjets, 2007; Schwartz, 2004). Even experienced learners may become overwhelmed and demotivated by an excessive amount of choice, for instance, when they are provided with hundreds of tasks to choose from (Iyengar & Lepper, 2000; Schwartz, 2004). In this respect, learners may experience difficulties in selecting, sequencing, and pacing huge amounts of information because of cognitive overload (Scheiter & Gerjets, 2007). However, several studies (Fry, 1972; Kinzie & Sullivan, 1989; Lahey, Hurlock, & McCann, 1973) show that the majority of learners actually prefer some control. Hence, providing learners with a limited amount of choice may avoid the potential pitfalls of a too high amount of choice and yet grant a desired amount of learner control. This is realized when an instructional agent (e.g., a teacher or a computer program) and the learner share control over the process of task selection (Corbalan, Kester, van Merriënboer, 2006; Tennyson & Buttery, 1980). In this two-step process, a computer program first selects a subset of learning tasks with desirable task features (e.g., surface features) based on task features of previously selected tasks (program control). Second, the learner selects from this subset one task to work on (learner control). This prevents overloading the learners' cognitive resources, while still giving them a sense of control and preventing them from making wrong instructional decisions (Kinzie & Sullivan, 1989). Hence, this study implements shared control to prevent cognitive overload by reducing the amount of choice given to learners.

To sum up, for learner control to be effective, learners must both perceive they are in control of something and be supported to make optimal choices – otherwise learner control may even hamper learning. Variability in the surface features of learning tasks helps to meet these requirements since (a) it enhances the perception of control and, in combination with shared control, (b) it prevents learners of making suboptimal choices because the subset of tasks they may choose from ensures a high variability, which is a prerequisite to abstraction and transfer. Furthermore, shared control prevents cognitive overload by reducing the amount of possible choices given to the learners. The purpose of this study is to investigate under which conditions shared control is optimized. Shared control is hypothesized to increase learners' involvement in learning and yield higher performance on transfer tasks provided that high variability over surface features of tasks is ensured; shared control in combination with low variability may even decrease learners' involvement and yield lower performance on transfer tasks.

Method

Participants

Ninety-four first year students (90 females and 6 males) in the Health Sciences domain of a Dutch school for secondary Vocational Education and Training (VET) participated in this study. Their mean age was 17.48 years (*SD* = 2.04). A 2 x 2 factorial design was used to study the effects of control (program control vs. shared control) and variability of surface features in the learning tasks (low variability vs. high variability). All participants received a movie DVD for their participation. They were randomly assigned to one of the four experimental groups: program control with low variability (n = 22); program control with high variability (n = 25); shared control with low variability (n = 23); and shared control with high variability (n = 24).

Materials

Electronic learning environment. The learning environment was a web application containing a basic introduction to the domain of genetics, a factual knowledge test, the learning tasks, the transfer test, and mental effort questions to estimate cognitive load.

Basic introduction. In the basic introduction, participants studied the main concepts in the domain of genetics included in the training (i.e., dominant and recessive genes, homozygous or heterozygous gene pairs, genotype and phenotype) and a worked-out example containing all the problem-solving steps of a representative inheritance task.

Factual knowledge test. The factual knowledge test contained 16 multiple-choice questions. Five items were not included in the analysis because those items had a negative item-test correlation. The maximum test score was 11 points. The reliability of the test was .62 (Cronbach's alpha).

Learning tasks. The learning environment was connected to a database which contained the learning tasks. The task database consisted of 162 completion tasks about inheritance mechanisms of acquired traits that apply to different species and traits (e.g., hair, fur or leaf colour). Completion tasks present a given state, a goal state, and a partial solution that learners have to complete. Table 4.1 shows the features of the learning tasks in the database: the first column contains species and species type (within brackets), and the second column contains traits per species type and trait parts (within brackets). These are the surface features of the tasks. In addition, depending on the parents' gene forms for a certain trait, three crossing types can be distinguished for each trait in each species, which leads to three different types of solutions. Crossing types depend on whether the parents are homozygous (i.e., the two gene forms - dominant or recessive - for one trait are identical) or heterozygous (i.e., the two gene forms for one gene are different) for a trait. These are the structural features of the tasks.

Table 4.1.
Composition of the Database with Learning Tasks

Species (species type)	Traits (part traits)	Crossing type I	Crossing type II	Crossing type III
		Homozygous parent x Homozygous parent	Homozygous parent x Heterozygous parent	Heterozygous parent x Heterozygous parent
Humans (European/African/Asian)	Colour (hair/eyes) Shape (eyelashes/nose) Length (hair/thumb finger)	18 Tasks	18 Tasks	18 Tasks
Animals (dog/cat/guinea pig)	Colour (hair/eyes) Shape (ear/hair) Length (tail/fur)	18 Tasks	18 Tasks	18 Tasks
Plants (pea/corn/bean)	Colour (flower/leaf) Shape (fruit/pod) Length (axis/fruit)	18 Tasks	18 Tasks	18 Tasks

Each participant completed twelve learning tasks. After each learning task, two multiple-choice questions with four answer options were presented to the learners. Each correct question scored one point and each wrong question scored zero points, leading to a maximum score of twenty-four points. The reliability of the learning-tasks questions was .95 (Cronbach's alpha). An example question is: "what are the expected genotypes for coat length of the puppies?"

Transfer test. The transfer test consisted of 12 transfer tasks, divided in 6 near transfer tasks and 6 far transfer tasks. The near transfer tasks were analogous to the learning tasks but contained different surface features (i.e., other subjects within the species, e.g., a fish, and other traits, e.g., the swimming pattern of a fish) and determined whether participants were able to apply the learned procedures. The far transfer tasks differed in structural features and were meant to determine whether participants were able to apply the learned procedures to new situations. The following tasks were used: (a) a dihybrid crossing task which required two different traits to be treated separately; (b) a family tree task in which participants had to infer the genotype of one of the parents based on the information of one of the grandparents; (c) a task in which participants had to infer the genotype of an individual from information from the relatives given; (d) a second generation task that required learners to determine the offspring of the offspring, (e) a co-dominant genes task, that is, genes that are equally strong and both expressed, and (f) a task in which participants must apply the acquired knowledge in a bottom-up way, that is, using the information of the offspring to find out information of the parents. The maximum test score was 12 points. The reliability of the test was .83 (Cronbach's alpha).

Mental effort. After each learning task and transfer test task, participants' perceived mental effort was measured as the 'effort required to complete the task' with a one-item 7-point rating scale (Paas, 1992; Paas, Tuovinen, Tabbers, & van Gerven, 2003). Reliability of the mental effort measures reported during the training was .97 and during the transfer test .95 (Cronbach's alpha).

Motivation questionnaires. After the 12 learning tasks, participants completed five 7-point rating scales (i.e., perceived ability, effort, interest, usefulness, and intrinsic motivation) of the Intrinsic Motivation Inventory (IMI; Deci, Eghrari, Patrick, & Leone, 1994). The scales contained, in order, five, five, seven, four, and nine items. Reliability analysis (Cronbach's alpha) of the IMI scales yielded internal consistencies of .76, .70, .90, .83, and .67, respectively. In addition, participants completed two 7-point rating scales (i.e., control beliefs - 4 items - and self-efficacy - 8 items -) of the Motivated Strategies for Learning Questionnaire (MSLQ; Garcia & Pintrich, 1994). Reliability analysis of the MSLQ scales yielded, in order, Cronbach's alphas of .76 and .91.

Time logging. The learning environment kept track of the time (in minutes) participants needed to complete each learning task and transfer test task.

Learners' involvement. In this study, task involvement (Paas et al., 2005) was used as an indicator of learners' involvement. Based on the assumption that motivation, mental effort and performance are positively related, task involvement is computed by combining transfer test performance and mental effort invested during training. According to Paas et al. (2005), low mental effort combined with low test performance can be considered indicative of low task involvement, whereas high mental effort combined with a high test performance is indicative of a high task involvement. To calculate task involvement, transfer test scores and mental effort scores are first standardized and the *z*-scores are entered into the formula (Paas et al., 2005):

$$\textit{Task Involvement} = \frac{Z_{\textit{Transfer Score}} + Z_{\textit{Mental Effort}}}{\sqrt{2}}$$

In the formula, task involvement is computed for each learner as the perpendicular distance between a dot in the cross of axes (i.e., the z value for mental effort on the x-axis and the z value for transfer score on the y-axis) and the diagonal, T = 0, where mental effort and transfer scores are proportionally related to each other.

Procedure

Training phase. Prior to the training phase, participants received the basic introduction and the factual knowledge test. After that, participants received the 12 learning tasks (i.e., completion tasks) and the associated mental effort measures.

During the training phase, the 12 tasks were dynamically selected as follows: in the program control conditions, the program selected and presented each time only *one* task to the participant. In the shared control conditions, the program randomly pre-selected three tasks from six possible tasks resulting from the combination of species types (x3) and part traits (x2) and presented these *three* tasks to the participant, from which the participant selected one task to perform. In the low variability conditions, the species (i.e., human, animal or plant) and trait (i.e., colour, shape or length) of the *first* (range of) learning task(s) was randomly selected and counterbalanced over the participants. Each next (range of) task(s) was selected from all available tasks featuring the *same* species and trait as the prior task, thus species and trait remained invariable throughout the training. In the high variability conditions, the species and trait of the *first* (range of) learning task(s) were randomly selected. Each next (range of) task(s) was selected from tasks featuring a *different* species and a *different* trait than the previously performed task. Figure 4.1 shows an example of a possible sequence of three learning tasks performed in the low and high variability conditions.

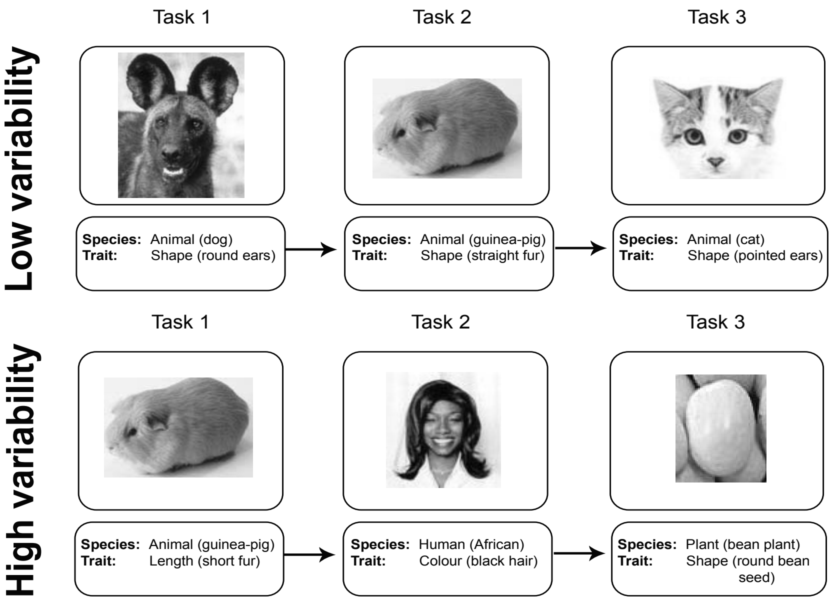


Figure 4.1. Example of a sequence of three learning tasks in the low variability condition (upper row) and the high variability condition (lower row).

Moreover, each of the three crossing types (i.e., homozygous x homozygous, homozygous x heterozygous, and heterozygous x heterozygous) was randomly selected to avoid a practice bias. Once the learner performed a learning task with a certain crossing type four times, remaining tasks with this crossing type were deleted from the database, leading to a total of twelve learning tasks (four learning tasks for each of the three crossing types). This assured that every learner practiced all the solution-relevant aspects.

Test phase. Directly after the training, participants filled out the scales corresponding to the IMI and the MSLQ. One week after the first session, the participants performed the transfer test. After each test item, mental effort measures were taken.

Results

An ANOVA on the factual knowledge test filled out by the participants prior to the training revealed no difference between conditions (F (1, 90) < 1). Therefore, all subsequent analyses are performed using ANOVAs with between-subjects factors variability and control. For all statistical tests a significance level of .05 was maintained. Table 4.2 provides an overview of the mean scores and standard deviations for the dependent variables during training and transfer.

Table 4.2
Results of the Training Phase and the Test Phase

	Low Variability				High Variability			
	Program control		Shared control		Program control		Shared control	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Factual knowledge test	6.05	2.50	5.60	2.57	5.65	2.29	5.79	2.05
Training								
Time (min.)	17.92	3.48	16.92	4.17	18.15	5.45	16.51	2.82
Performance (max. = 24)	17.73	5.87	15.64	6.73	15.39	6.75	18.75	5.34
Mental Effort (max. = 7)	3.35	1.41	3.22	1.41	3.36	1.36	3.44	1.05
Test								
Time (min.)	36.24	9.42	38.15	13.35	37.93	11.55	39.62	10.74
Performance (max. = 12)	6.43	2.86	5.26	2.46	5.70	3.04	7.04	2.80
Mental Effort (max. = 7)	4.35	1.14	4.60	1.23	4.30	1.40	4.24	1.07
Task Involvement	0.14	0.79	-0.27	0.79	-0.10	0.81	0.24	0.68

Training Results

Due to technical difficulties, there were seven system-missing values corresponding to seven participants on one of the mental effort values measured across the 12 learning tasks. Each missing value was replaced by the mean mental effort on the remaining 11 tasks of that particular participant during the training phase.

An ANOVA revealed no main effects for control (F (1, 90) = .25, MSE = 9.48, *ns*) and variability (F (1, 90) = .091, MSE = 3.51, *ns*) on training performance, but there was a significant interaction effect (F (1, 90) = 4.51, MSE = 173.85, $p < .05$, $\eta^2_p = .05$). As Figure 4.2 shows, with high variability shared control yields better training performance than program control, but with low variability program control yields better performance than shared control. Post-hoc multiple comparisons using Bonferroni's adjustment shows no significant differences between the different conditions.

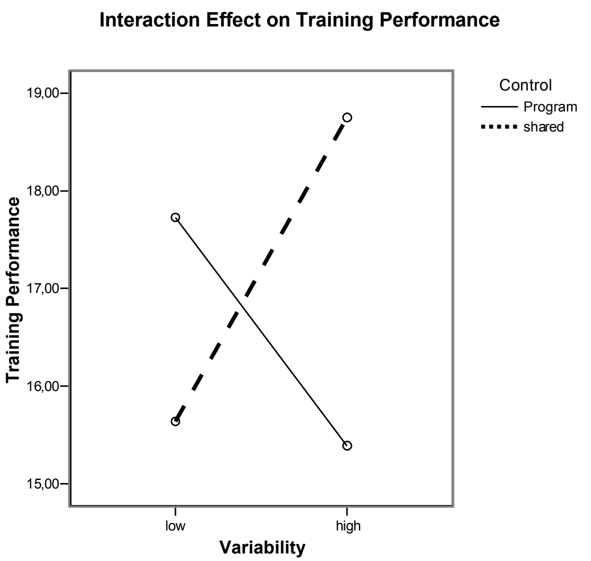


Figure 4.2. Interaction of variability and control on training performance.

No effects on training time were found for control (F (1, 90) = 2.24, MSE = 147410.44, *ns*), variability (F (1, 90) = .012, MSE = 716.34, *ns*), and their interaction (F (1, 90) = .14, MSE = 8690.87, *ns*). Also no effects on mean mental effort invested during training were found for control (F (1, 90) = .009, MSE = .016, *ns*), variability (F (1, 90) = .16, MSE = .275, *ns*), and their interaction (F (1, 90) = .155, MSE = .270, *ns*).

Test Results

Only the data of participants who attended the second session of the experiment (*N* = 86) could be included in the analysis of test results. The number of participants that dropped out (*n* = 8) was evenly distributed over the conditions (X^2 = .29, *p* = .96). This resulted in the following group composition: program control with low variability (*n* = 22); program control with high variability (*n* = 21); shared control with low variability (*n* = 22); and shared control with high variability (*n* = 21).

No main effects were found for control (*F* (1, 82) = .018, *MSE* = .143, *ns*) and variability (*F* (1, 82) = .751, *MSE* = .5.865, *ns*) on transfer test performance, but there was a significant interaction effect (*F* (1, 82) = 4.310, *MSE* = 33.68, $p < .05$, η^2_p = .05). As Figure 4.3a shows, with high variability shared control yields better transfer test performance than program control, but with low variability this pattern is reversed. Post-hoc multiple comparisons were conducted using Bonferroni's adjustment. These comparisons revealed that participants in the shared control with high variability condition outperformed participants in the shared control with low variability condition (*F* (1, 82) = 4.33, $p < .05$). Thus, as hypothesized, shared control yields highest transfer test results provided that high variability over surface features is ensured. The other comparisons revealed no differences.

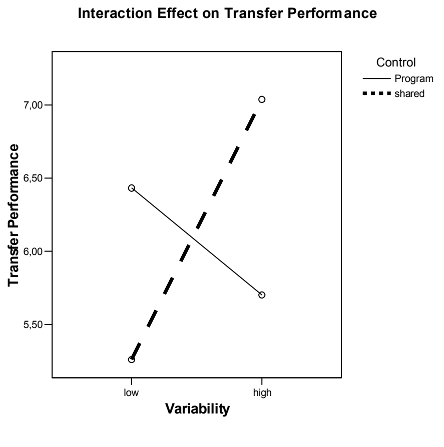


Figure 4.3a

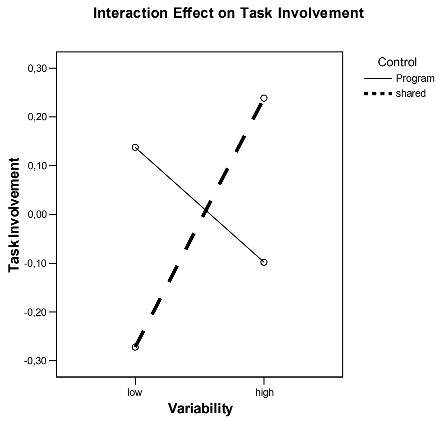


Figure 4.3b

Figure 4.3. Graphical representation of the interaction of variability and control on transfer test performance (a) and task involvement (b).

No main effects were found for control (*F* (1, 82) = .829, *MSE* = .047, *ns*) and variability (*F* (1, 82) = .690, *MSE* = .405, *ns*) on task involvement, but again an interaction effect occurred (*F* (1, 82) = 5.043, *MSE* = 2.992, $p < .05$, η^2_p = .06). As Figure 4.3b shows, with high variability shared control yields higher task involvement than program control, whereas with low variability this pattern is reversed. Post-hoc multiple comparisons were conducted using Bonferroni's adjustment. These comparisons revealed that participants in the shared control with high variability condition were more involved in the tasks than participants in the shared control with low variability condition (F (1, 82) = 4.719, p < .05). The other comparisons revealed no differences.

Analyses revealed no significant main effects for control (F (1, 82) = .538, MSE = 249827.26, *ns*), variability (F (1, 82) = .418, MSE = 194348.13, *ns*), and their interaction (F (1, 82) = .02, MSE = 918.80, *ns*) on the time spent on the transfer test. Also no effects on mean mental effort invested during the transfer tests were found for control (F (1, 82) = .003, MSE = .006, *ns*), variability (F (1, 82) = .056, MSE = .120, *ns*), and their interaction (F (1, 82) = .717, MSE = 1.524, *ns*).

Motivation questionnaire

Table 4.3 provides an overview of the mean scores and standard deviations for the scales of the motivation questionnaire.

Table 4.3
Results of the Motivation Questionnaire

	Low Variability				High Variability			
	Program control		Shared control		Program control		Shared control	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Perceived ability	4.03	1.03	4.27	1.18	4.32	1.25	4.12	1.10
Effort	4.58	1.04	4.85	1.26	4.48	1.19	4.46	1.01
Interest	3.85	1.24	3.71	1.49	4.01	1.56	1.48	1.31
Use	4.25	1.26	3.80	1.27	4.07	1.57	3.58	1.32
Intrinsic Motivation	4.43	.73	4.59	1.02	4.56	.96	4.37	.86
Control Beliefs	5.60	1.10	4.85	1.47	5.38	1.18	5.15	1.04
Self-Efficacy	5.12	1.07	4.5	1.34	4.84	1.15	4.48	1.08

A main effect of variability on participants' reported self-efficacy was found, F (1, 90) = 4.294, MSE = 5.675, p < .05, η^2_p = .05). Participants in the program control conditions reported higher self-efficacy (M = 4.98, SD = 1.11) than participants in the shared control conditions (M = 4.49, SD = 1.17). No effects were found on any of the other scales.

Discussion

This study investigated whether shared control over the selection of learning tasks, in which a computer program first makes a pre-selection of tasks and the learner makes the final selection, might be optimized by ensuring variability of the surface features of the learning tasks (i.e., species and traits in inheritance tasks). Learners who made a selection from learning tasks with surface features different from the previous task were expected to reach higher transfer test performance, and be more involved in learning than learners who made a selection from learning tasks with surface features very similar to the previous task. We found support for this hypothesis. The observed interactions show that the shared control groups were influenced by the low/high variability provided whereas the program control groups did not. In addition, participants in the group with shared control and high variability of surface features achieved higher transfer test performance, and were more involved in learning than participants in the group with shared control and low variability of surface features.

These findings are in line with the explanation for the effect of control and variability on transfer suggested by Gay (1986). According to this author, learner control is more efficient under conditions in which learners have a well established conceptual understanding of the content domain. In our study, participants in the shared control with high variability condition – in which schema construction is believed to be promoted by the successive presentation of varied instances – may have quickly achieved superior understanding of the genetic content. Therefore, they might have profited more from shared control than participants in the shared control with low variability condition, in which schema construction was not facilitated because the tasks to choose from were very similar to the previous tasks.

Moreover, despite the fact that in the shared control conditions the available choice concerned task aspects irrelevant for goal attainment (i.e., surface features which were more perceptible or more salient for the learners than the structural features), participants profited more from the given control when the surface features of the tasks to choose from were different from the surface features of the previous task. If participants have control but can only choose from tasks similar to the previous task, they may not see the meaning of making a selection, possibly leading to frustration (Burger & Cooper, 1979), less involvement in learning, and lower performance. Accordingly, when choices are provided, the degree to which individuals perceive they are in control seems to be related to the level of attractiveness of the choices available (Kehoe, 1979). Choosing from tasks that are different from the previous one appears to be more attractive than choosing from tasks that are very similar to the previous one. This may have made participants in the shared control with high variability condition be more receptive to the instructional material than participants in the shared control with low variability condition, who probably perceived less control. This supports the idea that *perception* of control is a condition sine qua non. Without the perception of control, which may be optimized by task variability, learners' involvement in learning is low and shared control does not work.

Nevertheless, participants in the shared control with high variability condition performed not higher than participants in the two program-controlled conditions on any of the variables. A plausible reason could be that the choices provided to the learners, including learners in the high variability conditions, were rather restricted (i.e., including a different species and trait than the prior task, but still limited to one species and trait) and learners may have seen the provided choices as not sufficiently different from each other. Consequently, although our results on the transfer test and on involvement supported our hypothesis, conclusions based on these interaction effects should be drawn with caution. Future studies should include direct measures of perceived control to uncover whether learners' perception of control in a shared control with high variability condition is indeed higher than in other conditions.

With regard to the learners' performance during training, the found interaction of control and variability is not surprising and further supports the added value of shared control, provided that the perception of control is enhanced by high variability of the surface features of the learning tasks. Whereas the value of shared control was enhanced when the tasks to choose from were different from the previous task (high variability), program control could have been influenced by such a higher variability. For the program control groups, repeatedly presenting a task with similar surface features might have made surface features more salient to learners than repeatedly presenting a task with new surface features, which could have benefited performance.

Participants in the program control conditions reported higher self-efficacy than participants in the shared control conditions. In the MLSQ, the definition of self-efficacy involves the expectancy for success and judgments of one's ability to accomplish a task and confidence to perform a task (Garcia & Pintrich, 1994). The results of this study seem to indicate that when the sequence of tasks is program-controlled, learners have more confidence that they will be able to perform the task. Selecting own tasks may thus cause learners to become more insecure. As in other studies on learner control (e.g., Williams, 1996), no differences on the other scales of the motivation questionnaire were found. This could be a typical disadvantage of short laboratory experiments. Future studies may investigate the effects of shared control in more authentic learning environments and may also study the effects of giving learners a higher amount of control than in the current study, where it was limited to choosing between three tasks. A higher level of control could be achieved, for instance, by allowing learners to choose between more learning tasks that differ in their - similar and dissimilar - surface features and possibly also their strufctural features.

Another possible explanation for the lack of effects on measures of motivation is that after twelve learning tasks participants were used to the mode of task selection, which might have reduced possible differences at the end of the training. For instance, measuring motivation after three or four tasks, instead of after the whole training, could have been a more sensitive measure of the effects of variability of surface features of tasks on motivation. Future studies may also investigate learners' perception of control over task selection. The 'perceived autonomy' scale included in the IMI, which was included in this study, turned out to be not very valuable because it was highly unreliable ($\alpha = .28$) and only partially measured the perception of control as applied in this study.

Regarding the results on the mean mental effort, no significant effects were found but the observed pattern is in accordance with cognitive load theory (CLT; Sweller, 1988; Sweller, van Merriënboer, & Paas, 1998; van Merriënboer & Sweller, 2005). According to CLT, mental effort during training would be higher when variability is high, as a result of the germane load associated with the construction of more general rules, which in turn leads to lower mental effort during transfer test performance. In this study, overall cognitive load was reduced by using completion tasks, by limiting the choice available, and by randomizing variability on structural features (the mean mental effort invested during training lies below the neutral score of 4 in all conditions). Nevertheless, in the low variability conditions the mean mental effort during training was slightly lower than in the high variability conditions, and the mean mental effort invested on transfer test performance showed the reverse pattern. Moreover, shared control with low and high variability led, in order, to the lowest and the highest mean mental effort during training, and, again in order, to the highest and the lowest mean mental effort during transfer test performance. Future studies should further investigate the relationship between variability over surface features and mean mental effort. Furthermore, more sensitive cognitive load measures that make a distinction between germane and overall load may reveal more powerful results, especially when the amount of learning tasks to choose from is increased.

In this study it has been assumed that learners in the shared control with high variability condition would *recognize* the differences between learning tasks and that this would increase their *perceived control*. However, none of these two variables was directly measured. In general, future studies may also investigate the effects of providing learners with shared control over the selection of more complex tasks or the selection of tasks in other learning domains. Another route for future research pertains to the effects of further increasing the variability of surface features of learning tasks. Increasing variability, for instance, by including other groups of humans, animals, and plants, and by broadening the number of inherited traits, such as diseases or physical variations (e.g., being right or left handed in humans, ears that stand up or hang down in animals, or position of the flower in some plants), may have revealed more powerful effects of variability as a result of stronger schema formation. In addition, in our study variability on the structural features was randomized across *all* conditions. Whereas in the low variability conditions the variability was limited to one species and one trait, in the high variability conditions the surface features varied from each precedent task. Varying both structural and surface features could have masked any positive effect of the variability in surface features. Future studies should study the effects of providing learners with variability over the surface features when variability on structural features is predetermined.

Besides, the maximum amount of learner control that may be provided, especially when both surface and structural features vary highly, requires further investigation since they may easily overwhelm learners when provided together. Besides, completion tasks integrate the presentation of new information (in the given part of the solution) and the practice of problem-solving steps that have already been introduced (in the to-be-completed part of the solution). In our study, variability only referred to the presentation of surface features in the given part of the solution. Futures studies may investigate next to the effects of presentation variability, the effects of "practice variability" (i.e., practicing problem-solving steps that have not been practiced in the previous task) on transfer test performance and cognitive load. A final observation concerns the fact that 93,75 % of the participants in our study was female. Future research is needed to determine whether the results can also be found in other domains and with another population of learners (e.g., with mostly males or with a similar proportion of males and females).

To conclude, this study shows that when learners are provided with shared control, transfer test performance and involvement in learning are enhanced if learners can choose from learning tasks with surface features that are distinctly different from surface features of previously performed tasks. In this respect, this study complements the attempts of other authors to determine guidelines for implementing learner control (Hannafin, 1984; Scheiter & Gerjets, 2007). Our findings are particularly important for instructional designers because more and more educational curricula use forms of on-demand education, in which learners can plan their own learning trajectory by choosing from authentic, real-life tasks. Our study shows that those curricula should ensure that available learning tasks are sufficiently different from each other and, most importantly, teachers should provide learners with selection options that ensure an optimal selection.

5

Learner Controlled Selection of Tasks with Different Surface and Structural Features: Effects on Transfer and Efficiency[5]

Surface task features are more salient than structural task features and thus easier to recognize for novice learners. The more salient the task features the better learners can choose personally relevant and varied tasks, which enhances transfer of learning. A 2 x 2 factorial experiment with 72 participants studied the effects of control over tasks with different surface features (learner control, program control) and different structural features (learner control, program control). Learner control over the selection of tasks with salient surface features enables learners to select personally relevant and varied tasks. This is believed to yield higher effectiveness (i.e., higher near and far transfer test performance) as well as higher efficiency (i.e., higher transfer test performance combined with lower associated mental effort). Learner control over the selection of tasks with non-salient structural features does not enable learners to select personally relevant and varied tasks and is therefore not expected to yield beneficial effects on learning. The results confirm the expected positive effects of learner control over the selection of tasks with salient surface features for efficiency on the far transfer test but not for effectiveness. Theoretical and practical implications are discussed.

---

[5] This chapter is based on: Corbalan, G., Kester, L., & van Merriënboer, J.J.G. (2007). *Learner controlled selection of tasks with different surface and structural features: Effects on transfer and efficiency*. Manuscript submitted for publication.

Learner controlled instruction gives learners the opportunity to make selections according to their current knowledge, interests, and preferences (Merrill, 1980; van Merriënboer, Schuurman, de Croock, & Paas, 2002). This is believed to positively influence learning and motivation (Flowerday & Schraw, 2000; Schnackenberg & Sullivan, 2000). Research shows that learner control is beneficial for learning even when choices are trivial, such as choosing the context in which the task is presented or the density of the text in which it is written (Cordova & Lepper, 1996; Kinzie, 1990; Lepper, 1985). So, merely the perception of control might be enough for learner control to work (Savage, Perlmunter, & Monty, 1979; Skinner, 1996). However, studies report both beneficial and detrimental effects of learner control on learning (e.g., Katz & Assor, 2007; Kopcha & Sullivan, in press; Williams, 1996). It seems that learner control functions differently depending on *what* (e.g., pace, display, task features) is being controlled by *whom* (e.g., novice or experienced learners), and only works if learners recognize the control that is given to them (Morrison, Ross, O'Dell, & Schultz, 1988; Scheiter & Gerjets, 2007).

The aim of this study is to investigate under which conditions the effects of learner control over the selection of learning tasks are optimized. More specifically, it investigates whether learner control is more effective and efficient when learners select tasks on the basis of their surface features or on the basis of their structural features. In this introduction, first the potential effects of providing learners with control over the selection of tasks with different surface and structural features are described, and second the potential overloading effects of an excessive amount of control over task selection are discussed.

Learner Controlled Selection of Tasks with Different Surface Features

Surface task features refer to task aspects that are not relevant to reach a solution (e.g., species in inheritance tasks because Mendel's laws are the same for animals, plants, and humans). They generally are *salient* even for domain novices (Chi, Feltovich, & Glaser, 1981; Cummins, 1992; Gick & Holyoak, 1987; Quilici & Mayer, 1996, 2002). For example, in inheritance tasks, novice learners will distinguish tasks dealing with a cat's eye colour from tasks dealing with a pea plant's flower shape (i.e., surface task features). The saliency of the surface task features makes it easier to perceive control over the selection of tasks that differ from each other on these features (Corbalan, Kester, & van Merriënboer, in press-b). In other words, learners who are aware of the different surface features of the tasks they may select to perform, will probably be more aware of their motives for choosing between them.

Since learner control over the selection of tasks with different surface features is likely to be perceived by the learners, they may use it to make instruction more personally relevant for them (Katz & Assor, 2007). The chosen tasks facilitate them to connect new information to their prior knowledge without affecting the way the task is solved (Hanaffin, 1984; Kinzie, 1990; Ross & Morrison, 1989; Wouters, Tabbers, & Paas, 2007). This promotes elaboration, a process in which existing knowledge is used as an assimilative context (or 'schema') to integrate new information in (van Merriënboer, Kirschner, & Kester, 2003). In a study carried out by Ross, Morrison, and O'Dell (1989), students in a learner control group who could choose between four task themes (i.e., sports, medical, educational, abstract) to learn statistics performed better than students in a control group who were given standard themes. In another study carried out by Cordova and Lepper (1996), participants who could generate the names of their spacecraft and their opponent in a computer game designed to teach arithmetical and problem solving skills outperformed participants who received predetermined names. The authors argued that this control over surface features probably increased the personal relevance of the task.

In addition, learners who recognize the different surface features of the tasks they perform may use these features to select a varied set of tasks. Variations in surface features of tasks that share the same solution steps help learners see beneath surface features and recognize the solution steps of a task, enhancing schema induction and thus fostering transfer of learning (Chen & Mo, 2004; Shute & Gawlick, 1995; Quilici & Mayer, 1996). So, learner controlled selection of tasks with different surface features facilitates learning if learners vary their choice of surface task features (Morrison, Ross, & Baldwin, 1992; Ross & Morrison, 1989; Tennyson & Buttery, 1980). A study in which learners could choose between pedagogical agents with five different ethnicities showed that students who chose to learn with agents of *different* ethnicities had higher transfer test scores than those who chose to learn with agents of the same ethnicity (Moreno & Flowerday, 2006).

To sum up, learner controlled selection of tasks that differ in surface features fosters learning because it enables learners to select a varied set of personally relevant tasks which enhances, in order, induction and elaboration and, eventually, transfer of learning. In this study, we characterize surface features as the *contexts* (i.e., different species and traits) in which a series of inheritance tasks is presented.

Learner Controlled Selection of Tasks with Different Structural Features

Structural task features refer to task aspects that are directly relevant to reach a solution (e.g., a solution step such as 'determine the genotype of a parent' in an inheritance task) (Chen & Mo, 2004; Cummins, 1992; Gick & Holyoak, 1987; Holyoak & Koh, 1987; Novick, 1988; Quilici & Mayer, 1996, 2002; Ross, 1989). In contrast to surface task features, structural task features are often less salient, or not salient at all, for novice learners. In inheritance tasks, for example, it is difficult for novice learners to distinguish tasks in which they are required to 'determine the genotype of a parent' from tasks in which they are required to 'determine the phenotype of a parent'. Consequently, it is difficult for novice learners to perceive the control they are given over the selection of tasks with different structural features. Learners who do not recognize the non-salient structural task features, will not be aware of any valid motives for choosing between tasks that differ in these features and are thus unable to distinguish between tasks that are necessary for learning and tasks that could just as well be omitted (Ross & Morrison, 1989). This could negatively influence the learning process (Chung & Reigeluth, 1992; Tennyson & Buttery, 1980; Williams, 1996) and even enlarge individual differences between low and high ability learners (Kopcha & Sullivan, in press; Merrill, 2002; Snow, 1980).

Since learner control over the selection of tasks with different structural features will not be perceived as such by the learners, it is practically impossible for them to select a varied set of personally relevant tasks. Consequently, they are not likely to profit from the effects of elaboration and induction on transfer of learning. Moreover, they are also not likely to profit from variation in practicing tasks with different *structural* features which is also considered to have positive effects on learning (e.g., Holladay & Quiñones, 2003). When - in a sequence of practiced tasks - the structural task features vary, repetition of specific structural task features occurs at longer intervals and learners have to retrieve the appropriate schema each time a feature (e.g., a specific solution step) needs to be performed. This variation in the practice of structural task features may result in reconstructive activities that will eventually yield more accessible representations in memory, with beneficial effects on learning (Lee & Magill, 1985; Reder & Klatzky, 1994; van Merriënboer, Kester, & Paas, 2006).

To sum up, learner controlled selection of tasks with different structural features does not foster learning because learners do not recognize the structural task features and therefore are not able to select a varied set of personally relevant tasks. Consequently, this type of control is not expected to enhance elaboration, induction and, eventually, transfer of learning. In this study, we characterize structural task features as the procedural solution steps learners must complete to solve inheritance tasks.

Learner Control over Task Selection and Cognitive Overload

The effectiveness of learner control is influenced by the amount of choice learners have. A too high amount of choice may cause cognitive overload (Borsook & Higginbotham-Wheat, 1991; Scheiter & Gerjets, 2007; Schwartz, 2004). Even expert learners experience difficulties in selecting, sequencing, and pacing huge amounts of information (Scheiter & Gerjets, 2007). This problem could be solved when an instructional agent (e.g., a teacher or a computer program) and the learner share control over the process of task selection. In this two-step process, a computer program first selects a subset of learning tasks with desirable task features (e.g., surface and structural features) based on task features of previously selected tasks (program control). Second, the learner selects from this subset one task to work on (learner control). This avoids the potential pitfall of a too high amount of choice and yet grants some learner control (Corbalan et al., 2006, in press-a; Tennyson & Buttery, 1980). This study implements shared control to prevent cognitive overload by reducing the amount of choice given to learners.

The purpose of the present study is to investigate the effects of learner controlled selection of tasks with different surface features (i.e., the context) and structural features (i.e., the to-be-completed solution steps) on learning effectiveness (i.e., transfer test performance) and efficiency (i.e., transfer test performance in relation to the mental effort invested to reach this performance). It is hypothesized that learners profit from learner controlled selection of tasks that differ in their *surface* features, because the saliency of those task features enables learners to select a varied set of personally relevant tasks which fosters learning and transfer. In contrast, learners will not profit from learner controlled selection of tasks that differ in *structural* features, because the non-saliency of those features impedes learners to select a varied set of personally relevant tasks. Therefore, learners who control the selection of tasks that differ in surface features are expected to achieve a higher transfer test performance and to show higher efficiency than learners who receive tasks which are selected by the program on the basis of the surface features. In addition, with regard to structural features we expect no differences between learner and program control over the selection of tasks with different structural features.

Method

Participants

Seventy-two first year students (61 females and 11 males; mean age = 17.33 years; *SD* = 1.07) in the Health Sciences domain of a Dutch school for secondary vocational education participated in this study. They received €20 (approximately $27) for their participation. A 2 x 2 factorial design was used to study the effects of control over the selection of learning tasks in the genetics domain that differ in surface features (program control vs. learner control) and structural features (program control vs. learner control). The students were randomly assigned to one of the four experimental groups: in the 'program surface, program structural' condition (n = 18) the program fully controlled task selection based on both surface task features and structural task features; in the 'program surface, learner structural' condition (n = 17) the program controlled task selection based on surface task features and the learners controlled task selection based on structural task features; in the 'learner surface, program structural' condition (n = 17) the learners controlled task selection based on surface task features and the program controlled task selection based on structural task features; and in the 'learner surface, learner structural' condition (n = 20) the learners fully controlled task selection based on both the surface task features and the structural task features.

Materials

Electronic learning environment. The learning environment especially developed for this study was a web application written in the web scripting language PHP. A MySQL database connected to the learning environment contained a basic introduction to the domain of genetics, a prior factual knowledge test, all learning tasks, a perceived control questionnaire, a transfer test, and mental effort measurements.

Basic introduction. The basic introduction included the main concepts of the domain of genetics necessary to begin the training (i.e., dominant and recessive genes, homozygous and heterozygous gene pairs, genotype and phenotype) and a worked-out example containing all the solution steps of a representative inheritance task.

Prior factual knowledge test. This test contained eight multiple-choice questions and assessed participants' prior factual knowledge. The maximum test score was eight points.

Learning tasks. The learning environment was connected to a database which contained 54 completion tasks in the genetics domain, more specifically inheritance tasks (e.g., inheritance of the hair, the fur, or the leaf colour). Completion tasks present a given state, a goal state, and a partial solution that learners must complete by adding the missing *solution steps* (van Merriënboer, 1997; van Merriënboer & Kirschner, 2007). The learning tasks varied in: (a) the contextual features species and traits, which were the surface task features, and (b) the to-be-completed solution steps, which were the structural task features. Table 5.1 shows the surface features and structural features of the learning tasks. The first and second columns contain species (e.g., animal, plant) with species types (e.g., cat, pea) and traits (e.g., colour, shape) with trait parts (e.g., fur, tail). In addition, the third column describes the seven solution steps that are necessary to reach the solution for an inheritance problem. Steps five, six and seven appeared two times in each task. Each completion task contained three to-be-completed solution steps and four steps that were already completed. For instance, if steps one, three, and four had to be completed by the learner, the right solution of the four remaining solution steps, that is, steps two, five, six, and seven was presented on the computer screen.

Table 5.1
Composition of the Database with Learning Tasks

Surface Task Features		Structural Task Features
Species (species type)	Traits (part traits)	Solution Steps To-be-Completed
Humans (European/African/Asian)	Colour (hair/eyes) Shape (hair/nose) Length (nose/lips)	Determine the genotype of one parent based on information of the individual given Determine the genotype of one parent based on the given percentage in his/her generation
Animals (dog/cat/guinea pig)	Colour (fur/eyes) Shape (ear/fur) Length (tail/fur)	Determine the genotype of one of the offspring of the first generation Determine the genotype based on the information of the prior partner and related offspring Draw a Punnett's square by combining the genotype of the parents
Plants (pea/corn/bean)	Colour (flower/leaf) Shape (fruit/pod) Length (axis/fruit)	Determine the genotype of the offspring and calculate their percentage Determine the phenotype of the offspring and calculate their percentage

Each participant completed twelve learning tasks. Each correctly completed step scored one point, except for steps five, six and seven, which appeared twice in each task. For these steps each correct answer scored a half point, leading to a total of one point per step. This led to a maximum score of 3 points per learning task and of 36 points for the whole training phase. The reliability of the scores for the learning tasks was .92 (Cronbach's alpha).

Task selection. The first (subset of) task(s) was/were randomly selected by the program. Subsequently, each (set of) learning task(s) presented to the learners was/were dynamically selected from the learning-tasks database and differed on the surface and the structural task features, selected either by the learner or by the program. Tasks varied depending on how dissimilar surface task features and structural task features were as compared to the previous task. Regarding the *surface task features,* two levels of dissimilarity were distinguished: (1) low dissimilarity, which contained tasks with either (a) same species, species type, trait, and trait part, or (b) same species and trait but different species type and trait part, and (2) high dissimilarity, which contained tasks with either (a) different species or trait, and different species type and trait part, or (b) different species and trait and different species type and trait part. Regarding the *structural task features,* also two levels of dissimilarity were distinguished: (1) low dissimilarity, which contained tasks with either (a) zero, or (b) one to-be-completed solution steps different from the previous task, and (2) high dissimilarity, which contained tasks with (a) two, or

(b) three to-be-completed solution steps different from the previous task. When the surface task features were selected by the *program, one* task with high dissimilarity was presented. When the surface task features were selected by the *learner, four* tasks including each combination of low and high dissimilarity levels were presented, from which the learner selected one. When the structural task features were selected by the *program, one* task with high dissimilarity was presented. When the structural task features were selected by the *learner, four* tasks including each combination of the low and high dissimilarity levels were presented, from which the participant selected one.

Transfer test. The transfer test consisted of eight transfer tasks, divided in four near transfer tasks and four far transfer tasks. The near transfer tasks were structurally similar to the learning tasks but contained different surface features (i.e., other subjects within the species, for example, fruit flies, and other traits, for example, position of the wings) and determined whether participants were able to apply the learned procedures in the same way as in the training tasks. The far transfer tasks required participants to flexibly use the learned solution procedures during training to structurally different tasks. More specifically, the following far transfer tasks were used: (a) a dihybrid crossing task which required two different traits to be treated separately; (b) a family tree task in which participants had to infer the genotype of several of the individuals based on the information given in the tree; (c) a task in which participants had to infer the genotype of an individual from information of the father; and (d) a task with co-dominant genes, that is, genes that are equally strong and both expressed. The maximum score on the near transfer test and on the far transfer test was four points each. The reliability of the tests were, in order, .88 and .74 (Cronbach's alpha) for the near and far transfer test.

Mental effort. Mental effort was used as an index for cognitive load, which refers to the amount of cognitive capacity that is allocated to problem solving. Mental effort was measured after each learning task and after each transfer test task with a one-item 7-point rating scale (Paas, 1992; Paas, Tuovinen, Tabbers, & van Gerven, 2003). Reliability of the mental effort measures reported during the training was .98 and during the transfer test .91 and .86 (Cronbach's alpha) for the near and far transfer test, respectively.

Efficiency. Participants' test performance on near and far transfer and mental effort invested during the performance of those tests were combined using the procedure of Paas and van Merriënboer (1993) to calculate efficiency (E) on near and far transfer. Performance and mental effort scores are first standardized, and then the *z*-scores are entered into the formula:

$$E = \frac{Z_{Performance} - Z_{MentalEffort}}{\sqrt{2}}$$

In a two-dimensional space defined by the standardized test performance and mental effort scores, efficiency is computed for each condition as the perpendicular distance between a point representing the condition (i.e., the z-score for transfer Performance and the z-score for Mental Effort) and the diagonal, E = 0, where Performance and Mental Effort are proportionally related to each other. When performance is higher than might be expected based on perceived mental effort, the instructional condition is more efficient. Conversely, when performance is lower than might be expected based on perceived mental effort, the instructional condition is less efficient.

Perceived control. After the 12 learning tasks, participants completed 4 items with a 7-point scale designed to rate self-reports of perceived control on task selection. The four items were: "*I was able to choose the inheritance task I wanted to perform", "I could decide by myself what I wanted to learn about solving inheritance tasks*", "*I could decide by myself how I wanted to learn about solving inheritance tasks*", and "*I could decide by myself which information about the inheritance tasks I wanted to consult*". Reliability of the perceived control questionnaire was .89 (Cronbach's alpha).

Time logging. The learning environment kept track of the time (in seconds) participants needed to complete the training tasks and the transfer test tasks.

Procedure

In the *pre-training phase*, participants received the basic introduction and completed the prior factual knowledge test. Subsequently, participants started the *training phase.* Participants were not informed on how the tasks were selected or pre-selected (for the program control and learner control conditions, respectively). In each learning task, participants could press a continue button after a to-be-completed solution step was solved. The remaining solutions steps until the next to-be-completed solution step appeared. After each training task, mental effort was measured by asking the learner to fill out the 7-point rating scale. Participants could access the basic information from the introduction at all times by pressing a button labelled 'basic information' always visible on the left-hand side of the screen. It was emphasized that they were not allowed to skip any part of the answer of the tasks and cognitive load questions and that in such a case the program would prompt them to answer the questions before they could continue. After the training phase was completed, participants started the *test phase,* in which they completed the perceived control questionnaire and the transfer test. After each transfer task, mental effort was measured with the 7-point rating scale. During the test phase, the 'basic information' button disappeared from the screen. Participants were allowed to work at their own pace. The times spent during the training and transfer phase were logged.

Results

An ANOVA on the prior factual knowledge test revealed no differences between conditions ($F(3, 68) < 1$). Therefore, all subsequent analyses are performed using ANOVAs with the between-subjects factors control over task selection based on surface task features (program control, learner control) and control over task selection based on structural task features (program control, learner control). For all statistical tests a significance level of .05 was maintained. Table 5.2 provides an overview of the mean scores and standard deviations for the prior factual knowledge test and the dependent variables during the training phase and transfer phase.

Table 5.2
Overview of Results from the Prior Factual Knowledge Test, the Training Phase, and the Test Phase

	Program surface				Learner surface			
	Program structural (n = 18)		Learner structural (n = 17)		Program structural (n = 17)		Learner structural (n = 20)[a]	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Prior Factual knowledge test	4.00	1.78	4.41	1.12	4.06	1.14	4.65	1.63
Training Phase								
Time (sec.)	3135	945.13	3006	816.18	2879	651.70	2666	783.21
Performance (max. = 36)	20.42	8.63	23.09	8.45	16.76	9.25	19.08	9.11
Mental Effort (max. = 7)	4.51	1.68	4.97	1.03	4.76	1.68	4.32	1.47
Transfer Phase								
Time (sec.)	2564	744.28	2170	832.24	2341	1025.25	2290	1024.33
Performance Near Transfer (max. = 4)	1.67	1.44	1.62	1.43	1.71	1.33	1.72	1.14
Performance Far Transfer (max. = 4)	1.39	.94	1.16	.76	1.46	1.22	1.48	1.13
Mental Effort Near Transfer (max. = 7)	5.01	1.76	5.50	1.13	4.35	1.58	4.35	1.39
Mental Effort Far Transfer (max. = 7)	5.53	1.29	5.85	.72	4.68	1.41	4.59	1.53
Efficiency								
On Near Transfer	-.11	1.54	-.36	1.19	.22	1.31	.23	1.05
On Far Transfer	-.19	1.20	-.51	.71	.30	1.34	.37	1.31

[a]n = 19 in the Transfer Phase

Transfer Phase

One participant in the 'learner surface, learner structural' condition had to leave the session earlier and did not perform the transfer test.

Test time. No effects were found on the time spent during training for control over task selection based on surface task features, $F(1, 67) = .06$, $MSE = 47363.26$, *ns*; control over task selection based on structural task features, $F(1, 67) = 1.06$, $MSE = 890229.32$, *ns*, and their interaction, $F(1, 67) = .62$, $MSE = 526428.38$, *ns*.

Test performance. Analyses on the near transfer test revealed no statistically significant main effects for control over task selection based on surface task features, $F(1, 67) = .05$, $MSE = .09$, *ns*; control over task selection based on structural task features, $F(1, 67) = .003$, $MSE = .006$, *ns*, and the interaction between these factors, $F(1, 67) = .009$, $MSE = .02$, *ns*. Similarly, on the far transfer test no effects were found for control over task selection based on surface task features, $F(1, 67) = .67$, $MSE = .71$, *ns;* control over task selection based on structural task features, $F(1, 67) = .17$, $MSE = .18$, *ns*, and their interaction, $F(1, 67) = .27$, $MSE = .29$, *ns*.

Mental effort. A significant main effect of control over task selection based on surface task features was found on mental effort invested during the near transfer test, $F(1, 67) = 6.65$, $MSE = 14.71$, $p < .025$, $\eta^2_p = .09$. Participants in the 'learner surface' conditions experienced lower mental effort during the near transfer test ($M = 4.35$, $SD = 1.46$) than participants in the 'program surface' conditions ($M = 5.25$, $SD = 1.49$). No effects on mental effort during the near transfer test were found for control over task selection based on structural task features, and the interaction between control based on surface features and control based on structural features. For the mental effort during the far transfer test, again a significant main effect of control over task selection based on surface features was found, $F(1, 67) = 12.01$, $MSE = 19.75$, $p < .010$, $\eta^2_p = .15$. Participants in the 'learner surface' conditions experienced lower mental effort during the far transfer test ($M = 4.63$, $SD = 1.45$) than participants in the 'program surface' conditions ($M = 5.69$, $SD = 1.05$). No effects on mental effort during the far transfer test were found for control over task selection based on structural task features, or the interaction between control based on surface features and control based on structural features.

Efficiency. No main effects on the efficiency of the near transfer test were found for control over task selection based on surface task features, $F(1, 67) = 2.25$, $MSE = 3.72$, *ns*; control over task selection based on structural task features, $F(1, 67) = .16$, $MSE = .26$, *ns*; and the interaction between those factors, $F(1, 67) = .18$, $MSE = .29$, *ns*. With regard to the efficiency of the far transfer test, a significant main effect of control over task selection based on surface task features was found, $F(1, 67) = 6.02$, $MSE = 8.32$, $p < .025$, $\eta^2_p = .08$. Efficiency was higher in the 'learner surface' conditions ($M = .34$, $SD = 1.31$) than in the 'program surface' conditions ($M = -.35$, $SD = .99$). No effects on efficiency of the far transfer test were found for control over task selection based on structural task features, and the interaction of control based on surface features and control based on structural features.

Training Phase

Due to technical difficulties, 8 participants had a missing mental effort value on one of their training tasks. The missing value was replaced by the participant's own mean mental effort computed over the whole training phase.

Training time. No effects on training time were found for control over task selection based on surface task features, $F(1, 68) = 2.45$, $MSE = 1598754.36$, *ns*; control over task selection based on structural task features, $F(1, 68) = .81$, $MSE = 524607.13$, *ns*; and their interaction, $F(1, 68) = .05$, $MSE = 651385.83$, *ns*.

Training performance. An ANOVA revealed no main effects on training performance for control over task selection based on surface task features, $F(1, 68) = 3.34$, $MSE = 263.24$, *ns*; control over task selection based on structural task features, $F(1, 68) = 1.41$, $MSE = 111.20$, *ns*; and their interaction, $F(1, 68) = .01$, $MSE = .59$, *ns*.

Mental effort. Analyses revealed no significant main effects on the mental effort invested during training for control over task selection based on surface task features, $F(1, 68) = .31$, $MSE = .70$, *ns;* control over task selection based on structural task features, $F(1, 68) = .001$, $MSE = .002$, *ns*; and the interaction between those factors, $F(1, 68) = 1.62$, $MSE = 3.63$, *ns*.

Learner's selection of tasks with different surface and structural features. The level of dissimilarity of the tasks selected by participants in the 'learner *surface*, program structural' condition ($n = 17$) was calculated and compared to the level of dissimilarity of the tasks selected by participants in the 'program surface, learner *structural*' condition ($n = 17$). Each selected task received zero, one, two, or three points corresponding to the low and high levels of dissimilarity for both surface and structural task features as described above. For example, a task with a different species and trait and a different species type and trait part would score 3 points because it corresponds to the high dissimilarity level of the surface task features. Similarly, a task containing one different step would score one point because it corresponds to the low dissimilarity level of structural task features. This led to a dissimilarity level of minimally zero points and maximally 33 points (maximally three points times 11 tasks). A *t*-test showed a significant difference between the level of dissimilarity of the tasks selected by the learners on the basis of their surface features and the tasks selected on the basis of their structural features, $t(29.44) = 4.94$, $p < .001$, $d = 1.68$ (which indicates a large effect size). The level of dissimilarity of the tasks learners selected on the basis of their surface task features ($M = 19.78$, $SD = 4.46$) was much higher than the level of dissimilarity of the tasks learners selected on the basis of their structural task features ($M = 10.88$, $SD = 6.03$).

Perceived Control Questionnaire

Table 5.3 provides an overview of the mean scores and standard deviations of the perceived control questionnaire.

Table 5.3
Overview of Results from the Perceived Control Questionnaire

	Program surface				Learner surface			
	Program structural (n = 18)		Learner structural (n = 17)		Program structural (n = 17)		Learner structural (n = 20)	
	M	SD	M	SD	M	SD	M	SD
Perceived control	2.97	1.50	5.94	.94	5.91	1.10	5.94	1.15

A one-way ANOVA revealed an overall effect of condition on perceived control ($F(3, 68) = 26.88$, $MSE = 39.39$, $p < .001$). Post-hoc tests, using Tukey's HSD, indicated that only the 'program surface, program structural' condition differed from the other three conditions on perceived control (all p's < .001). As indicated in Table 5.3, participants in the fully program-controlled condition reported significantly lower perceived control than participants in the other conditions.

Discussion

The goal of this study was to investigate the effects of learner controlled selection of learning tasks that differ in surface task features and/or structural task features on learning effectiveness and efficiency. It was expected that learner controlled selection of tasks with different surface features would enhance learning as compared to program controlled task selection. This hypothesis was supported for efficiency on the far transfer test. With regard to effectiveness, transfer test performance was not higher for students who selected tasks with different surface features, although their mental effort scores were in the expected direction. That is, it took them significantly less effort to perform the near and far transfer test. Thus, it is better to give learners the freedom to select tasks with their preferred surface task features. Learners who are given this type of control probably construct general cognitive schemas which enable them to flexibly apply the learned solution procedure to solve unfamiliar inheritance tasks. Furthermore it was expected that learner controlled selection of tasks with different structural features would not enhance learning. Indeed, no differences were found for invested mental effort, transfer test performance, and efficiency between learners who could select tasks with different structural features and learners who could not.

Although participants in the 'learner surface' conditions invested less mental effort to reach a similar performance on the near transfer test than participants in the 'program surface' conditions, no differences between these groups were found for efficiency on the near transfer test. A possible explanation for the higher efficiency on the far transfer test but no higher efficiency on the near transfer test concerns the general information available in the schemas constructed. This general information is particularly useful to deal with tasks that require learners to *flexibly* apply the learned solution procedure, but is of less use for familiar tasks that require learners to apply the learned solution procedure similarly to the practiced tasks (Kester, Kirschner, & van Merriënboer, 2006; Sweller, van Merriënboer, & Paas, 1998). Additionally, participants in all conditions were required to direct their attention towards the solution steps. Hence, all participants had the opportunity to compile problem-specific schemas, but only participants in the 'learner surface' conditions, who were encouraged to see the connections between tasks, were enabled to construct generalized schemas.

The main hypothesis of this study was based on the assumption that learner control has beneficial effects on learning, provided that it is perceived by the learners and actually used to select a varied set of personally relevant tasks. Results on the perceived control questionnaire show that participants in the three conditions with some form of learner control reported a similar perceived control over task selection. It was assumed that learners would recognize surface task features more easily than structural task features and therefore perceive control over task selection based on surface features more easily than control over task selection based on structural features. This is not supported by the results on perceived control. A possible explanation is that in all three conditions with some form of learner control participants were required to select one task from a set of four tasks that was pre-selected by the program. The selection screen thus always showed four tasks, presented in two rows of two tasks. Each task contained a description of the surface features and a description of the to-be-completed solution steps. So, even participants in the conditions in which the control over task selection was based on the structural features must have noticed that they had a choice even if they did not recognize the structural task features. In future studies, the perceived control questionnaire should specifically address if learners recognize the surface task features *or* the structural task features – and it should also be studied which information is precisely used by the learners to select their next task.

Although the perceived control questionnaire does not provide specific information on the perception of surface and structural task features, it does show that participants who could exercise control also perceived this control. And as predicted, only participants in the 'learner surface' conditions profited from this control. Apparently these participants were better able to select personally relevant and varied tasks. The analysis of which tasks were selected by the participants confirms this. These results clearly show that participants chose much more varied tasks when they selected them on the basis of their surface features than when they selected them on the basis of their structural features.

A final finding that needs to be discussed is why learner controlled selection of tasks based on structural features, which might yield low variation because these features are not salient for the learners, did not yield inferior learning than program controlled selection of tasks based on structural features, which by definition presented tasks with dissimilar task features (i.e., high variation). A possible explanation is provided by Gick and Holyoak (1987). The authors hold that initial exposure to relatively similar learning elements (e.g., high similarity on to-be-completed solution steps) helps establish generalized rules and more dissimilar elements should only be used to elaborate the rule set once the initial rules have been firmly established and strengthened. Accordingly, early practice of tasks with similar structural features, followed by subsequent exposure to more variable practice, is expected to optimize learning and transfer. In our study, such a sequence occurred in neither the ‘learner structural’ nor the ‘program structural’ conditions.

The results of our study have several implications for future research. First, to gain better insight in the way people experience learner control and use it to construct cognitive schemas, more direct measures of their mental processes such as verbal protocols, retrospective reports, and eye-tracking data should be gathered during training. Second, the effects on transfer performance should be examined over a more extended period of time because transfer may not be apparent immediately after practice, but may be present at a later time if the same or additional transfer tasks are repeated (Gick & Holyoak, 1987). Third, it seems plausible that learners with a higher level of expertise are better able to recognize structural task features than novice learners; thus, differences between novice and experienced learners should be studied with regard to their ability to select a varied set of personally relevant tasks on the basis of structural features. Fourth, future studies might investigate if supporting learners in recognizing - surface and structural - task features, for example by explicit instruction to compare different tasks or by giving them feedback on their task selections, would yield better results. Fifth, learners’ perceptions of the personal relevance of tasks and their motives for the selections they make should be assessed in more detail. Finally, it should be investigated to what extent practice of tasks with dissimilar surface features enhances transfer to a completely new domain (e.g., Chen & Mo, 2004).

To conclude, this study showed that the effectiveness and efficiency of learner control clearly depends on *what* this control is based on: learner control based on surface features is beneficial while learner control based on structural features is not. This finding is particularly important for instructional designers because educational curricula increasingly use forms of on-demand education, in which learners plan their own learning trajectory.

6

Dynamic Task Selection: Effects of Feedback and Learner Control on Transfer, Efficiency, and Motivation[6]

Structural features of learning tasks are relevant for problem solving but not salient for novice learners. Feedback in the form of Knowledge of Correct Response (KCR) during practice is expected to help learners recognize the structural features and to profit from learner control over the selection of learning tasks. A 2 x 2 factorial experiment (*N* = 118) was conducted to study the effects of KCR (present, absent) and control over the selection of learning tasks (learner control, program control). The presence of KCR yielded higher performance and efficiency on a near transfer test as well as higher learner motivation. An interaction between feedback and control, indicating extra beneficial effects of feedback when learners control the selection of learning tasks, was not found. Theoretical and practical implications are discussed.

---

[6] This chapter is based on: Corbalan, G., Kester, L., & van Merriënboer, J.J.G. (2008). *Dynamic task selection: Effects of feedback and learner control on transfer, efficiency, and motivation.* Manuscript submitted for publication.

Recent instructional theories advocate on-demand methods of education which give learners freedom to choose their own learning path (Hannafin, 1984; Williams, 1996). Optimal learner control allows learners to make selections according to their current knowledge, interests, and preferences (Merrill, 1980; van Merriënboer, Schuurman, de Croock, & Paas, 2002). This is believed to positively influence learning and motivation (Flowerday & Schraw, 2000; Schnackenberg & Sullivan, 2000). Studies report both positive and negative effects of learner control on learning (Katz & Assor, 2007; Williams, 1996). It seems that the effectiveness of learner control depends on *what* (e.g., which elements of instruction, such as pace, display or task features) is controlled by *whom* (e.g., novice or more experienced learners) and, moreover, is only realized if learners recognize the control that is given to them (Morrison, Ross, O'Dell, & Schultz, 1988; Scheiter & Gerjets, 2007).

This study investigates how learner control over the selection of learning tasks with different structural features (i.e., what) can be optimized for novice learners (i.e., whom). More specifically, it will be studied if feedback helps learners to recognize structural task features and thus enables them to select personally relevant tasks with beneficial effects on learning and motivation. The next paragraphs describe what structural task features are, how providing feedback on structural task features may facilitate learning and motivation, and how feedback might interact with different types of control over task selection.

Structural task features refer to task aspects that are necessary to reach a solution for a particular problem (e.g., solution steps in inheritance tasks or the underlying mathematical procedure in statistical problems; Chen & Mo, 2004; Novick & Holyoak, 1991; Ross, 1989; Vosniadou & Ortony, 1989). These task features are generally *not salient* for, especially, novice learners in a domain. Therefore, novice learners are not able to spontaneously distinguish between tasks that differ on structural features (Cummins, 1992; Quilici & Mayer, 1996, 2002) and are also not able to strategically select learning tasks that differ from each other on their structural features. In this study, we characterize structural task features as the solution steps learners must complete to solve inheritance tasks. In such tasks, for example, it will be difficult for novices to distinguish tasks for which they are required to "determine the genotype of a parent" from tasks for which they are required to "determine the phenotype of a parent". Obviously, the lack of saliency of structural features causes problems when learners are required to select their own learning tasks: they might unknowingly select equivalent tasks over and over again and not select structurally different tasks that also need to be practiced.

If learners do not recognize non-salient structural task features this may have negative effects on both learning and motivation. With regard to learning, learners will not be aware of any valid motives for choosing between tasks that differ in these features and are thus unable to distinguish between tasks that are necessary for learning and tasks that could just as well be omitted (Ross & Morrison, 1989). This will negatively influence the learning process (Chung & Reigeluth, 1992; Kopcha

& Sullivan, in press; Tennyson & Buttery, 1980; Williams, 1996) and even enlarge individual differences between low and high ability learners (Merrill, 2002; Snow, 1980). With regard to motivation, learners who are not aware of structural differences between tasks will probably not see the meaning of choosing between these tasks. For them, all tasks look the same which will negatively influence their motivation. Furthermore, these learners will not be able to choose personally relevant tasks. Personally relevant instruction facilitates learners to connect new information to their prior knowledge (Hanaffin, 1984; Ross & Morrison, 1989; Wouters, Tabbers, & Paas, 2007), with positive effects on transfer of learning (van Merriënboer, Kirschner, & Kester, 2003). Personally relevant instruction is also prerequisite for enhancing and maintaining learners' motivation (Katz & Assor, 2007).

In the current study, the provision of feedback during practice is presented as a method to counteract the negative effects of the lack of saliency of structural task features. The form of feedback used is Knowledge of Correct Response (KCR). KCR provides learners with worked-out correct solution steps, regardless of the correctness of the solution steps generated by the learner. Research has shown that practice with KCR leads to better learning than practice with no-feedback and practice with Knowledge of Response (KOR), which only states whether a response is correct or incorrect (Ross & Morrison, 1993).

KCR emphasizes structural features in the form of correct solution steps, and thus enables learners to focus their attention on those features and better recognize necessary solution steps for future tasks (Cummins, 1992; Mory, 2003; Quilici & Mayer, 1996). Structural task features that are recognized as shared aspects throughout a series of tasks promote generalization and abstraction of a common relational structure that can be stored in cognitive schemas (Loewenstein, Thompson, & Gentner, 1999). In addition, since feedback informs learners about their achievement, it gives them the opportunity to adjust and improve their cognitive strategies and to rectify misconceptions while progressing through the training (Azevedo & Bernard, 1995; Gagné et al., 1987).

Furthermore, many authors have recognized the motivational effects of feedback (e.g., Azevedo & Bernard, 1995; Chai, 2003; Gagné et al., 1987; Hyland, 2001; Keller, 1983b; Mory, 2003; Ross & Morrison, 1993). Since feedback in the form of KCR helps learners to focus on structural task features or solution steps, it will eventually promote the perceived *relevance* of the learning material because it enables learners to see the connection between what they need to learn and the learning opportunities presented to them (Keller, 1983b, 1987; Margueratt, 2007). To sum up, the provision of KCR emphasizing structural task features fosters learning and motivation because it enables learners to recognize the structural features in future tasks and to relate those tasks to what they already know.

Since KCR helps learners to recognize the structural task features it makes these features more salient to the learners. So if learners who have control over the selection of tasks with different structural features are also given KCR on structural aspects, they will be better able to perceive the control given to them, and thus to choose personally relevant tasks. Consequently, the beneficial effects of providing KCR on learning and motivation may be higher in combination with learner controlled selection of learning tasks than with program controlled selection of learning tasks.

Concluding, this study investigates the effects of KCR when learners work on a series of learning tasks with different structural features (i.e., to-be-completed solution steps) on transfer, efficiency, and motivation. KCR helps the learners to recognize structural features of tasks and to build generalized and abstract cognitive schemas. Therefore, learners who are provided with KCR are expected to show higher transfer of learning, higher efficiency (i.e., higher transfer test performance combined with less mental effort to reach it), and higher motivation because they are better able to see the connection between what they need to know and the presented tasks. In addition, if feedback makes the structural task features more salient for learners and helps them to select personally relevant learning tasks, the beneficial effects of providing KCR on learning and motivation may be higher in combination with learner controlled selection of learning tasks than with program controlled selection of learning tasks.

Method

Participants

First-year students (N = 118; 93 females and 25 males; mean age = 18.73 years; SD = 4.67) enrolled in the Health Science program of a Dutch school for secondary vocational education participated in this study. In order to make participation attractive, they took part in a lottery making them eligible to win one of 20 cinema tickets. A 2 x 2 factorial design was used to study the effects of control over the selection of learning tasks on the basis of their structural features (program control, learner control) and feedback in the form of Knowledge of Correct Response (KCR; present, absent). Students were randomly assigned to one of the four experimental groups: program control/KCR (n = 30); program control/no-KCR (n = 29); learner control/KCR (n = 30), and learner control/no-KCR (n = 29).

Materials

Electronic learning environment. The learning environment especially developed for this study was a web application written in the web scripting language PHP. A MySQL database connected to the learning environment contained a basic introduction to the domain of genetics, a factual knowledge test, the learning tasks, a transfer test, mental effort measurements, a perceived relevance item, a perceived control questionnaire, and a motivational questionnaire.

Basic introduction. The basic introduction contained the main concepts in the domain of genetics included in the training (i.e., dominant and recessive genes, homozygous and heterozygous gene pairs, genotype and phenotype) and a worked-out example containing all the solution steps of a representative inheritance task.

Factual knowledge test. This test contained eight multiple-choice questions and assessed participants' prior factual knowledge about the domain of inheritance. The maximum test score was 8 points.

Learning tasks. The learning environment was connected to a database which contained 54 completion tasks in the genetics domain, dealing with the inheritance of particular features (e.g., inheritance of the color of hair, fur, or leafs). Completion tasks are learning tasks that present a given state, a goal state, and a partial solution (i.e., a number of *solution steps*) that learners have to complete by adding the missing steps (van Merriënboer, 1997; van Merriënboer & Kirschner, 2007). Each learning task could be solved following the same basic structure that comprised five solution steps, in order: (1) determine the genotype of the male parent based on the given information, that is, whether it is a homozygous or a heterozygous organism; (2) determine the genotype of the female parent based on the given probabilities in her generation; (3) draw a punnett's square by combining the genotypes of the two parents; (4) determine the genotype of the offspring and calculate the proportion; and (5) determine the phenotype of the offspring and calculate the proportion. In each completion task, three solution steps were given by the program and the remaining two solution steps had to be completed by the learner. For instance, if solution steps 2, 4, and 5 were given by the program, steps 1 and 3 had to be completed by the learner. Each participant completed 12 learning tasks. Each correctly completed solution step scored 1 point, leading to a maximum score of 2 points per learning task and 24 points for the whole training phase. The reliability of the learning tasks was .85 (Cronbach's alpha).

Participants in the learner control conditions were given three tasks, which were pre-selected by the program, to choose from. This limited amount of control avoids overloading learners as a result of a (too) great amount of choice (Corbalan, Kester, & van Merriënboer, 2006; Iyengar & Lepper, 2000; Scheiter & Gerjets, 2007; Schwartz, 2004; Tennyson & Buttery, 1980). The first learning task in the program control conditions, as well as the first set of three tasks in the learner control conditions, was randomly selected by the program. In the program control conditions, the program pre-selected three tasks containing either 0, 1, or 2 solution steps not completed by the learner in the preceding task, and then randomly selected and presented *one* task from this subset to the learners. In the learner control conditions, the program also pre-selected three tasks containing either 0, 1, or 2 solution steps not completed by the learner in the preceding task, but then presented all three tasks to the learner who made a final selection of one task to work on.

In the no-KCR conditions, no information was provided after participants completed the two required solution steps. In the KCR conditions, the correct solution of the steps completed by the learner was presented together with the given responses immediately after the learning task was finished (i.e., after the last step was completed by either the learner or the program; see Figure 6.1 for an example). Participants were prompted to compare their own responses with the correct responses given. If there was a mismatch, participants were advised to restudy the basic information, which also contained a worked-out example showing how the solution steps should be applied.

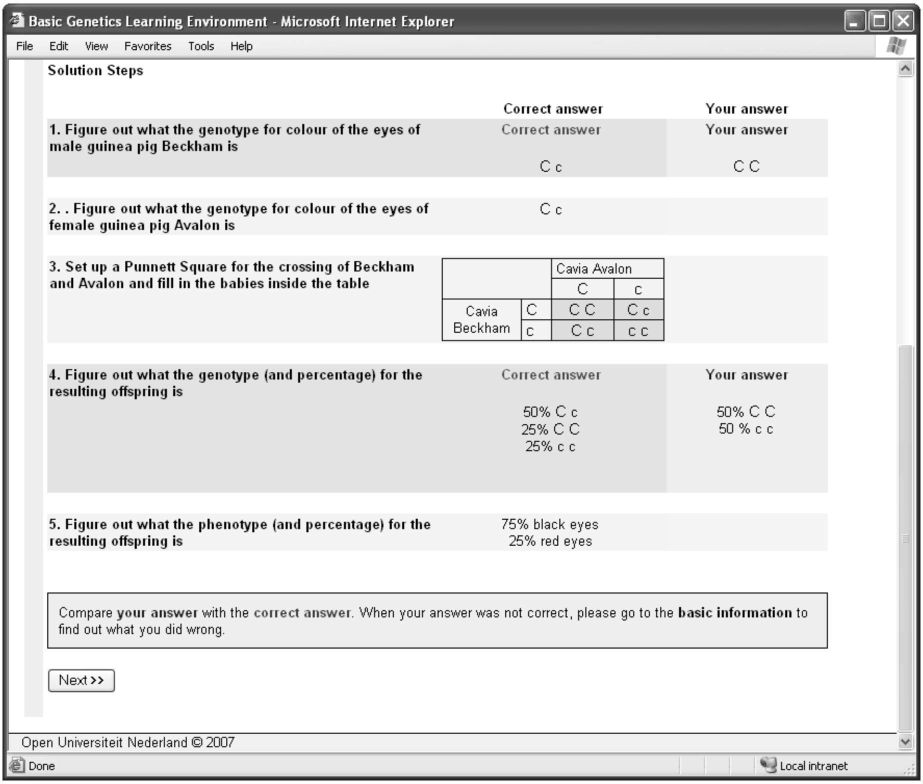


Figure 6.1. Partial screendump illustrating the completed solution steps of a training task in the KCR conditions.

Transfer test. The transfer test consisted of ten transfer tasks, divided in four near transfer tasks and six far transfer tasks. The near transfer tasks were structurally similar to the learning tasks but contained different surface features (i.e., other members within the species, for example, fruit flies, and other traits, for example, position of the wings). They determined if participants were able to apply the learned procedures in the same way as in the learning tasks. The far transfer tasks required participants to flexibly apply the learned solution procedures to structurally different tasks. More specifically, the following far transfer tasks were used: (a) determine if the baby of two parents with the same disease will also have this disease; (b) infer the genotype and phenotype of the offspring of two parents from information given of one parent and about the father of the other parent; (c) infer the genotype of several family members based on the information given in a family tree; (d) determine the genotype and phenotype of the offspring of two individuals with co-dominant genes, that is, genes that are equally strong and both expressed; (e) use the information of the phenotype of the offspring (i.e., bottom-up) and of one of the parents to find out the genotype of the other parent and of the offspring, and (f) determine the genotype and phenotype of the offspring of two individuals in a dihybrid crossing task which requires the separate treatment of two different traits. The maximum score was 4 points for the near transfer test and 6 points for the far transfer test. The reliability was, in order, .90 for the near transfer test and .80 for the far transfer test (Cronbach's alpha).

Mental effort. Mental effort reflects the amount of cognitive capacity allocated to problem solving and was used as an index for cognitive load. Mental effort was measured after each learning task and after each transfer task with a one-item 7-point rating scale (Paas, Tuovinen, Tabbers, & van Gerven, 2003). Reliability of the reported mental effort measures during training was, in order, .96, during the near transfer test .96, and during the far transfer test .94 (Cronbach's alpha).

Efficiency. Participants' transfer test performance and associated mental effort were combined using the procedure of Paas and van Merriënboer (1993) to calculate instructional efficiency (E). Performance and mental effort scores are first standardized, and then the *z*-scores are entered into the formula:

$$E = \frac{Z_{Performance} - Z_{MentalEffort}}{\sqrt{2}}$$

In a two-dimensional space defined by the standardized test performance and mental effort scores, efficiency is computed for each condition as the perpendicular distance between a point representing the condition (i.e., the *z*-score for transfer test performance and the *z*-score for mental effort) and the diagonal, E = 0, where performance and mental effort are proportionally related to each other. When performance is higher than might be expected on the basis of perceived mental effort, the instructional condition is relatively more efficient. Conversely, when performance is lower than might be expected on the basis of perceived mental effort, the instructional condition is relatively less efficient.

Perceived relevance measure. The perceived relevance of the three choices provided in the two learner control conditions was measured with a 5-point rating-scale (i.e., "*The three inheritance tasks I could choose from were relevant to my* *interests"*). Answers ranged from 'not true' (1) to 'completely true' (5). Reliability of the measure was .95 (Cronbach's alpha).

Instructional Materials Motivation Survey (IMMS). The IMMS (Keller, 1983a, see also Margueratt, 2007) assesses the motivational effects of instructional situations and asks students to rate 36 ARCS related statements (Attention, Relevance, Confidence, and Satisfaction) about the learning materials. Reliabilities of the measures were .86, .71, .91, and .89 (Cronbach's alpha) for, in order, the attention, relevance, confidence, and satisfaction scales.

Perceived control measure. A 5-point rating scale containing 4 items was designed to rate participants' perceived control over the selection of learning tasks. The four items were: "*I was able to choose the inheritance task I wanted to perform",* "*I could decide by myself what I wanted to learn about solving inheritance tasks*", "*I could decide by myself how I wanted to learn about solving inheritance tasks*", and "*I could decide by myself which information about the inheritance tasks I wanted to consult*". Reliability of the perceived control questionnaire was .67 (Cronbach's alpha).

Time logging. The learning environment kept track of the time (in seconds) participants needed to complete the learning tasks and the transfer tasks.

Procedure

In the *pre-training phase*, participants received the basic introduction and completed the prior factual knowledge test. Subsequently, participants started the *training phase.* Participants were not informed on how the tasks were selected or pre-selected (for the program control and learner control conditions, respectively). While working on a learning task, participants could press a continue button after each step, whereafter the next given or to-be-completed step appeared on the computer screen. After each learning task, mental effort and perceived relevance were measured. Participants could always access the basic information by pressing a button that was at all times visible on the left-hand side of the screen. It was emphasized that they were not allowed to skip solution steps or self-rating questions: the program would prompt them to provide an answer before they were allowed to continue. After the training phase was completed, participants started the *test phase*, in which they completed the perceived control questionnaire, the IMMS questionnaire, and the transfer test. During the test phase, the 'basic information' button disappeared from the screen. After each transfer task, mental effort was measured with the 7-point rating scale. Participants were allowed to work at their own pace. The times spent during the training phase and transfer phase were logged automatically.

Results

Table 6.1 presents the means and standard deviations of the prior factual knowledge test, the time spent on the learning tasks, and the dependent variables measured during the test phase and the training phase.

Table 6.1
Overview of Results from the Factual Knowledge Test, the Training Phase, and the Test Phase

	Program control				Learner control			
	No-KCR (n = 30)		KCR (n = 29)		No-KCR (n = 30)		KCR (n = 29)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Factual knowledge test	4.23	1.77	4.51	1.88	4.83	1.82	4.62	1.86
Time on training phase (sec.)	1823	585.19	1815	455.68	1541	553.34	1711	471.71
Test Phase								
Test time (sec.)	2141	492.57	2217	571.95	2234	763.18	2223	571.957
Performance Near transfer (max. = 4)	2.57	1.16	2.89	1.04	2.03	1.50	2.61	1.20
Performance Far transfer (max. = 6)	3.37	1.63	3.72	1.52	2.97	1.67	3.45	1.88
Mental effort Near transfer (max. = 7)	3.66	1.71	3.53	1.95	4.25	1.67	3.30	1.68
Mental effort Far transfer (max. = 7)	4.28	1.64	4.12	1.95	4.74	1.69	4.11	1.59
Efficiency Near transfer	.04	1.17	.27	1.21	-.51	1.45	.21	1.20
Efficiency Far transfer	.02	1.22	.23	1.34	-.35	1.32	.12	1.28
Training Phase								
Performance (max. = 24)	16.16	5.32	16.95	4.81	16.52	6.17	17.41	5.50
Mental effort (max. = 7)	3.19	1.65	2.77	1.45	3.40	1.73	2.56	1.28

Note: Estimated marginal means are presented with total training time as covariate (except for the factual knowledge test and time on training).

An ANOVA on the factual knowledge test filled out by the participants prior to the training revealed no differences between conditions, $F(1, 114) < 1$, *ns*. A significant main effect of control was found on time on the learning tasks, $F(1, 114) = 4.05$, $MSE = 1095979.79$, $p < .05$, $\eta^2_p = .03$. Participants in the program control conditions needed more time to perform the learning tasks ($M = 1819.07$, $SD = 521.04$) than participants in the learner control conditions ($M = 1624.78$, $SD = 517.54$). No effects on time on the learning tasks were found for KCR and the interaction between control and KCR. Therefore, the test scores - time, performance, and mental effort – and the training scores – performance and mental effort - were analyzed with ANCOVAs with the between-subjects factors control (program control, learner control) and KCR (present, absent) and the covariate time on learning tasks. For all statistical tests a significance level of .05 was maintained.

Test Phase

Test time. ANCOVA revealed no main effects on test time for control, $F(1, 113) = .15$, $MSE = 60003.81$, *ns*; KCR, $F(1, 113) = .27$, $MSE = 106571.28$, *ns*; and their interaction, $F(1, 113) = .02$, $MSE = 8436.49$, *ns*.

Test performance. ANCOVA showed a significant main effect of KCR on near transfer performance, $F(1, 113) = 4.15$, $MSE = 5.85$, $p < .05$, $\eta^2_p = .04$. Participants in the KCR conditions scored higher on the near transfer test ($M = 2.72$, $SD = 1.12$) than participants in the no-KCR conditions ($M = 2.33$, $SD = 1.34$). No effects on the near transfer test were found for control, $F(1, 113) = 3.03$, $MSE = 4.70$, *ns*; and the interaction between control and KCR, $F(1, 113) = .36$, $MSE = .51$, *ns*. In addition, no effects on far transfer performance were found for control, $F(1, 113) = 1.30$, $MSE = .3.21$, *ns;* KCR, $F(1, 113) = 2.05$, $MSE = 5.07$, *ns*; and their interaction, $F(1, 113) = .06$, $MSE = .14$, *ns*.

Mental effort. ANCOVA revealed no effects on mental effort during the near transfer test for control, $F(1, 113) = .36$, $MSE = .96$, *ns;* KCR, $F(1, 113) = 3.2$, $MSE = 8.56$, *ns*; and their interaction, $F(1, 113) = 1.88$, $MSE = 5.03$, *ns*. Similarly, no effects on mental effort during the far transfer test were found for control, $F(1, 113) = .36$, $MSE = .96$, *ns;* KCR, $F(1, 113) = 3.2$, $MSE = 8.56$, *ns*; and their interaction, $F(1, 113) = 1.88$, $MSE = 5.03$, *ns*.

Efficiency. ANCOVA showed a significant main effect of KCR on efficiency on near transfer, $F(1, 113) = 4.70$, $MSE = 6.51$, $p < .05$, $\eta^2_p = .04$. Efficiency on the near transfer test was higher in the KCR conditions ($M = .20$, $SD = 1.20$) than in the no-KCR conditions ($M = -.20$, $SD = 1.31$). No effects on efficiency on near transfer were found for control, $F(1, 113) = 1.93$, $MSE = 3.67$, *ns*; and the interaction between KCR and control, $F(1, 113) = 1.24$, $MSE = 1.71$, *ns*. Similarly, no effects on efficiency on far transfer were found for control, $F(1, 113) = 1.12$, $MSE = 1.57$, *ns*; KCR, $F(1, 113) = 2.42$, $MSE = 3.40$, *ns*; and their interaction, $F(1, 113) = .35$, $MSE = .49$, *ns*.

Training Phase

Training performance. ANCOVA revealed no effects on training performance for control, $F(1, 113) = .18$, $MSE = 4.67$, *ns*; KCR, $F(1, 113) = .78$, $MSE = 20.43$, *ns;* and their interaction, $F(1, 113) = .003$, $MSE = .07$, *ns*.

Mental effort. ANCOVA revealed a significant main effect of KCR on mental effort during training, $F(1, 113) = 5.69$, $MSE = 11.501$, $p < .025$, $\eta^2_p = .05$. Participants in the KCR conditions reported lower mental effort during training ($M = 2.71$, $SD = 1.37$) than participants in the no-KCR conditions ($M = 3.24$, $SD = 1.68$). No effects on mental effort during training were found for control, $F(1, 113) = .00$, $MSE = .00$, *ns*; and the interaction between control and KCR, $F(1, 113) = .164$, $MSE = 1.28$, *ns*.

Motivation

Table 6.2 provides an overview of the mean scores and standard deviations for the perceived relevance item and the four IMMS scales.

Table 6.2
Overview of Results from the Perceived Relevance Item and the Instructional Materials Motivation Survey (IMMS)

	Program control				Learner control			
	No-KCR (n = 30)		KCR (n = 29)		No-KCR (n = 30)		KCR (n = 29)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
During training								
Perceived Relevance	-	-	-	-	2.61	.77	3.16	.66
After training - IMMS								
Attention	2.89	.66	3.16	.83	2.78	.71	3.42	.59
Relevance	3.30	.63	3.38	.77	3.28	.59	3.70	.56
Confidence	3.28	1.03	3.52	1.09	3.39	.96	3.93	.80
Satisfaction	2.72	.89	3.10	1.03	2.82	.78	3.45	.82

Perceived relevance. A t-test showed a significant difference between participants' perceived relevance of the three choices provided in the two learner control conditions, $t(57) = -2.95$, $p < .01$, $d = 0.77$, indicating a medium to large effect size. Perceived relevance was significantly higher in the learner control/KCR condition ($n = 29$; $M = 3.16$, $SD = .66$) than in the learner control/no-KCR condition ($n = 30$; $M = 2.61$, $SD = .77$).

IMMS. ANOVAs showed significant main effects for KCR on attention, $F(1, 114) = 12.31$, $MSE = 6.08$, $p < .001$, $\eta^2_p = .10$; relevance, $F(1, 114) = 4.51$, $MSE = 1.86$, $p < .05$, $\eta^2_p = .04$; confidence, $F(1, 114) = 4.72$, $MSE = 4.51$, $p < .05$, $\eta^2_p = .04$; and satisfaction, $F(1, 114) = 9.65$, $MSE = 7.57$, $p < .01$, $\eta^2_p = .08$. Participants in the KCR conditions reported higher attention ($M = 3.29$, $SD = .73$), higher relevance ($M = 3.54$ $SD = .60$), higher confidence ($M = 3.73$, $SD = .97$), and higher satisfaction ($M = 3.27$, $SD = .94$) than participants in the no-KCR conditions (in order, $M = 2.83$, $SD = .68$; $M = 3.28$, $SD = .60$; $M = 3.33$, $SD = .99$, and $M = 2.77$, $SD = .83$ for attention, relevance, confidence, and satisfaction). Post hoc tests using Tukey's HSD revealed that participants in the learner control/KCR condition reported significantly higher attention ($M = 3.42$, $SD = .59$) than participants in both the program control/no-KCR condition ($M = 2.89$, $SD = .66$; $p < .025$) and the learner control/no-KCR condition ($M = 2.78$, $SD = .71$; $p < .01$). In addition, participants in the learner control/KCR condition reported significantly higher satisfaction (*M* = 3.45, *SD* = .82) than participants in both the program control/no-KCR condition (*M* = 2.72, *SD* = .89; p $<$ .05) and the learner control/no-KCR condition (*M* = 2.82, *SD* = .78; *p* $<$ .05).

Perceived Control

Because the assumption of homogeneity of variances was violated, a Kruskal-Wallis test was used to compare the four conditions on perceived control. A significant effect of condition was found, $H(3) = 45.57$, $p < .001$. Multiple comparisons among groups were performed using a Conover-Inman test, which is a non-parametric alternative to Fisher's least significance difference method performed on ranks (Conover, 1999). Results showed a significantly lower perceived control for the program control/no-KCR condition (mean rank = 36.60) than the learner control/no-KCR condition (mean rank = 69.23; $p < .001$) and than the learner control/KCR condition (mean rank = 89.66; $p < .001$); a significantly lower perceived control for the program control/KCR condition (mean rank = 42.97) than the learner control/no-KCR condition (mean rank = 69.23; $p < .001$) and than the learner control/KCR condition (mean rank = 89.66; $p < .001$); and a significantly lower perceived control for the learner control/no-KCR condition (mean rank = 69.23) than the learner control/KCR condition (mean rank = 89.66; $p < .01$).

Discussion

This study investigated the effects of giving feedback in the form of KCR, which emphasized structural features of learning tasks (i.e., to-be-completed solution steps), on transfer, efficiency, and motivation. Moreover, it examined if potential beneficial effects of providing KCR are higher in combination with learner controlled selection of learning tasks, which gives learners the opportunity to select personally relevant learning tasks, than with program controlled selection of learning tasks.

First, we hypothesized that providing learners with KCR emphasizing structural features leads to higher transfer, efficiency, and motivation. This hypothesis is largely supported by our results. Learners provided with KCR performed better on the near transfer test and also showed higher efficiency on near transfer, indicating that the near transfer performance of participants provided with KCR was higher than could be expected on the basis of their invested mental effort. In addition, participants provided with KCR scored significantly higher on all four scales of the IMMS questionnaire (attention, relevance, confidence, and satisfaction) than participants who were not provided with KCR. This seems to indicate that KCR helps learners to recognize structural features, which enables them to connect what is presented to them (i.e., learning tasks) to what they already know. In addition, results support Keller's (1983a, 1983b) theory of motivation, which argues that the motivation of a learner can be manipulated by the instructional design of the materials. They are also in line with Bassok (1990), who stated that it is possible to foster transfer of learning by increasing the relative weight of structural aspects, for example, by giving learners information about the relevance of structural features.

In contrast to the results on performance and efficiency for near transfer, no differences between the experimental conditions were found on performance and efficiency for *far* transfer. This may have been caused by the fact that KCR located where the error exactly was and what the learner could do to solve that problem (i.e., consult the basic information). This yielded 'restricted' cognitive schemas that allowed learners to perform the steps of the near transfer test - as indicated by the higher performance and efficiency for near transfer - as 'routines' (van Merriënboer, 1997). However, far transfer does not require learners to merely apply a routine: deep understanding of the rationale behind the solution steps is crucial. Providing learners not only with the correct solution steps, but also with the rationale behind and/or the purpose of the solution steps could have enabled them to more flexibly use those steps which is essential for far transfer (van Gog, Paas, & van Merriënboer, 2004, 2006, in press).

Second, we studied if the beneficial effects of providing KCR would be higher in combination with learner control than in combination with program control, because the given feedback has the potential to make the structural task features more salient for learners, enabling them to select personally relevant tasks which further enhances learning and motivation. Our results do not support the superiority of combining KCR with learner control for transfer and efficiency, but the effects on motivation are in the expected direction. Participants in the learner control/KCR condition reported higher attention and satisfaction than participants in the learner control/no-KCR condition. Moreover, participants in the learner control/KCR condition reported to perceive the choices provided as more relevant than participants in the learner control/no-KCR condition. Additionally, participants in the learner control/KCR condition perceived more control than participants in the learner control/no-KCR condition, which indicates that the provision of KCR enhances the perception of control.

Nevertheless, with regard to transfer and efficiency the learner control conditions did not profit more from KCR than the program control conditions. Apparently, KCR did not sufficiently support learners in making more effective task selections which seems to support the idea that less experienced learners are not able to make effective selections regarding structural features and must thus be explicitly guided in how to achieve learning objectives (Butler & Winne, 1995). Alternatively, the small number of tasks to choose from (only three) and the small variety between the three tasks to choose from may also have limited the learners' opportunities to select a range of personally relevant tasks with a genuine effect on learning, although it positively influenced motivation.

A final unexpected result that needs to be discussed is that participants in the learner control conditions invested less time in performing the learning tasks than participants in the program control conditions. Descriptive analysis reveals that in the learner control/no-KCR condition, the time invested was lower (*M* = 1541, *SD* = 553.34) than in the learner control/KCR condition (*M* = 1711, *SD* = 471.71). This lower time invested in the learner control/no-KCR condition seems to account for the largest part of the lower time invested in both learner control conditions together compared with the program control conditions. Participants in the learner control/no-KCR condition were not supported in recognizing the structural features, perceived the choices provided as being significantly less relevant, and scored relatively low on all motivational measures. Their lack of motivation may well explain the lower time invested in training.

Our findings yield some important implications for future research. First, more sophisticated process-tracking methods, such as eye-tracking or thinking-aloud protocols, may uncover whether learners who are provided with feedback indeed focus their attention to more relevant task aspects than learners who are not provided with feedback. Second, the issue of near and far transfer should be addressed in forthcoming studies. Possibly, richer types of feedback than KCR yield not only effects on performance and efficiency for near transfer, but also for far transfer. Similarly, the effects on transfer performance should be examined over a more extended period of time because transfer may not be apparent immediately after practice, but may only be present at a later time if the same or additional transfer tasks are repeated (Gick & Holyoak, 1987). Third, Bell and Kozlowski (2002) found that adaptive guidance which provided learners with diagnostic and interpretive information (i.e., a sort of guidance), supported learners in making more effective learning decisions. Future studies may examine the effects of adaptive feedback on learners' selection of learning tasks on the basis of their structural features.

To conclude, this study clearly shows the beneficial effects of feedback in the form of KCR on near transfer performance and efficiency, as well as its added value for motivating learners when they are given the freedom to select tasks that differ with regard to their structural features. These results are particularly important because more and more educational approaches stress the importance of providing learners with control over the learning tasks they perform. This could easily hamper learners' motivation if they are brought into a position in which they cannot see valid motives for choosing between different tasks.

7

General Discussion

In this chapter, a general overview of the dissertation is presented, the main conclusions of the studies presented in Chapters 3 to 6 are described, and limitations of the studies are acknowledged. Furthermore, theoretical implications for the improvement of the tested task selection models in future studies as well as practical implications of the studies are discussed.

Overview of the Results

Current educational approaches focus on flexible curricula which offer a unique sequence of learning tasks selected according to a learner's needs or preferences. Task sequence can be personalized by either a computer program or the learners themselves. The studies in this dissertation investigated under which conditions learner control over task selection is most effective.

First of all, the studies all used a specific form of learner control, namely, shared control. With shared control over task selection, there is partial program control combined with partial learner control, that is, the program first selects a subset of tasks from which the learner can subsequently choose one task to work on. The model presented in Chapter 2 underlines the trade-off between program control and learner control over task selection as a basis for shared control. The effectiveness of learner control largely depends on which task features are selected by the program and which task features are selected by the learners themselves. With a good trade-off, shared control should reduce the choice provided to learners, prevent cognitive overload, and ensure that learners select from a subset of tasks that is optimal for their learning.

The studies described in Chapters 3 to 6 tested the effects of different shared control models for dynamic task selection on effectiveness (i.e., transfer test performance), efficiency (i.e., test performance combined with invested mental effort), and motivational effects of the learning material. In Chapter 3, difficulty and support were tailored to changing levels of *learner expertise*. In the studies described in Chapters 4 to 6, learning tasks were (pre-) selected on the basis of previous surface or structural *task features*.

In general, four conclusions can be drawn from the studies reported in this dissertation. First, adapting critical task aspects such as task difficulty and embedded support to changing levels of expertise, prevents cognitive overload and associated negative effects on learning (Chapter 3). Second, choosing tasks with preferred surface features has beneficial effects on learning and motivation (Chapters 3 to 5). However, the range of tasks to choose from should not include tasks with surface features very similar to those of the previous task (Chapter 4). Third, learners should receive extra support if they are given control over the selection of tasks that differ in their non-salient structural features (Chapter 5). Support in the form of feedback emphasizing structural features, such as 'Knowledge of Correct Response' (KCR), makes learner control over task selection on the basis of those features more motivating (Chapter 6). Fourth, in order to have positive effects on learning outcomes, more powerful feedback strategies than KCR are needed to support learners in making optimal selections from tasks with different structural features. The next sections summarize general findings and provide explanations for expected and unexpected results.

Adaptive task selection

In all studies reported in this dissertation, each task or subset of tasks was dynamically adapted to either learner variables (Chapter 3) or task variables (Chapters 4 to 6). First, tailoring task difficulty and given support to the *learner variables* competence level and reported task load led to more effective and efficient learning (Chapter 3). Probably, adapting task selection to learner variables lowered cognitive load to an acceptable level, which then enabled the allocation of freed-up cognitive resources to learning. Unexpectedly, training time was found to be lower in the non-adaptive conditions. Training time can be considered as 'cost' associated with learning and in principle a lower training time may indicate more effective training. A possible explanation has been provided in the Discussion of Chapter 3. Participants in the adaptive conditions could have noticed the relationship between their performance and the difficulty and/or embedded support of each subsequent task, whereas participants in the non-adaptive conditions probably lacked this association. This might have negatively influenced their time investment.

Second, adaptation to *task features* was based on the assumption that effectiveness of learner control partly depends on preventing learners from being overloaded by a too high amount of choice. In addition, the control should be actually perceived as something valuable by the learners. Hence, the program should pre-select tasks with different salient *surface* features to enable learners to select a varied set of personally relevant tasks (Chapter 5). Accordingly, if the program pre-selects tasks with highly similar surface features, the effectiveness of learner control will be hampered because learners will recognize the lack of differences in the choices provided (Chapter 4). Additionally, if the program pre-selects tasks with non-salient *structural* task features, learners will probably not recognize those features and will be unable to select a varied set of personally relevant tasks. This annuls the effectiveness of learner control (Chapter 5). These assumptions will be examined more closely in the next section.

Learner control over task selection

It has been noted that the added value of learner control largely depends on *what* (e.g., surface features or structural features) is selected by learners. Learners will profit from the control exercised by them only when they use it to select a varied set of personally relevant tasks (Katz & Assor, 2007). A first remarkable conclusion concerns the beneficial effects of learner control over the selection of tasks with different *surface* features. Because those features are salient and thus perceptible for learners, they can use this information to select a varied set of personally relevant tasks which promotes induction and elaboration and, eventually, transfer of learning. In Chapter 3 evidence was found that allowing learners to select their preferred surface features enhances task involvement. Furthermore, providing learners with a subset of tasks with surface features dissimilar from the previous task was found to enhance transfer test performance and task involvement (Chapter 4) as well as the efficiency on far transfer (Chapter 5). In Chapter 5, it was argued that selecting one's preferred surface features may have elicited personal relevance which facilitated learners to connect new information to their prior knowledge. This process of elaboration enables the construction of general cognitive schemas which are especially useful when solving unfamiliar tasks (i.e., far transfer tasks).

Learner control over the selection of tasks with different *structural* features is difficult for learners. They do not spontaneously recognize these non-salient features and thus cannot select a varied set of personally relevant tasks based on them (Chapter 5). Yet, learners should *not* just like that be given control over tasks with similar structural features because, probably, they will not see the point of choosing between outwardly similar tasks. They may even become frustrated from choosing between things that are seen as equivalent. Hence, learners need to be supported in identifying the structural task features, for instance, by giving them feedback. The hypothesis that providing learners with feedback over structural features would lead to more efficient learning was indeed supported for a near transfer test although the expected pattern found for the far transfer test did not reach significance (Chapter 6). It is possible that the feedback helped learners construct restricted cognitive schemas, allowing them to perform routine aspects of the task. But learners did possibly not fully *understand* how to solve the problems, which is a prerequisite for far transfer. Additionally, feedback increased motivation but it did not support learners on the task selection process (Chapter 6).

The effects of having learners select from a subset of tasks differing in surface features and from a subset of tasks differing in structural features, might well be captured by the distinction between 'picking' (i.e., selecting without preferences) and 'choosing' (i.e., selecting as meaningful realization of preferences; Ullmann-Margalit & Morgenbesser, 1997). Picking is less motivating than choosing and also undermines the effectiveness of learner control (see for a review, Katz & Assor, 2007). The difference between picking and choosing might well explain the superior effects on transfer test performance and task involvement of learner control when learners choose from pre-selected tasks with surface features that were different from the surface features of previous tasks, found in the study reported in Chapter 4.

It was assumed that surface features are easier to recognize than structural features, so that learners who select tasks on the basis of their surface features would perceive a higher level of control (Chapter 5). Although the answers to the perceived control questions did not directly support this assumption, as expected, only participants who selected their preferred tasks on the basis of surface features benefited from the control exerted; and moreover, they selected much more varied tasks. Future studies should explicitly address the issue if learners indeed recognize surface features and structural features, and if so, if they actually use this information to select the next task(s) to work on. Additionally, the perceived control and the perceived relevance of the given choices were found to be higher when feedback emphasizing structural features was provided. Thus, it can be concluded that feedback is helpful to recognize structural features and should thus be provided when learners have to choose between tasks which differ on those features (Chapter 6).

Our studies yielded both positive and unexpected results on motivation. Learner control enhanced task involvement (Chapters 3 and 4), and the provision of feedback increased the perceived relevance and motivational effects of making choices between tasks with different structural features (Chapter 6). However, no effects were found on interest and other scales of the Intrinsic Motivation Inventory (Chapters 3 and 4). Possibly, a wider range of tasks to choose from or a higher variety between the tasks to choose from in the learner control conditions could have revealed differences in the motivational scales between the experimental conditions. The fact that positive effects were found on the task involvement measure but not on interest and other scales of the Intrinsic Motivation Inventory seem to support Paas et al.'s (2005) argument that combining cognitive load and performance measures is supplementary to the use of inventories collecting motivational data. Another explanation concerns the fact that learner control was only limited to *task selection.* That is, learner control while working on the tasks was similar in all conditions. Hence, the provision of learner control over task selection may be considered as only one method to enhance interest, which needs to be combined with other methods to reach effects on motivation.

Summary of methods to optimize learner control

Whereas learner control enhances learning and motivation, it should not be used unconditional. Learners should *perceive* the control provided, choices should enable learners to select personally *relevant* tasks, learners should select relatively *varied* tasks, and they should *not* be *overwhelmed* by the amount of choice. Shared control limits the choice to manageable levels. This section briefly summarizes proposed strategies to enhance learner control.

First, *learner control should be provided over tasks with varied surface features.* Surface features are often more perceptible for learners than structural features. Learners who are presented a subset of tasks that differ in surface features from previous tasks will most likely recognize those differences, and be able to select a varied set of personally relevant tasks which enhances learning and transfer because of improved induction and elaboration. However, *learner control should not be provided over tasks with highly similar surface features* because this may hamper task involvement and transfer.

Second, learner control over tasks with different non-salient structural features does *not* enhance learning because learners are not likely to recognize those features and are not able to select a varied set of personally relevant tasks (Chapter 5). It has been suggested that *feedback should be provided in combination with learner control over tasks with different structural features* because this underlines those features and enables learners to select personally relevant tasks. In the study reported in Chapter 6, feedback enhanced motivation and made the choices provided (i.e., tasks differing in structural features) more personally relevant for learners. But feedback did not support learners in making better task selections. Giving feedback in the form of correct responses is thus insufficient to develop task selection skills. More powerful feedback strategies in the form of *advice* on task selection might better help learners to engage in appropriate actions – self-selecting *optimal* learning tasks – to close the gap between actual and desired performance (Butler & Winne, 1995).

Limitations

Three methodological issues need to be acknowledged regarding the use of the efficiency and task involvement measures reported in Chapter 3. First, the original efficiency measure proposed by Paas and van Merriënboer (1993) relies on the combination of performance and invested mental effort during a *test* and examines the relative instructional efficiency of experimental conditions in terms of learning outcomes. More efficient learning outcomes are indicated by relatively high performance combined with low mental effort during the test, and less efficient learning outcomes are indicated by relatively low performance combined with high mental effort during the test. A review of the use of the instructional efficiency measure (van Gog & Paas, 2008) underlines the widely inappropriate use (33 out of 36 reviewed studies) of the originally proposed measure by cognitive load researchers. These studies, as well as the study described in Chapter 3, adopted an 'adapted' measure to compute efficiency that combines *test* performance with mental effort invested during *training* (referred to as 'task load' in Chapter 3), which analyzes instructional efficiency in terms of the learning process rather than the learning outcomes. Although the use of the adapted measure is not problematic in studies that only aim at reducing extraneous load, its use in studies aiming at increasing germane load is inappropriate because it is then an explicit goal to increase the investment of mental effort during training.

A second, related issue with regard to the efficiency measure concerns its combined use for selecting tasks and computing efficiency. As in the study reported in Chapter 3, in the review of van Gog and Paas (2008) 6 from the 33 studies applying the adapted measure used mental effort during training not only to compute efficiency but *also* to select learning tasks. Although the combination of performance and mental effort during training provides valuable information for dynamic task selection, van Gog and Paas make clear that using the same mental effort invested during training to compute efficiency and to select tasks, poses an additional limitation because mental effort has been directly manipulated by the instructional strategy. That is, in Chapter 3 cognitive load was optimized in the adaptive conditions but not in the non-adaptive conditions, and this inevitably affects the adopted efficiency measure. Hence, a replication of the study should compute efficiency using mental effort invested during the test phase, which provides a better indicator of the quality of the learning outcomes.

The third issue concerns the use of the *task involvement* measure in Chapter 3 which is based on the assumption that higher involvement is reflected by the investment of more mental effort and higher resulting performance. The original computation of task involvement (Paas et al., 2005) uses mental effort as a general term, not making a distinction between the investment of effort to deal with extraneous/intrinsic load and the investment of effort to deal with germane load. However, the study in Chapter 3 used a scale that solely intended to measure the effort learners invested *in learning* (i.e., germane load) to compute task involvement. Future efforts must address whether perceived mental effort in general, which may also be influenced by aspects such as the difficulty of the task (affecting intrinsic load) or the availability of support (affecting extraneous load), or a measure of mental effort directly related to learning (i.e., germane load), should best be used to compute task involvement.

A limitation of the studies reported in Chapters 4 and 5 concerns the lack of direct measures of some underlying variables. First, the higher perception of control when learners could select amongst tasks with different surface features accounted for the positive effects reported in Chapter 4. Second, the unexpected similar perceived control reported by the participants in all learner control conditions (i.e., control over tasks with different surface features and with different structural features) in the study described in Chapter 5, suggested that not only perception of control but also perceived relevance of the given choices is a prerequisite for learner control to work. However, none of these variables were directly measured. Although the variables that were hypothesized to be affected by increasing perceived control and perceived relevance were directly measured, replication of the studies must include measures of perception of control (Chapter 4) as well as measures of personal relevance (Chapters 4 and 5).

Finally, as a more general measurement issue, it is not yet clear if learners are actually able to distinguish between different types of cognitive load. In fact, cognitive load theorists are still facing the challenge to distinguish the different types of cognitive load through self-reporting instruments. The multi-dimensionality of the concept of cognitive load may even rise the question whether it is possible at all to empirically separate the constituent types of load.

Implications and Future Research

The results from our studies provide a number of issues that should be considered in future research and have some practical implications as well. First, adaptive task selection on the basis of *learner variables* should include variables representing other costs than mental effort, such as time on task. Also, to get better insights in the quality of learners' constructed cognitive schemas, more direct measures - such as verbal protocols, retrospective reports, and eye tracking data - of their mental processes could be taken during training. The inclusion of these variables would further refine the basis for adaptive task selection, which in turn may provide superior learning results. Although in the study reported in Chapter 3 some additional rules to select challenging tasks were included, a more sophisticated selection algorithm should include a combination of cognitive and motivational measures to dynamically select challenging tasks for learners (e.g., a more difficult task or a task with less support). In addition, future efforts should implement task selection rules to tailor the practice of tasks with different *structural features* to the learner's progress, for example, by providing additional practice or feedback when a learner repeatedly makes an error in a specific solution step.

Second, it has been argued that the amount of learner control should increase as learner expertise develops (Chapter 2). However, the reported studies kept the level of learner control constant. They did not take this aspect of the model into account. A fine-tuned model of shared control could first offer learners control over surface features, then over surface features plus embedded support, and, finally, over surface features plus support plus task difficulty. As learner control increases, learners may receive *advice* – rather than merely knowledge of correct responses - to guide them through the complex process of task selection (Kicken, Brand-Gruwel, & van Merriënboer, 2007). Another possibility would be to explicitly teach learners to self-assess themselves and to select optimal tasks (Kostons, van Gog, & Paas, 2007). This seems a promising approach which will allow learners to select tasks on the basis of their *self-assessed* competence in combination with their reported mental effort. Nevertheless, instructional designers should keep in mind that too much choice causes cognitive overload, even for expert learners.

Third, the transfer test scores reported in the study described in Chapter 5 were relatively low, which may have attenuated learning and made it more difficult to find differential effects. In that study, all tasks involved seven types of solution steps, from which three steps were repeated twice. This could have resulted in incoherence caused by redundant information (i.e., the same step repeated twice) or to too detailed information (i.e., an overly specific description of the steps) which may have hampered learning and transfer (Nadolski, Kirschner, & van Merriënboer, 2005). Future studies should reveal whether optimizing the number and the level of detail of steps required to solve a problem better enables learners to understand the problem, which will ultimately enhance the learning outcomes and especially transfer.

Fourth, whereas in Chapters 3 and 4 the transfer test was administered one week after the training, in Chapters 5 and 6 the transfer test was administered immediately after the training. Hence, nothing can be concluded about short-term effects in Chapters 3 and 4 and about long-term effects in Chapters 5 and 6. Future research is needed to determine whether the results can also be found with immediate (Chapters 3 and 4) and delayed (Chapters 5 and 6) assessments of transfer test performance. In addition, although all the studies of this dissertation adopted a quantitative approach, a more qualitative approach may complement and corroborate the results. For example, interviews may provide more insight in learners' criteria for their task selections (e.g., whether they 'pick' or 'choose' a task) and their motives to invest effort and time in training.

Our findings have important educational implications. Effects of perceived control have been widely studied in medicine (e.g., on cancer, stress, asthma) and even in gambling, but less in educational contexts. Giving learners control over surface features in such a way that they perceive this control is yet beneficial for learning. Although surface features are not considered to affect the way a task is solved, instructional designers should acknowledge their instructional importance because of their strong influence on learning and transfer. A final implication pertains to the implementation of dynamic task selection in the field of lifelong learning (van Gerven, 2002). Given the ability to adapt the sequence of tasks on the basis of prior performance and mental effort invested, this approach may be especially appropriate and valuable for the elderly, because this group generally shows larger individual differences in those variables than young people do.

References

Alexander, P. A., & Jetton, T. L. (1996). The role of importance and interest in the processing of text. *Educational Psychology Review, 8*, 89–121.

Averill, J. R. (1973). Personal control over aversive stimuli and its relationship to stress. *Psychological Bulletin, 80*, 286-303.

Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research, 13,* 111-127.

Baartman, L. K. J., Bastiaens, T. J., Kirschner, P.A., & van der Vleuten, C. P.M. (2007). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. *Educational Research Review, 2,* 114-129.

Baddeley, A. D. (1992). Working memory. *Science, 255,* 556-559.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W. H. Freeman.

Bassok, M. (1990). Transfer of domain-specific problem-solving procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 522-533.

Beal, C. R. (2004). Adaptive user displays for intelligent tutoring systems. *CyberPsychology and Behavior, 7,* 689-693.

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.

Borsook, T. K., & Higginbotham-Wheat, N. (1991). Interactivity: What is it and what can it do for computer-based instruction? *Educational Technology*, *31*, 11-17.

Brna, P., & Cox, R. (1998). Adding intelligence to a learning environment: Learner-centered design? *Journal of Computed Assisted Learning*, *14*, 268-277.

Burger, J. M. (1989). Negative reactions to increases in perceived personal control. *Journal of Personality and Social Psychology, 56*, 246-256.

Burger, J. M., & Cooper, H. M. (1979). The desirability of control. *Motivation and Emotion, 3*, 381-393.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65,* 245-281.

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.

Chai, L. (2003). To have or have not: An examination of feedback, learner control and knowledge type in online learning. In R. H. Sprague (Ed.), *Proceedings of the 36th Hawaii international Conference on System Sciences*. Hawaii: IEEE Computer Society: CD-Rom, 10 pp.

Chen, Z., & Mo, L. (2004). Schema induction in problem solving: A multidimensional analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 583-600.

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

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

Clark, R. C., Nguyen, F., & Sweller, J. (2005). *Efficiency in learning*. San Francisco, CA: Pfeiffer.

Conover, W. J. (1999). *Practical nonparametric statistics* (3rd Ed.). New York: John Wiley & Sons Inc.

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.

Corbalan, G., Kester, L., & van Merriënboer, J. J. G. (in press-a). Selecting learning tasks: Effects of adaptation and shared control on learning efficiency and task involvement. *Contemporary Educational Psychology.*

Corbalan, G., Kester, L., & van Merriënboer, J. J. G. (in press-b). Combining shared control with variability over surface features: Effects on transfer test performance and task involvement. *Computers in Human Behavior*.

Corbett, A. T., Koedinger, K. R., & Hadley, W. H. (2001). Cognitive tutors: From the research classroom to all classrooms. In P. S. Goodman (Ed.), *Technology enhanced learning: Opportunities for change (*pp. 235-263). Mahwah, NJ: Lawrence Erlbaum Associates.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology, 88,* 715-730.

Corno, L., & Snow, R. E. (1986). Adapting teaching to individual differences among learners. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd Ed.) (pp. 605-629). New York: MacMillan.

Crocker, L, & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston, Inc.

Cummins, D. D. (1992) The role of analogical reasoning in the induction of problem categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1103-1124.

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality, 62,* 119-142.

Deci, E. L., & Ryan. R. M. (1985). *Intrinsic motivation and self-determination in human behaviour.* New York: Plenum.

Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist, 26*, 325-346.

De Croock, M. B. M., & van Merriënboer, J. J. G. (2003, April). *TASKI: A system to adapt 4C/ID training programs to a learners' learning capability.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Eraut, M. (1994). *Developing professional knowledge and competence*. London: Routledge Falmer.

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

Fisher, S. L., & Ford, J. K. (1998). Differential effects of learner effort and goal orientation on two learning outcomes. *Personnel Psychology, 51,* 397-420.

Flowerday, T., & Schraw, G. (2000). Teachers' beliefs about instructional choice: A phenomenological study. *Journal of Educational Psychology, 92,* 634–645.

Freitag, E. T., & Sullivan, H. J. (1995). Matching learner preference to amount of instruction: An alternative form of learner control. *Educational Technology, Research, and Development, 43*, 1042-1629.

Fry, J. P. (1972). Interactive relationship between inquisitiveness and learner control of instruction. *Journal of Educational Psychology*, *63*, 459-465.

Gagné, E. D., Crutcher, R. J., Anzelc, J., Geisman, C., Hoffman, V. D., Schultz, P., & Lizcano, L. (1987). The role of student processing of feedback in classroom achievement. *Cognition and Instruction, 4,* 167-186.

Garcia, T., & Pintrich, P. R. (1994). Regulating motivation and cognition in the classroom: The role of self-schemas and self-regulatory strategies. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications* (pp. 127-153). Hillsdale, NJ: Erlbaum.

Gay, G. (1986). Interaction of learner control and prior understanding in computer-assisted video instruction. *Journal of Educational Psychology, 78*, 225-227.

Gick, M., & Holyoak, J. (1987). The cognitive basis for knowledge transfer. In S. Cormier & J. Hagman (Eds.), *Transfer of learning: Contemporary research and applications* (pp. 9-46). San Diego, CA: Academic Press.

Gray, S. J. (1987). The effects of sequence control on computer learning. *Journal of Computer-based Instruction, 14*(2)*,* 54-56.

Green, B. A., & Miller, R. B. (1996). Influences on achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology, 21*, 181-192.

Hannafin, M. J. (1984). Guidelines for using locus of instructional control in the design of computer-assisted instruction. *Journal of Instructional Development, 7*(3), 6-10.

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology, 21,* 713-729.

Hidi, S. (2001). Interest, reading and learning: Theoretical and practical considerations. *Educational Psychology Review*, *13,* 191–210.

Hidi, S. (2006). Interest: A unique motivational variable. *Educational Research Review, 1,* 69-82.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41,* 111-127.

Holladay, C. L., & Quiñones, M. A. (2003). Practice variability and transfer of training: The role of self-efficacy generality. *Journal of Applied Psychology, 88,* 1094-1103.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition, 15,* 332-340.

Hyland, F. (2001). Providing effective support: Investigating feedback to distance language learners. *Open Learning, 16,* 233-247.

Iyengar, S. S., & Lepper, M. (2000). When choice is demotivating: Can one desire too much of a good thing?. *Journal of Personality and Social Psychology, 79*, 995-1006.

Jochems, W. , van Merriënboer, J. J. G., & Koper, R. (2005). An introduction to integrated e-learning. In W. Jochems, J. J. G. van Merriënboer, & R. Koper, (Eds.). *Integrated e-learning: Implications for pedagogy, technology, and organization, technology and organization* (pp. 1-12). London, UK: Routledge Falmer.

Judd, W. A. (1972). Learner-controlled computer-assisted instruction. ERIC #ED 072-635.

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

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.

Katz, I., & Assor, A. (2007). When choice motivates and when it does not. *Educational Research Review, 19,* 429-442.

Kehoe, J. F. (1979). Choice time and aspects of choice alternatives. In L. C. Perlmunter & R. A. Monty (Eds.), *Choice and perceived control* (pp. 67-82). Hillsdale, NJ: Erlbaum.

Keller, J. M. (1983a). *Development and use of the ARCS model of motivational design.* Enschede, The Netherlands: Toegepaste Onderwijskunde, Technische Hogeschool Twente.

Keller, J. M. (1983b). Motivational design of instruction*.* In C. M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current status* (pp. 383-436). Hillsdale, NJ: Erlbaum.

Keller, J. M. (1987). Development and use of the ARCS model of motivational design. *Journal of Instructional Development, 10*(3), 2-10.

Kester, L., Kirschner, P. A., & van Merriënboer, J. J. G. (2006). Just-in-time information presentation: Improving learning a troubleshooting skill. *Contemporary Educational Psychology*, *31,* 167-185.

Kicken, W., Brand-Gruwel S, & van Merriënboer, J. J. G. (2007). *Scaffolding advice on task selection: A safe path toward self-directed learning in on-demand education*. Manuscript submitted for publication.

Kinzie, M. B. (1990). Requirements and benefits of effective interactive instruction: Learner control, self-regulation, and continuing motivation. *Educational Technology, Research and Development, 38*, 5-21.

Kinzie, M. B., & Sullivan, H. J. (1989). Continuing motivation, learner control, and CAI. *Educational Technology, Research and Development*, *37*, 5-14.

Kinzie, M. B., Sullivan, H. J., & Berdel, R. L. (1988). Learner control and achievement in science computer-assisted instruction. *Journal of Educational Psychology, 80,* 299-303.

Kopcha, T. J., & Sullivan, H. (in press). Learner preferences and prior knowledge in learner-controlled computer-based instruction. *Educational Technology, Research and Development.*

Kostons, D., van Gog, T., & Paas, F. (2007). *Self-assessment and task selection as determinants of effective learner-controlled instruction: A cognitive load perspective.* Manuscript submitted for publication.

Lahey, G. F. (1976). Learner-control of lesson strategy: A model for PLATO IV system lessons. ERIC #ED 125-543.

Lahey, G. F., Hurlock, R. E., & McCann, P. H. (1973). Post-lesson remediation and learner control of branching in computer-based learning. ERIC #ED 083-797.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32*, 311-328.

Lawless, K. A., & Brown, S. W. (1997). Multimedia learning environments: Issues of learner control and navigation. *Instructional Science, 25,* 117-131.

Lee, T. D., & Magill, R. A. (1985). Can forgetting facilitate skill acquisition? In D. Goodman, R. B. Wilberg, & I. M. Franks (Eds), *Differing perspectives in motor learning, memory, and control* (pp. 3–22). Amsterdam, The Netherlands: Elsevier.

Lepper, M. R. (1985). Microcomputers in education. *American Psychologist, 40*, 1-18.

Lin, B., & Hsieh, C. (2001). Web-based teaching and learner control: A research review. *Computers and Education, 37*, 377-386.

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin and Review, 6,* 586-597.

Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology, 88*, 49-63.

Margueratt, D. (2007). *Improving learners motivation through enhanced instructional design*. Unpublished master's thesis. Athabasca University, Athabasca, Canada.

Martens, R. L., & Kirschner, P. A. (2004, October). *What predicts intrinsic motivation?* Paper presented at the Annual Meeting of the Association for Educational Communications and Technology, Chicago, IL.

Merrill, M. D. (1980). Learner control in computer based learning. *Computers and Education, 4,* 77-95.

Merrill, M. D. (1994). *Instructional design theory*. Englewood Cliffs, NJ: Educational Technology Publications.

Merrill, M. D. (2002). Instructional strategies and learning styles: Which takes precedence?. In R.A. Reiser, & J.V. Dempsey (Eds.), *Trends and issues in instructional technology*. Upper Saddle River, NJ: Merrill Prentice Hall.

Moreno, R., & Flowerday, T. (2006). Students' choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology, 31,* 186-207.

Morrison, G. R., Ross, S. M., & Baldwin, W. (1992). Learner control of context and instructional support in learning elementary school mathematics. *Educational Technology, Research and Development, 40*, 5-13.

Morrison, G. R., Ross, S. M., O'Dell, J. K., & Schultz, C. W. (1988). Adapting text presentation to media attributes: Getting more out of less in CBI. *Computers in Human Behavior, 4,* 65-75.

Mory, E. H. (2003). Feedback research revisited. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 745-783). Mahwah, NJ: Lawrence Erlbaum.

Nadolski, R. J., Kirschner, P. A., & van Merriënboer, J. J. G. (2005). Optimizing the number of steps in learning tasks for complex skills. *British Journal of Educational Psychology, 75,* 223-237.

Niemiec, 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.

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 14,* 510-520.

Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 398- 415.

Opfermann, M., Gerjets, P., & Scheiter, K. (2005, August). *Exploration of hypermedia environments: Learner control and adaptive user strategies*. Paper presented at the Bi-annual Conference of the European Association for Research on Learning and Instruction (EARLI), Nicosia, Cyprus.

Overskeid, G., & Svartdal, F. (1996). Effects of reward on subjective autonomy and interest when initial interest is low. *The Psychological Record, 46*, 319-331.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skills in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84,* 429-434.

Paas, F. (1993). Instructional control of cognitive load in the training of complex cognitive tasks. Doctoral dissertation, University of Twente, Enschede, The Netherlands.

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

Paas, F., Tuovinen, J., van Merriënboer, J. J. G., & Darabi, A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *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.

Paas, F., & van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6,* 51-71.

Paas, F., & van Gog, T. (2006). Optimising worked example instruction: Different ways to increase germane cognitive load. *Learning and Instruction, 16,* 87-91.

Park, O., & Lee, J. (2003). Adaptive instructional systems. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd Ed.) (pp. 651-684). Mahwah, NJ: Erlbaum.

Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology, 88,* 144-161.

Quilici, J. L., & Mayer, R. E. (2002). Teaching students to recognize structural similarities between statistics word problems. *Applied Cognitive Psychology, 16*, 325-342.

Reder, L., & Klatzky, R. L. (1994). Transfer: Training for performance. In D. Druckman & R. A. Bjork (Eds.), *Learning, remembering, believing: enhancing human performance* (pp. 25-56). Washington, DC: National Academy Press.

Reeve, J., Hamm. D., & Nix, G. (2003). Testing models of the experience of self-determination in intrinsic motivation and the conundrum of choice. *Journal of Educational Psychology, 95*, 375-392.

Reeves, T. C. (1993). Pseudoscience in computer based instruction: The case of learner control research. *Journal of Computer Based Instruction, 20*(2), 39-46.

Reigeluth, C. M. (1999) (Ed.). *Instructional design theories and models: A new paradigm of instructional theory* (Vol. II). Mahwah, NJ: Lawrence Erlbaum.

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., Stark, R., Gruber, G., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23,* 90-108.

Richman, W., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*, *84*, 754-775.

Rieber, L. P. (2005). Multimedia learning in games, simulations, and microworlds. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 549-567). New York: Cambridge University Press.

Ross, T. C. (1989). Distinguishing types of superficial similarities: Different effects on the access and the use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 456-468.

Ross, S. M., McCormick, D., & Krisak, N. (1986). Adapting the thematic context of mathematical problems to students interests: Individual versus group-based strategies. *Journal of Educational Research, 79,* 245-252.

Ross, S. M., & Morrison, G. R. (1989). In search of a happy medium in instructional technology research: Issues concerning external validity, media replication, and learner control. *Educational Technology, Research and Development, 37,* 19-33.

Ross, S. M., & Morrison, G. R. (1993). Using feedback to adapt instruction for individuals. In J. Dempsey & G. Sales (Eds.), *Interactive instruction and feedback*. Englewood Cliffs, NJ: Educational Technology Publications.

Ross, S. M., Morrison, G. R., & O'Dell, J. K. (1989). Uses and effects of learner control of context and instructional support computer-based instruction. *Educational Technology, Research and Development, 37*, 29-39.

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). A comparison of approaches to learning task selection in the training of complex cognitive skills. *Computers in Human Behavior, 22*, 321-333.

Salomon, G. (1983). The differential investment of mental effort in learning from different sources, *Educational Psychologist, 18,* 42-50.

Salomon, G. (1985). Information technologies: What you see is not (always) what you get. *Educational Psychologist, 20,* 207-216.

Savage, R. E., Perlmunter, L. C., & Monty, R. A. (1979). Effect of reduction in the amount of choice and the perception of control on learning. In L. C. Perlmutter & R. A. Monty (Eds.), *Choice and perceived control* (pp. 91-102). Hillsdale, NJ: Erlbaum.

Scheiter, K., & Gerjets, P. (2007). Learner control in hypermedia environments. *Educational Psychology Review, 19,* 285-307.

Schnackenberg, H. L., & Sullivan, H. J. (2000). Learner control over full and lean computer-based instruction under differing ability levels. *Educational Technology, Research and Development, 48*, 19-35.

Schraw, G., Flowerday, T., & Lehman, S. (2001). Increasing situational interest in the classroom. *Educational Psychology Review, 13,* 211-224.

Schraw, G., Flowerday, T., & Reisetter, M. F. (1998). The role of choice in reader engagement. *Journal of Educational Psychology, 90*, 705-714.

Schwartz, B. (2004). *The paradox of choice: Why more is less*. New York: Ecco.

Shute, V. J., & Gawlick, L. A. (1995). Practice effects of skill acquisition, learning outcome, retention, and sensitivity to relearning. *Human Factors*, *37*, 781–803.

Shute, V., & Towle, B. (2003). Adaptive e-learning. *Educational Psychologist, 38*, 105-114.

Skinner, E. A. (1996). A guide to constructs of control. *Journal of Personality and Social Psychology, 71*, 549-570.

Snow, R. E. (1980). Aptitude, learner control, and adaptive instruction. *Educational Psychologist, 15,* 151-158.

Song, S. H., & Keller, J. M. (2001). Effectiveness of motivational adaptive computer-assisted instruction on the dynamic aspects of motivation*. Educational Technology, Research and Development, 49*, 5-22.

Straetmans, G., Sluijsmans, D., Bolhuis, B., & van Merriënboer, J. J. G. (2003). Integratie van instructie en assessment in competentiegericht onderwijs [Integration of instruction and assessment in competence based education]. *Tijdschrift voor Hoger Onderwijs*, *3,* 171-197.

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

Sweller, J. (1994). Cognitive theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295-312.

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

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*, 193-210.

Tennyson, R. D., & Buttery, T. (1980). Advisement and management strategies as design variables in computer-assisted instruction. *Educational Communication and Technology Journal*, *28*, 169-176.

Ullmann-Margalit, E., & Morgenbesser, S. (1997). Picking and choosing. *Social Research, 44,* 757–785.

Van Gerven, P. W. M. (2002). *Efficient complex skill training into old age: Exploring the benefits of cognitive load theory*. Doctoral dissertation, Maastricht University, Maastricht, The Netherlands.

Van Gog, T., Paas, F., & van Merriënboer, J. J. G. (2004). Process-oriented worked examples: Improving transfer performance through enhanced understanding. *Instructional Science, 32,* 83-98.

Van Gog, T., Paas, F., & van Merriënboer, J. J. G. (2006). Effects of process-oriented worked examples on troubleshooting transfer performance. *Learning and Instruction, 16*, 154-164.

Van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 43*, 1-11.

Van Gog, T., Paas, F., & van Merriënboer, J. J. G. (in press). Effects of studying sequences of process-oriented and product-oriented worked examples on troubleshooting transfer efficiency. *Learning and Instruction.*

Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills*. Englewood Cliffs, NJ: Educational Technology Publications.

Van Merriënboer, J. J. G., Clark, R., & de Croock, M. B. M. (2002). Blueprints for complex learning: The 4C/ID-model. *Educational Technology, Research and Development*, *50*, 39-64.

Van Merriënboer, J. J. G., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology, 20,* 1-10.

Van Merriënboer, J. J. G., & Kirschner, P. (2001). Three worlds of instructional design: State of the art and future directions. *Instructional Science, 29,* 429-441.

Van Merriënboer, J. J. G., & Kirschner, P. (2007). *Ten steps to complex learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist, 38,* 5-13.

Van Merriënboer, J. J. G., & Luursema, J. (1996). Implementing instructional models in computer-based learning environments: A case study in problem selection. In T. Liao (Ed.), *Advanced educational technology: Research issues and future potential (*pp. 184-206). Berlin, Germany: Springer Verlag.

Van Merriënboer, J. J. G., Schuurman, J., de Croock, M. B. M., & Paas, F. (2002). Redirecting learner's attention during training: Effects on cognitive load, transfer test performance, and training efficiency. *Learning and Instruction*, *12*, 11-37.

Van Merriënboer, J. J. G., Sluijsmans, D. M. A., Corbalan, G., Kalyuga, S., Paas, F., & Tattersall, C. (2006). Performance assessment and learning task selection in complex learning environments. In J. Elen & R. E. Clark (Eds.), *Handling complexity in learning environments: Theory and research* (pp. 201-220). Amsterdam, The Netherlands: Elsevier.

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

Volet, A. (1997). Cognitive and affective variables in academic learning: The significance of direction and effort in students' goals. *Learning and Instruction, 7,* 235-254.

Vosniadou, S., & Ortony, A. (1989). Similarity and analogical reasoning: A synthesis. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 1–17). Cambridge: Cambridge University Press.

Williams, M. D. (1996). Learner-control and instructional technologies. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 957-982). New York: Simon & Schuster Macmillan.

Wolters, C. A. (2003). Regulation of motivation: Evaluating an underemphasized aspect of self-regulated learning. *Educational Psychologist, 38,* 189-205.

Wong, W. K., Chan, T. W., Chou, C. Y., Heh, J. S., & Tung, S. H., (2003). Reciprocal tutoring using cognitive tools. *Journal of Computer Assisted Learning, 19,* 416-428.

Wouters, P., Tabbers, H. K., & Paas, F. (2007). Interactivity in video-based models. *Educational Psychology Review, 19,* 327-342.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, *25*, 82-91.

Zimmerman, B. J. (2002). Achieving academic excellence: A self-regulatory perspective. In M. Ferrari (Ed.), *The pursuit of excellence through education* (pp. 85-110). Mahwah, NJ: Erlbaum.

Summary

Modern education emphasizes the need to flexibly personalize learning tasks to the needs and interests of individual learners. Rather than one curriculum for all learners, such approaches allow each learner to have her own curriculum. The sequence of learning tasks performed by the learners can be personalized either by a computer program, by the learners themselves, or by a combination of both, that is, the program and the learner may share control over task selection. The aim of the studies presented in this dissertation is to investigate under which conditions learner control over task selection is optimized.

Chapter 2 introduces a personalized task selection model with shared instructional control. Taking the 4C/ID-model (van Merriënboer, 1997) as a starting point, the *first* component comprises (a) task characteristics – level of complexity (i.e., from easy to difficult), embedded learner support (i.e., from full support to no support), and other task features (e.g., surface features) – and (b) learner characteristics – task performance and invested mental effort – documented in a learner portfolio. The *second* component refers to the personalization mechanism. Program controlled instruction includes task-selection rules used by an instructional agent to base its decisions on. Learner controlled instruction lets the learner select the learning tasks from a smaller or larger subset of - pre-selected - tasks. The *third* component includes the learning-task database with tasks with diverse levels of complexity, embedded support, and other task features. The model combines the strong points of program control and learner control into a model with shared control over task selection, which is expected to make learning more effective (i.e., higher transfer test performance), more efficient (i.e., higher transfer test performance combined with less invested mental effort), and more motivating. Chapter 2 also reports the results of a pilot study carried out with a twofold purpose: (a) to test whether adaptive task selection with shared control yields better results than adaptive task selection with full program control, and (b) to test a web application developed according to the model.

Taking the model described in Chapter 2 as a basis, Chapter 3 (domain: dietetics) describes a study in which 55 health science students participated in an experiment with a 2 x 2 factorial design with the factors adaptation (present or absent) and control over task selection (program control or shared control). It was predicted that adapting the selection of tasks to learner variables would lower cognitive load to an acceptable level, which enables the allocation of learners' freed-up cognitive resources to learning. This hypothesis was confirmed. Specifically, the results show that adapting task difficulty and support to learners' level of competence and perceived task load leads to more effective and efficient learning. However, training time in the non-adaptive conditions was unexpectedly low. This could be explained by a lack of learners' willingness to invest more time in training, probably because

in comparison with participants in the adaptive conditions they could not perceive the relationship between their performance and the levels of difficulty and support of the tasks they were working on. The second hypothesis stated that shared control over task selection – which provided learners a choice amongst three tasks varying in surface features – would enhance motivation. As expected shared control yielded higher task involvement (i.e., higher learning outcomes combined with more effort directly invested in learning). However, learners' interest in training was not enhanced. This was partially explained by the relatively small amount of choice provided and the fact that learner control was limited to task selection.

Chapters 4 and 5 (domain: genetics) tested the effects of control over task selection based on surface task features (shared or learner control and program control). The study described in Chapter 4 is based on the assumption that the positive effects of learner control decrease when learners do not perceive the control given to them, make didactically unsound choices, or are overwhelmed by the amount of choice. Ninety-four students participated in a 2 x 2 factorial experiment with the factors control (program, shared) and variability of surface features (low, high). Two interaction effects reveal that learner control over surface features of selected tasks enhances transfer test performance and task involvement, but *only* when surface tasks features *differ* from the previous task. When learners were required to make a selection amongst highly similar tasks, transfer test performance and task involvement were hindered. The notion of perception of control was used to account for this effect, although no direct measures of perceived control were included in this study. Again, no differential effects were found on learners' interest in training or the other motivational scales measured after the training. Measuring motivation during training, instead of after completing the whole training, could have been a more sensitive measure of motivation. Reported self-efficacy was found to be higher in the conditions with program control, which seems to indicate that when the sequence of tasks is extrinsically controlled, learners have more confidence that they will be able to perform the tasks.

Similarly, Chapter 5 describes a 2 x 2 factorial experiment carried out with 72 participants, to study the effects of control over the selection of learning tasks that differ in surface task features (program control or learner control) and control over the selection of learning tasks that differ in structural task features (program control or learner control) of a series of completion tasks. Whereas in the study reported in Chapter 4 learners could select from a range of tasks with either similar surface features *or* dissimilar surface features as compared to each prior task, in the study reported in Chapter 5 learners could select from a range of tasks with both similar *and* dissimilar surface features and/or similar *and* dissimilar structural features. It was predicted that learner control over the selection of tasks with salient surface features would enable learners to select a varied set of personally relevant tasks which fosters learning and transfer. This hypothesis was supported for efficiency on the far transfer test. It was argued that learners who were given control over surface features probably constructed general cognitive schemas which enabled them to flexibly apply the learned solution procedure to solve unfamiliar inheritance tasks. In addition, learner control over the selection of tasks with non-salient structural features does not enable learners to select a varied set of personally relevant tasks and therefore was not expected to yield beneficial effects on learning. As expected, learner controlled selection of tasks with different structural features did not enhance learning. It was concluded that learners should be explicitly supported in recognizing structural task features.

Consequently, a final study (Chapter 6; domain: genetics) with 118 students investigates whether feedback on the structural features (operationalized as 'knowledge of correct response') would support learners to recognize those features. Feedback was found to enhance efficiency on the near transfer test, although the expected pattern found on the far transfer test failed to reach significance. Probably feedback supported learners in acquiring more or less automated cognitive schemas that allowed them to perform the steps of the near transfer test as 'routines'. However, far transfer does not allow learners to merely apply a routine but deep understanding of the rationale behind the solution steps is crucial. In addition, in agreement with our predictions the provision of feedback made training in general as well as learner controlled selection of tasks with different structural features more motivating and relevant for learners. However, no support was found for the hypothesis stating that in combination with learner control over task selection, feedback would enhance efficiency. No differences amongst participants in the learner controlled conditions were found on efficiency. This seems to support the idea that less experienced learners are not able to make effective task selections on the basis of structural features and must thus be guided in how to achieve learning objectives.

Finally, Chapter 7 presents an overview and a general discussion of the results of the studies presented in Chapters 3 to 6. The general conclusions are: (a) it is advisable to adapt task difficulty and support to learners' expertise, (b) learners should be provided with control over the selection of tasks that differ in surface features, provided that choices do not include surface features that are very similar to the surface features of the prior task, (c) learners do not benefit from control over the selection of tasks that differ in non-salient structural features unless they are supported in recognizing those features, and (d) feedback on structural features during task practice may supply this support. The provision of feedback made the given choices more motivating for the learners, but additional feedback strategies are needed to support learners in making an optimal selection of tasks with different structural features. In addition, Chapter 7 provides some explanations for unexpected results, followed by a discussion of some methodological issues concerning the efficiency measures and task involvement measures used in the study described in Chapter 3. The final Chapter closes with some considerations for future research and practical implications. A more fine-tuned task selection model might include other variables than mental effort and performance, such as time on task and motivation, for task selection purposes. Such a fine-tuned model should also increase the amount of learner control along with learners' expertise. Future studies must include direct measures of perceived control and personal relevance, as well as immediate and delayed measures of transfer of learning. Implications concern the further examination of the effects of perceived control in educational settings, the importance of the role of salient surface features, and the potential effects of dynamic task selection especially for elderly people.

Samenvatting

In het moderne onderwijs wordt benadrukt dat leertaken flexibel moeten worden afgestemd op de behoeften en interesses van de individuele student. In plaats van één curriculum voor iedereen kan iedere student dan zijn of haar eigen curriculum volgen. De volgorde van de leertaken die de studenten uitvoeren, kan worden gepersonaliseerd door een computerprogramma, of door de studenten zelf, of door een combinatie daarvan (waarbij het programma en de student de sturing van de taakselectie samen doen). In de studies die in dit proefschrift worden beschreven, is onderzocht onder welke omstandigheden zelfsturing van de taakselectie optimaal is.

Hoofdstuk 2 introduceert een gepersonaliseerd taakselectiemodel met gezamenlijke sturing, gebaseerd op het 4C/ID-model (Van Merriënboer, 1997). De *eerste* component omvat (a) taakkenmerken: complexiteit (van gemakkelijk tot moeilijk), ingebouwde ondersteuning (van volledige ondersteuning tot geen enkele ondersteuning) en andere taakkenmerken (bijv. oppervlaktekenmerken); en (b) studentkenmerken: taakprestatie en mentale inspanning, die worden vastgelegd in een studentportfolio. De *tweede* component heeft betrekking op het personalisatiemechanisme. Programmagestuurde instructie omvat taakselectieregels die een instructiegever gebruikt om zijn beslissingen op te baseren. Bij zelfgestuurde instructie kiest de student de leertaken uit een kleinere of grotere subset van – vooraf geselecteerde – taken. De *derde* component omvat de leertakendatabase met taken van diverse complexiteitsniveaus, ingebouwde ondersteuning en andere taakkenmerken. Het model combineert de sterke punten van programmasturing en zelfsturing tot een model met een gezamenlijke sturing van de taakselectie, dat het leerproces naar verwachting effectiever (betere transfer test-prestaties), efficiënter (betere transfer test-prestaties én minder mentale inspanning) en motiverender zal maken. Hoofdstuk 2 beschrijft ook de resultaten van een pilotstudie die werd uitgevoerd met het tweeledige doel om: (a) te toetsen of adaptieve taakselectie met gezamenlijke sturing betere resultaten oplevert dan adaptieve taakselectie met volledige programmasturing, en (b) een op basis van dit model ontwikkelde webapplicatie te toetsen.

Aan de hand van het in hoofdstuk 2 uitgewerkte model beschrijft hoofdstuk 3 (vakgebied: diëtetiek) een studie waarbij 55 studenten gezondheidskunde deelnamen aan een 2 x 2 factorieel experiment met als factoren adaptatie (aanwezig of afwezig) en sturing van taakselectie (programmasturing of gezamenlijke sturing). De veronderstelling was dat het aanpassen van de taakselectie aan studentvariabelen de cognitieve belasting zou verlagen tot een aanvaardbaar niveau, waardoor de student de vrijgemaakte cognitieve capaciteit zou kunnen gebruiken om te leren. Deze hypothese werd bevestigd. Meer specifiek laten de resultaten zien dat wanneer de taakmoeilijkheid en ondersteuning worden aangepast aan het competentieniveau en de vermeende taakbelasting van de studenten, dit tot effectiever en efficiënter leren leidt. Zonder aanpassing daarvan bleek de trainingstijd echter onverwacht kort. Dit zou verklaard kunnen worden door een gebrek aan bereidheid onder de studenten om meer tijd te investeren in de training, waarschijnlijk omdat zij in vergelijking met de deelnemers voor wie de aanpassing wél gold, geen verband zagen tussen hun prestatie en de moeilijkheidsgraad van en de ondersteuning bij de taken waaraan ze werkten. De tweede hypothese was dat gezamenlijke sturing van de taakselectie – waarbij de studenten de keuze kregen tussen drie taken met verschillende oppervlaktekenmerken – de motivatie zou verbeteren. Zoals verwacht, leidde gezamenlijke sturing tot meer taakbetrokkenheid (d.w.z. betere leerresultaten gecombineerd met meer directe leerinspanning). De belangstelling van de studenten voor de training nam echter niet toe. Dit werd deels verklaard door de relatief geringe keuzemogelijkheden en het feit dat de zelfsturing was beperkt tot de taakselectie.

De hoofdstukken 4 en 5 (vakgebied: genetica) gaan in op de effecten van sturing van taakselectie op basis van de oppervlaktekenmerken van taken (gezamenlijke of zelfsturing en programmasturing). De studie die in hoofdstuk 4 wordt beschreven, gaat uit van de veronderstelling dat de positieve effecten van zelfsturing afnemen als studenten niet doorhebben welke sturingsmogelijkheden zij hebben, of als zij didactisch onverantwoorde keuzes maken, of overweldigd worden door het aantal keuzes. Vierennegentig studenten namen deel aan een 2 x 2 factorieel experiment met als factoren sturing (programma, gezamenlijk) en variabiliteit van oppervlaktekenmerken (laag, hoog). Twee interactie-effecten laten zien dat zelfsturing van oppervlaktekenmerken van geselecteerde taken een positief effect heeft op de transfer test-prestaties en de taakbetrokkenheid, maar *alleen* als de oppervlaktekenmerken van de taken *verschillen* van de vorige taak. Wanneer studenten moesten kiezen tussen sterk overeenkomende taken, had dit een negatief effect op de transfertestprestaties en de taakbetrokkenheid. Dit effect werd verklaard aan de hand van het begrip ‘perceptie van controle’, hoewel dit in deze studie niet direct gemeten werd. Ook hier was geen sprake van differentiële effecten op de belangstelling van studenten voor de training of op de andere motivatie-indicatoren die na afloop van de training werden gemeten. Het was wellicht beter geweest om de motivatie tijdens de training te meten in plaats van na afronding van de hele training. De self-efficacy bleek hoger te zijn in het geval van programmasturing, hetgeen erop lijkt te duiden dat wanneer de volgorde van taken van buitenaf wordt gestuurd, studenten er meer vertrouwen in hebben dat zij in staat zullen zijn de taken uit te voeren.

Hoofdstuk 5 beschrijft een soortgelijk 2 x 2 factorieel experiment met 72 deelnemers. Dit was bedoeld om na te gaan wat de effecten zijn van sturing van de selectie van leertaken die qua oppervlaktekenmerken verschillen (programmasturing of zelfsturing) en sturing van de selectie van leertaken die qua structurele kenmerken verschillen (programmasturing of zelfsturing) bij een reeks completeeropdrachten. Terwijl studenten in de in hoofdstuk 4 beschreven studie konden kiezen uit een reeks taken met dezelfde *of* andere oppervlaktekenmerken dan iedere voorgaande taak, konden zij in de in hoofdstuk 5 beschreven studie kiezen uit een reeks taken met dezelfde *en* andere oppervlaktekenmerken en/of dezelfde *en* andere structurele kenmerken. Aangenomen werd dat, bij zelfsturing van de selectie van taken met opvallende oppervlaktekenmerken, de studenten in staat zouden zijn om een gevarieerd geheel van voor hen persoonlijk relevante te taken te selecteren dat bevorderlijk zou zijn voor leren en transfer. Deze hypothese werd bewezen voor efficiency bij de verre transfer test. Gesteld werd dat studenten die taakselectie konden sturen op basis van oppervlaktekenmerken, vermoedelijk algemene cognitieve schema's construeerden om de geleerde oplossingsprocedure flexibel toe te kunnen passen voor het oplossen van onbekende erfelijkheidsopdrachten. Bij zelfsturing van de selectie van taken met niet-opvallende structurele kenmerken zijn studenten niet in staat om een gevarieerd geheel van voor hen persoonlijk relevante taken te selecteren, en daarom werd niet verwacht dat dit gunstige leereffecten zou opleveren. Zelfgestuurde selectie van taken met verschillende structurele kenmerken leidde inderdaad niet tot beter leren. Geconcludeerd werd dat studenten expliciet moeten worden geholpen om structurele taakkenmerken te herkennen.

Vervolgens onderzoekt een laatste studie (hoofdstuk 6; vakgebied: genetica) met 118 studenten of feedback op de structurele kenmerken (geoperationaliseerd als 'informatie over het juiste antwoord') studenten zou helpen om die kenmerken te herkennen. Feedback bleek de efficiency te vergroten bij de nabije transfer test, hoewel voor het verwachte patroon bij de verre transfer test geen significant effect optrad. Waarschijnlijk hielp de feedback de studenten om min of meer geautomatiseerde cognitieve schema's te verwerven waarmee zij de stappen van de verre transfer test 'routinematig' konden uitvoeren. Bij verre transfer zijn studenten echter niet in staat om domweg routinematig te werk te gaan, maar is diepgaand inzicht in het principe achter de oplossingsstappen cruciaal. Geheel volgens onze voorspellingen maakte het geven van feedback bovendien de training in het algemeen en de zelfgestuurde selectie van taken met verschillende structurele kenmerken motiverender en relevanter voor studenten. Er werd echter geen onderbouwing gevonden voor de hypothese dat feedback in combinatie met zelfsturing van de taakselectie de efficiency zou vergroten. Ten aanzien van efficiency werden geen verschillen gevonden onder de deelnemers voor wie zelfsturing gold. Dit lijkt te bevestigen dat minder ervaren studenten niet in staat zijn tot effectieve taakselectie op basis van structurele kenmerken en dus moeten worden geholpen bij het bereiken van hun leerdoelen.

Hoofdstuk 7 omvat tot slot een overzicht en een algemene bespreking van de resultaten van de in hoofdstuk 3 t/m 6 beschreven studies. De algemene conclusies zijn: (a) het is raadzaam om taakmoeilijkheid en ondersteuning aan te passen aan het kennisniveau van studenten, (b) studenten moeten controle hebben over de selectie van taken die qua oppervlaktekenmerken verschillen, mits daarbij geen sprake is van oppervlaktekenmerken die sterk overeenkomen met die van de voorgaande taak, (c) studenten hebben geen baat bij zelfsturing van de selectie van taken die qua niet-opvallende structurele kenmerken verschillen, tenzij ze hulp krijgen bij het herkennen van die kenmerken, en (d) feedback over structurele kenmerken tijdens het uitvoeren van taken kan in deze hulp voorzien. Het geven van feedback maakte de gegeven keuzes motiverender voor de studenten, maar er zijn aanvullende feedbackstrategieën nodig om studenten te helpen bij het maken van een optimale selectie van taken met verschillende structurele kenmerken. Daarnaast biedt hoofdstuk 7 een aantal verklaringen voor onverwachte resultaten, gevolgd door een bespreking van enkele methodologische aspecten met betrekking tot de in hoofdstuk 3 gebruikte maten voor efficiency en taakbetrokkenheid. Het laatste hoofdstuk wordt afgesloten met een aantal overwegingen voor toekomstig onderzoek en praktische implicaties. Een verfijnder taakselectiemodel zou naast mentale inspanning en prestatie ook andere variabelen voor taakselectiedoeleinden kunnen omvatten, zoals ‘time-on-task’ en motivatie. Een dergelijk verfijnd model zou ook de mate van zelfsturing moeten vergroten naargelang van het kennisniveau van de student. Toekomstige studies zouden directe maten voor ‘perceptie van controle’ en persoonlijke relevantie moeten omvatten, en ook directe en uitgestelde maten voor de transfer van leren. Mogelijke implicaties hebben betrekking op nader onderzoek van de effecten van perceptie van controle in onderwijsomgevingen, het belang van de rol van opvallende oppervlaktekenmerken, en de potentiële effecten van dynamische taakselectie, met name voor ouderen.

Resumen

Los métodos de enseñanza actuales enfatizan la necesidad de personalizar, de una manera flexible, las tareas de aprendizaje según las necesidades e intereses de cada alumno. Más que un currículum para todos los alumnos, cada uno debe disponer de su propio currículum. La secuencia de las tareas de aprendizaje ejecutadas por los alumnos puede ser personalizada por un programa (en inglés, *program control*), por ellos mismos (*learner control*), o por una combinación de ambos; es decir, el programa y el alumno pueden compartir el control en la selección de tareas de aprendizaje (*shared control*). El objetivo principal de esta tesis es investigar qué condiciones ayudan a los alumnos a controlar mejor la selección de tareas de aprendizaje.

Para conseguirlo, el Capítulo 2 propone un modelo de personalización de tareas de aprendizaje basado en el control compartido. El modelo incluye tres elementos y se basa en el modelo de diseño instruccional *Four Components Instructional Design* Model (4C/ID; van Merriënboer, 1997). El primer elemento del modelo propuesto en el Capítulo 2 incluye (a) características de las tareas de aprendizaje: el nivel de dificultad (de fácil a difícil), la ayuda al alumno (*learner support*; de ayuda completa a sin ayuda), y otras características de las tareas de aprendizaje (por ejemplo, características superficiales o *surface features* que en principio son irrelevantes para resolver una tarea); y (b) características del alumno: el rendimiento y esfuerzo cognitivo (*mental effort*) documentado en el porfolio del alumno. El segundo elemento se refiere al mecanismo de personalización. Cuando el programa lleva a cabo la selección de tareas, incluye reglas que son utilizadas por un agente instruccional con el fin de decidir qué tareas deben seleccionarse. Cuando la selección de tareas es llevada a cabo por el alumno, éste las elige basándose en un subset de tareas previamente seleccionadas por el programa. El tercer elemento es una base de datos que incluye tareas de distintos niveles de dificultad, ayuda, y otras características. El modelo combina los puntos más fuertes del control llevado a cabo por el programa y por el alumno para seleccionar tareas de aprendizaje a través del control compartido, para facilitar un aprendizaje más efectivo (mayor rendimiento en la transferencia de conocimientos), más eficiente (mayor rendimiento en la transferencia de conocimientos combinado con menor esfuerzo cognitivo), y más motivante. El Capítulo 2 también reporta los resultados de un estudio piloto llevado a cabo con dos propósitos: (a) comprobar si la selección adaptativa de tareas combinada con el control compartido produce mejores resultados que la selección adaptativa de tareas combinada con un control total por parte del programa, y (b) examinar una aplicación web desarrollada según el modelo propuesto.

Partiendo del modelo descrito en el Capítulo 2, el Capítulo 3 describe un estudio (que toma la Dietética como dominio base) en el que 55 estudiantes de

Ciencias de la Salud participaron en un experimento factorial 2 x 2 con los factores Adaptación (presente o ausente) y Control sobre la selección de tareas (programa o compartido). La hipótesis planteada sostiene que, adaptando la selección de las tareas a las características de cada alumno, se reduciría la carga cognitiva a un nivel aceptable, lo que permitiría a los alumnos destinar al aprendizaje los recursos cognitivos que no utilizó. Esta hipótesis fue confirmada. De manera más específica, los resultados permiten observar que el hecho de adaptar el nivel de dificultad y de ayuda al nivel de competencias del alumno y, además, a la carga cognitiva percibida por el alumno, origina un aprendizaje más efectivo y eficiente. Sin embargo, el tiempo utilizado en las condiciones experimentales donde no existía dicha adaptación fue inesperadamente bajo. Esto podría deberse a la falta de disposición de los alumnos de invertir más tiempo en su formación, probablemente porque en comparación con los participantes que trabajaron en las condiciones donde hubo adaptación, aquellos no tuvieron la oportunidad de observar la relación entre su rendimiento y los niveles de dificultad y ayuda asociadas a las tareas recibidas. En una segunda hipótesis se predijo que el control compartido en la selección de tareas - ofreciendo al alumno tres opciones a elegir variando en las características superficiales - aumentaría su motivación. Tal y como se había planteado en la hipótesis, el control compartido conllevó una mayor implicación de los alumnos en las tareas de aprendizaje (mejores resultados con un esfuerzo cognitivo mayor invertido directamente al aprendizaje). Los alumnos, sin embargo, mostraron un interés similar en todas las condiciones experimentales, probablemente por el número reducido de alternativas ofrecidas y porque el control ofrecido al alumno se limitó a la selección de tareas.

En los Capítulos 4 y 5 (centrados en el campo de la Genética) se investigaron los efectos del control en la selección de tareas en base a sus características superficiales (programa o compartido). El estudio descrito en el Capítulo 4 se basa en la idea de que los efectos positivos de proveer de control al alumno disminuyen cuando éste no percibe el control recibido, sus elecciones no son didácticamente correctas, o el número elevado de elecciones ofrecidas los sobrecarga cognitivamente. Noventa y cuatro alumnos participaron en un experimento factorial 2 x 2 en el que los factores eran el control (programa o compartido) y la variabilidad de las características superficiales de las tareas (baja o alta). Dos efectos de interacción revelan que proveer de control al alumno sobre las características superficiales mejora la transferencia de conocimientos y la implicación de los alumnos, pero únicamente cuando las características superficiales difieren de las de la tarea anterior. Cuando los alumnos deben seleccionar entre tareas de aprendizaje demasiado similares, la transferencia de conocimientos y su implicación se pueden ver afectadas negativamente. La percepción del control ofrecido (*perceived control*) puede justificar este efecto, aunque en este estudio no se incluyó ninguna medida directa. Una vez más, no se encontraron diferencias en el interés o en las otras escalas incluidas posteriormente a la instrucción para medir la motivación. La evaluación de la motivación durante la instrucción, en lugar de después de la instrucción, podría haber sido una medida más sensible de la motivación. La auto-eficacia (*self-efficacy*) resultó superior cuando es el programa el responsable de seleccionar las tareas. Este hecho parece indicar que cuando la secuencia de tareas a ejecutar ha sido realizada por un agente externo, los alumnos confían más en su propia capacidad para realizar las tareas.

Del mismo modo, el Capítulo 5 describe un experimento factorial 2 x 2 en el que participaron 72 alumnos. En éste se estudiaron si los efectos del control sobre una serie de tareas que deben completarse difieren según si el control ofrecido a los alumnos es sobre las características superficiales (programa o alumno) o sobre las características estructurales (programa o alumno). Mientras que en el estudio descrito en el Capítulo 4 los alumnos debían elegir entre un abanico de tareas con características superficiales similares o distintas de las de la tarea ejecutada anteriormente, en el estudio descrito en el Capítulo 5 los alumnos tenían la opción de elegir entre tareas con características superficiales y estructurales similares y distintas con respecto a la tarea ejecutada anteriormente. La hipótesis fue que el control del alumno sobre la selección de tareas en función de las características superficiales le permitiría seleccionar un conjunto de tareas personalmente relevantes para ellos, lo que le ayudaría a mejorar el aprendizaje y la transferencia de conocimientos. Esta hipótesis se confirmó con respecto a la eficiencia en la transferencia lejana (*far transfer*). Los alumnos que eligieron las características superficiales de las tareas construyeron esquemas cognitivos generales que les permitieron utilizar de manera flexible el procedimiento de solución de los problemas (*solution procedure*) aprendido a la hora de resolver ejercicios de Genética que les eran menos familiares. Por otra parte, se predijo que proveer al alumno de control en la selección de tareas en función de sus características estructurales no influiría positivamente en su aprendizaje, ya que la baja perceptibilidad de estas características no les permitirá seleccionar un conjunto variado de tareas relevantes para ellos. Como se planteó en la hipótesis, el control dado al alumno en función de las características estructurales no mejoró su aprendizaje. Se concluyó, por tanto, que los alumnos necesitan ayuda para reconocer estas características.

Finalmente, se llevo a cabo un último estudio, que se describe en el Capítulo 6 (con la Genética nuevamente como campo), con 118 alumnos, para investigar si la provisión de *feedback* sobre las características estructurales (operacionalizado como "conocimiento de respuesta correcta" – *knowledge of correct response*) ayuda a los alumnos a reconocer dichas características. Los resultados muestran que el feedback facilitó la transferencia de conocimientos cercana (*near transfer*). Aunque los resultados de la transferencia lejana muestran el patrón esperado, éste no alcanzó relevancia estadística. Probablemente, el feedback ayudó a los alumnos a adquirir esquemas cognitivos más o menos automatizados que les permitieron realizar los pasos del test en forma de rutinas. Sin embargo, la transferencia lejana requiere más que aplicar una mera rutina, ya que es fundamental entender profundamente la lógica de los pasos para llegar a la solución. Además, tal y como se había sugerido en la hipótesis, proveer feedback trajo como consecuencia que los alumnos percibieran la instrucción como más motivadora y relevante. Sin embargo, la hipótesis que planteaba que la provisión de feedback, en combinación con la provisión de control al alumno sobre la selección de tareas en base a caracteristicas estructurales, ocasionaría un aprendizaje más eficiente, no se confirmó. No se encontraron diferencias significativas entre los participantes de las dos condiciones que dieron control al alumno (con y sin feedback). Esto podría indicar que alumnos con menos experiencia no son capaces de elegir tareas en función de las características estructurales de una manera efectiva y, por tanto, deben contar con apoyo guiado para mejorar sus objetivos educativos.

Para concluir, el Capítulo 7 presenta un resumen y una discusión general de los resultados de los estudios expuestos en esta tesis. En general, las conclusiones son: (a) es recomendable adaptar el nivel de dificultad y ayuda al nivel de experiencia de los alumnos, (b) los alumnos deben tener la posibilidad de elegir sus propias tareas de aprendizaje en función de las características superficiales, siempre y cuando las alternativas ofrecidas no incluyan características superficiales demasiado similares a las de la tarea anterior, (c) los alumnos no se beneficiarán de la elección de tareas basándose en las características estructurales a no ser que se les ayude a reconocer dichas características, y (d) este apoyo puede ofrecerse durante la ejecución de las tareas de aprendizaje mediante feedback sobre las características estructurales. Proporcionar feedback hace que los estudiantes perciban las alternativas ofrecidas como más motivadoras, aunque es necesario implementar estrategias de feedback adicionales con el fin de ayudarles a hacer una elección óptima. Adicionalmente, el Capítulo 7 incluye explicaciones de los resultados no esperados de esta tesis, y prosigue describiendo algunas cuestiones metodológicas relacionadas con las medidas de eficiencia y de implicación utilizadas en el estudio descrito en el Capítulo 3. Este último Capítulo 7 destaca diversas consideraciones para futuras investigaciones y finaliza subrayando una serie de implicaciones prácticas. Un modelo de selección de tareas más elaborado podría incluir otras variables aparte del esfuerzo cognitivo y el rendimiento como, por ejemplo, el tiempo invertido en las tareas y la motivación del alumno. Un modelo más elaborado también debería facilitar al alumno más control a medida que éste adquiera más experiencia. Estudios posteriores deben incluir medidas de percepción de control y relevancia personal, así como medidas de transferencia de conocimiento inmediatas y a largo plazo. Las implicaciones prácticas incluyen investigaciones adicionales sobre los efectos de percepción de control en entornos educativos, la importancia de las características superficiales en educación, y los efectos potenciales de la selección dinámica de tareas en personas de edad avanzada.

ICO dissertation series

In the ICO Dissertation Series dissertations are published of graduate students from faculties and institutes on educational research within the following universities: Eindhoven University of Technology, Leiden University, Maastricht University, Open University of the Netherlands, University of Amsterdam, University of Groningen, University of Twente, Utrecht University, VU University Amsterdam, and Wageningen University (and formerly Radboud University Nijmegen and Tilburg University).

134. Saab, N. (14-10-2005). *Chat and explore: The role of support and motivation in collaborative scientific discovery learning.* Amsterdam: University of Amsterdam.
135. Löhner, S. (11-11-2005). *Computer-based modeling tasks: The role of external representation.* Amsterdam: University of Amsterdam.
136. Beers, P.J. (25-11-2005). *Negotiating common ground: Tools for multidisciplinary teams.* Heerlen: Open University of the Netherlands.
137. Tigelaar, E.H. (07-12-2005). *Design and evaluation of a teaching portfolio.* Maastricht: Maastricht University.
138. Van Drie, J.P., (20-12-2005). *Learning about the past with new technologies. Fostering historical reasoning in computer-supported collaborative learning*. Utrecht: Utrecht University.
139. Walrecht, E.S. (09-01-2006). *Brede innovatie, passende strategie?: De Groninger Vensterschool als casus van onderzoek naar strategie en invoering.* Groningen University of Groningen
140. De Laat, M. (03-02-2006). *Networked learning*. Utrecht: Utrecht University.
141. Prince, C.J.A.H. (21-04-2006). *Problem-based learning as a preparation for professional practice.* Maastricht: Maastricht University.
142. Van Gog, T. (28-04-2006). *Uncovering the problem-solving process to design effective worked examples.* Heerlen: Open University of the Netherlands.
143. Sins, P.H.M. (18-05-2006). *Students' reasoning during computer-based scientific modeling.* Amsterdam: University of Amsterdam.
144. Mathijsen, I.C.H. (24-05-2006). *Denken en handelen van docenten.* Utrecht: Utrecht University.
145. Akkerman, S.F. (23-06-2006). *Strangers in dialogue: Academic collaboration across organizational boundaries.* Utrecht: Utrecht University.
146. Willemse, T.M. (21-08-2006). *Waardenvol opleiden: Een onderzoek naar de voorbereiding van aanstaande leraren op hun pedagogische opdracht.* Amsterdam: VU University Amsterdam.
147. Kieft, M. (19-09-2006). *The effects of adapting writing instruction to students' writing strategies.* Amsterdam: University of Amsterdam.
148. Vreman-de Olde, G.C. (27-09-2006). *Look experiment design: Learning by designing instruction.* Enschede: University of Twente.
149. Van Amelsvoort, M. (13-10-2006). *A space for debate: How diagrams support collaborative argumentation-based learning.* Utrecht: Utrecht University.
150. Oolbekking-Marchand, H. (9-11-2006). *Teachers' perspectives on self-regulated learning: An exploratory study in secondary and university education.* Leiden: Leiden University.
151. Gulikers, J. (10-11-2006). *Authenticity is in the eye of the beholder: Beliefs and perceptions of authentic assessment and the influence on student learning*. Heerlen: Open University of the Netherlands.
152. Henze, I. (21-11-2006). *Science teachers' knowledge development in the context of educational innovation.* Leiden: Leiden University.
153. Van den Bossche, P. (29-11-2006). *Minds in teams: The influence of social and cognitive factors on team learning.* Maastricht: Maastricht University.
154. Mansvelder-Longayroux, D.D. (06-12-2006). *The learning portfolio as a tool for stimulating reflection by student teachers.* Leiden: Leiden University

155. Visschers-Pleijers, A.J.S.F. (19-01-2007). *Tutorial group discussion in problem-based learning: Studies on the measurement and nature of learning-oriented student interactions.* Maastricht: Maastricht University.
156. Poortman, C.L. (16-02-2007). *Workplace learning processes in senior secondary vocational education.* Enschede: University of Twente.
157. Schildkamp. K.A. (15-03-2007). *The utilisation of a self-evaluation instrument for primary education.* Enschede: University of Twente.
158. Karbasioun, M. (20-04-2007). *Towards a competency profile for the role of instruction of agricultural extension professionals in Asfahan.* Wageningen: Wageningen University.
159. Van der Sande, R.A.W. (04-06-2007). *Competentiegerichtheid en scheikunde leren: Over metacognitieve opvattingen, leerresultaten en leeractiviteiten.* Eindhoven: Eindhoven University of Technology.
160. Pijls, M. (13-06-2007). *Collaborative mathematical investigations with the computer: Learning materials and teacher help.* Amsterdam: University of Amsterdam.
161. Könings, K. (15-06-2007). *Student perspectives on education: Implications for instructional design.* Heerlen: Open University of the Netherlands.
162. Prangsma, M.E. (20-06-2007). *Multimodal representations in collaborative history learning.* Utrecht: Utrecht University.
163. Niemantsverdriet, S. (26-06-2007). *Learning from international internships: A reconstruction in the medical domain.* Maastricht: Maastricht University.
164. Van der Pol, J. (03-07-2007). *Facilitating online learning conversations: Exploring tool affordances in higher education.* Utrecht: Utrecht University.
165. Korobko, O.B. (07-09-2007). *Comparison of examination grades using item response theory: A case study*. Enschede: University of Twente.
166. Madih-Zadeh, H. (14-09-2007). *Knowledge construction and participation in an asynchronous computer-supported collaborative learning environment in higher education.* Wageningen: Wageningen University
167. Budé, L.M. (05-10-2007). *On the improvement of students' conceptual understanding in statistics education.* Maastricht: Maastricht University.
168. Meirink, J.A. (15-11-2007). *Individual teacher learning in a context of collaboration in teams.* Leiden: Leiden Univesity.
169. Niessen, T.J.H. (30-11-2007). *Emerging epistemologies: Making sense of teaching practices.* Maastricht: Maastricht University.
170. Wouters, P. (07-12-2007). *How to optimize cognitive load for learning from animated models.* Heerlen: Open University of the Netherlands.
171. Hoekstra, A. (19-12-2007). *Experienced teachers' informal learning in the workplace.* Utrecht: Utrecht University.
172. Munneke-de Vries, E.L. (11-01-2008). *Arguing to learn: Supporting interactive argumentation through computer-supported collaborative learning*. Utrecht: Utrecht University.
173. Nijveldt, M.J. (16-01-2008). *Validity in teacher assessment. An exploration of the judgement processes of assessors*. Leiden: Leiden University.
174. Jonker, H.G. (14-02-2008). *Concrete elaboration during knowledge acquisition*. Amsterdam: VU University Amsterdam.
175. Schuitema, J.A. (14-02-2008). *Talking about values. A dialogue approach to citizenship education as an integral part of history classes*. Amsterdam: University of Amsterdam.
176. Janssen, J.J.H.M. (14-03-2008). *Using visualizations to support collaboration and coordination during computer-supported collaborative learning*. Utrecht: Utrecht University.
177. Honingh, M.E. (17-04-2008). *Beroepsonderwijs tussen publiek en privaat: Een studie naar opvattingen en gedrag van docenten en middenmanagers in bekostigde en niet-bekostigde onderwijsinstellingen in het middelbaar beroepsonderwijs.* Amsterdam: University of Amsterdam.
178. Baartman, L.K.J. (24-04-2008). *Assessing the assessment: Development and use of quality criteria for competence assessment programmes*. Utrecht: Utrecht University.