University of Amsterdam

UvA-DARE (Digital Academic Repository)

Comprehending process diagrams in biology education

Kragten, M.

Publication date
2015

Document Version
Final published version

Link to publication

Citation for published version (APA):
Kragten, M. (2015). *Comprehending process diagrams in biology education*. [Thesis, externally prepared, Universiteit van Amsterdam].

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

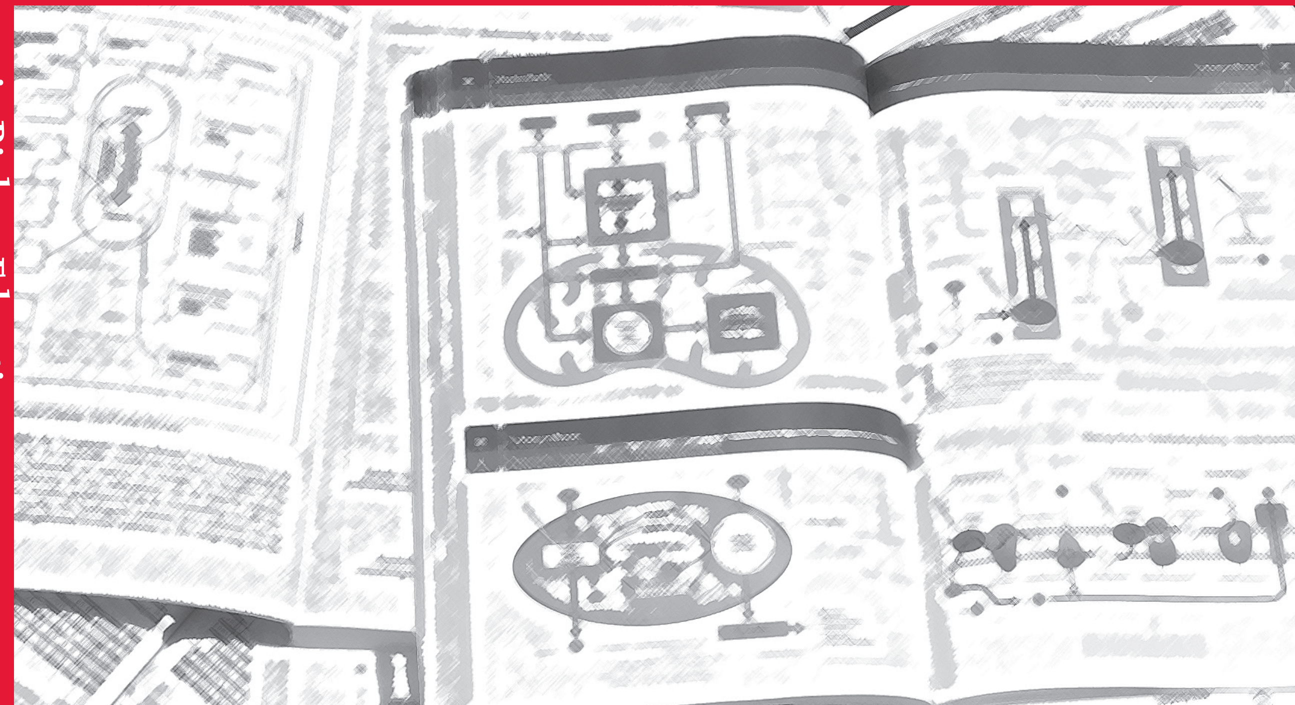
UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:06 Jun 2026

Comprehending Process Diagrams in Biology Education

Marco Kragten

Marco Kragten Comprehending Process Diagrams in Biology Education



University of Amsterdam

COMPREHENDING PROCESS DIAGRAMS IN BIOLOGY EDUCATION

Marco Kragten

Cover design: Monique van Hootegem
Printed by PrintPartners Ipskamp B.V., Amsterdam.

ISBN 978-90-9029317-2

This work is part of the Doctoral Grant for Teachers programme, which is financed by the Netherlands Organisation for Scientific Research (NWO).

COMPREHENDING PROCESS DIAGRAMS IN BIOLOGY EDUCATION

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel

op vrijdag 27 november 2015, te 12.00 uur

door

Marco Kragten

Geboren te Amsterdam

Promotiecommissie

Promotores:	Prof. dr. G.C.W. Rijlaarsdam Universiteit van Amsterdam
	Prof. dr. W.F. Admiraal Universiteit Leiden
Overige leden:	Prof. dr. M.L.L. Volman Universiteit van Amsterdam
	Prof. dr. C.A.M. van Boxtel Universiteit van Amsterdam
	Prof. dr. P.F. de Jong Universiteit van Amsterdam
	Prof. dr. J.J. Beishuizen Vrije Universiteit
	Prof. dr. F. Paas Erasmus Universiteit Rotterdam
	Prof. dr. J.H. van Driel Universiteit Leiden
Faculteit:	Faculteit der Maatschappij- en Gedragswetenschappen

TABLE OF CONTENTS

CHAPTER 1 Introduction	7
CHAPTER 2 Diagrammatic literacy in secondary science education	13
CHAPTER 3 Students' ability to solve process-diagram problems in secondary biology education	29
CHAPTER 4 Students' learning activities while studying biological process diagrams	43
CHAPTER 5 Effect of a multiple-strategy training with a stepwise working-routine on learning from process diagrams	65
CHAPTER 6 Discussion	85
REFERENCES	93
AUTHOR INDEX	99
SUMMARY	103
SUMMARY (DUTCH)	109
ACKNOWLEDGEMENTS	115

CHAPTER 1

INTRODUCTION

1. PROCESS DIAGRAMS

Process diagrams are a distinct type of diagram that are important in Biology education for explaining processes (e.g., photosynthesis, biogeochemical cycles) which are quite complex to communicate by text only. The number of diagrams that have entered scientific textbooks has increased substantially in the past years (e.g., Bezemer & Kress, 2008; Bowen & Roth, 2002; Campbell & Reece, 2002). Although diagrams are expected to facilitate search for information and to enhance learning (Carlson, Chandler, & Sweller, 2003; Larkin & Simon, 1987; Winn, 1991), students have difficulties with the interpretation of diagrams (e.g., Chittleborough & Treagust, 2007; Schönborn, Andersen, & Grayson, 2002).

Students are often not taught how to comprehend diagrams and difficulties with encoding diagrams are not readily solved by practice on standard, topic-oriented, instructional material (Chittleborough & Treagust, 2008).

1.1 Definition and description of process diagrams[*]

In the present thesis, process diagrams are defined as representations of how systems function by abstractions in components and arrows and that contain information that is spatial, dynamic, and schematic in nature.

The spatial organization of the ecological diagram in Figure 1.1 presents information to the reader that would be more difficult to convey by just text. The communication of information through individual elements (e.g., components, arrows, labels) and the way elements are arranged in space provides perceptual enhancement (Larkin & Simon, 1987). For instance, the location of the component 'atmosphere' in Figure 1.1 at the top of the diagram conveys the relative position of this compartment compared to the other compartments in the actual ecosystem: The component 'zooplankton' might be posted below the other components to provide information about its place in the hierarchical food chain of the ecosystem, i.e., zooplankton in the ecosystem presented in Figure 1.1 is a first- and second-order consumer as they eat algae, dissolved remains of organisms, and bacteria. In a review article, Winn

[*] An adapted version of this section has been published in a Dutch peer-reviewed research journal. Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2013). Geletterdheid in diagrammen in de bètavakken. *Tijdschrift voor Taalbeheersing*, *35*, 63–81.
doi:10.5117/TVT2013.1.KRAG

(1991) concludes that the perception of the configuration and discrimination of components is essential for comprehending diagrams.

Arrows in a process diagram represent the dynamic functioning of a system. In an ecological diagram like Figure 1.1, arrows indicate how components in an ecosystem relate via dynamic processes, e.g., photosynthesis, feeding, dissolving, etc. In ecological diagrams the dynamic functioning of the system is very concrete. In Figure 1.1 for instance the arrows represent the amount of carbon flux per hour. The aspect of time arrows convey might be less concrete in other than ecological process diagrams (e.g., biochemical process diagram), but the arrows always represent a dynamic process. When an arrow does not convey dynamic information, the diagram is not a process diagram. Compare for instance an anatomical depiction where an arrow is used as a pointer to zoom in or out of a specific part of the body: it does not convey dynamic information and therefore is not understood to be a process diagram.

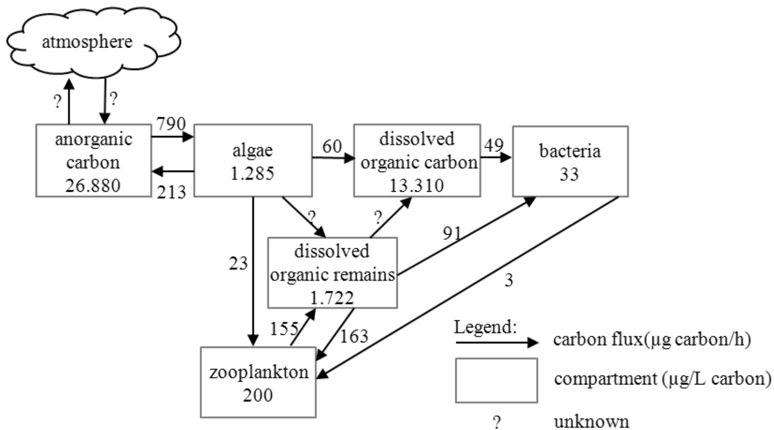


Figure 1.1. Example of a process diagram.

Process diagrams are schematic, i.e., simplified and symbolic, representations. The representation of an ecosystem in Figure 1.1 is simplified because in the real world the various compartments are, obviously, all dispersed throughout the lake. The convention used for the components in this diagram, i.e., text boxes, is symbolic. A less symbolic choice would have been to use iconic depictions of algae or bacteria, albeit a symbolic simplification.

The structure of a process diagram is largely determined by the placement of its components and the arrows that connect the components. Novick and Hurley (2001) analysed the structure of three spatial diagrams (i.e., matrices, networks, and hierarchies) and concluded that networks, i.e., process diagrams in the present study's terminology, have no formal structure and there is not necessarily a start or end point. Any component may be linked to any other component, thereby allowing

closed loops (e.g., feedback), and multiple arrows from one component to another are possible. Furthermore, the links between components can be directional (unidirectional, bidirectional, i.e., arrows), or signal associative meanings (e.g., clusters).

2. RESEARCH QUESTIONS

From the start of this research project, our aim was to design an intervention that facilitates students' comprehension of process diagrams. Previous research showed that students have difficulties with comprehending diagrams, but what difficulties do student have with *process* diagrams in particular? Before designing an effective intervention, we felt the need to better understand these difficulties. We therefore designed three studies to examine students' difficulties with theses diagrams. The general scheme was to study the interaction between three constituting elements: 1) the task, 2) the student, and 3) the diagram design. The first three studies provided insights that informed the design of the final intervention study. The research questions of this thesis are:

1) What is the relationship between features of the (1) task, (2) student, and (3) diagram, on the one hand, and difficulty of a diagram task, on the other hand?
2) What is the relationship between students' learning activities while studying various process diagrams, and their resulting comprehension of these diagrams?
3) What is the effect of a strategy training on learning from process diagrams.

The general assumptions concerning all research questions with regard to comprehending process diagrams are based on multimedia theories (Mayer, 2001; Schnotz & Bannert, 2003), theories that focus on learning from graphical representations (Hegarty, 2005; Larkin & Simon, 1987; Winn, 1991), and cognitive load theory (Sweller, 1988; Sweller, 1994; Sweller, Van Merriënboer, & Paas, 1998). Multimedia theories claim that text and diagrams are processed in a task-oriented manner through verbal and visual systems in working memory to construct an integrated mental model. Construction of a mental model of an external representation involves interacting top-down (i.e., knowledge-driven) and bottom-up (i.e., perceptual and encoding) processes (Hegarty, 2005). The cognitive load theory (Sweller, 1988; Sweller, 1994; Sweller et al., 1998) is based on the assumption of limited working memory and unlimited long-term memory. When students must process materials containing many interacting elements than intrinsic cognitive load is high and working memory limits makes it difficult to assimilate.

The first research question (first and second study) focuses on the relationship between students' difficulties with process-diagram *problem* solving tasks and the features of the task, the student, and the diagram design. Problem solving from the perspective of the task is theoretically grounded on studies that categorize cognitive demand of task conditions (e.g., Guthrie, Shelly, & Kimmerley, 1993, Krathwohl, 2002). Studies that focused on student features and problem solving with diagrams identified prior knowledge (e.g., Chi, Feltovich, & Glaser, 1981; Winn, 1988) and cognitive abilities (e.g., Kozhevnikov, Motes, & Hegarty, 2007; Mayer & Sims, 1994; Winn, 1982) as key elements for success. For the perspective on diagrams, we

will focus on studies that identified design features that affect students' task performance (e.g., labeling, Mayer & Gallini, 1990; representational style of components, Winn & Sutherland, 1989; configuration and discrimination of components, Winn, 1991).

The second research question (third study) focuses on students' difficulties while *studying* process diagrams. We will focus on studies on cognitive and metacognitive learning activities employed by students while studying text and/or diagrams (e.g., Azevedo & Cromley, 2004; Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Meijer, Veenman, & Van Hout-Wolters, 2006; Pressley, 2000; Pressley & Afflerbach, 1995).

The theoretical focus when answering the third research question (fourth study) is on learning strategies and self-regulated learning (Boekaerts, 1997; Schunk & Zimmerman, 1997), and observational learning (Raedts, Rijlaarsdam, Van Waes, & Daems (2007).

3. STUDIES IN THIS THESIS

This thesis consists of six chapters and presents four studies. Chapters 2 to 5 each address these four studies. The first two studies (chapter 2 and 3) answer the first research question, the third study (chapter 4) answers the second research question and the fourth study (chapter 5) answers the third research question. Each chapter has been submitted as an article to an international journal*; three have been published (chapter 2, 3 and 4) and one (chapter 5) is under review.

In chapter 2, we present a study on explanatory factors that predict students' difficulties with process diagram problem-solving tasks. For this study, 64 process diagram problem-solving tasks from 18 Dutch Biology national exams were used as items. Hierarchical regression analysis was performed how features of the task, the student, and the diagram design related to the difficulty of that particular task, indicated by the cohort mean exam score.

In chapter 3, we focus more specifically on the role of students' characteristic in difficulties with process diagram problem-solving tasks. For this purpose, we developed a test that covers process diagrams and adjacent tasks used in secondary education Biology. With this test, the relationship between the cognitive demand of the problem-solving task on the one hand and student characteristics, i.e., prior knowledge, spatial ability and working memory, on the other, were examined.

Chapter 4 addresses the second research question. The chapter presents a study on students learning activities while studying process diagrams. Students were submitted to three learning tasks where they had to comprehend the full content of a process diagram. We collected eye-tracking data during the learning tasks and verbal data was collected by cued retrospective think-aloud protocol. The verbal and eye-

* Writing a thesis in articles has both advantages and disadvantages. Clear advantage is that each chapter can be read separately. For this, first and subsequent citations are also formatted per chapter. A disadvantage is that there is sometimes overlap between chapters and that the consistency in, for instance, names of conditions is not optimal.

tracking data were indices of learning activities. A two-level multilevel model was applied to examine which learning activities distinguish more and less successful learners and whether these learning activities are stable across tasks.

In chapter 5, the third research question is addressed. We will report on an intervention study that examined the effect of strategy training on learning from process diagrams. The training is based on cognitive strategy instruction (Harris & Graham, 1996). Results from first three studies (Chapters 2, 3, and 4) provided insights into the design of the intervention. The training focusses on a stepwise working-routine that includes when and where to employ cognitive and metacognitive learning strategies. The training also aims at motivating students to invest effort in the implementation of this stepwise working-routine. Structured equation modeling was applied to examine effects on learning from process diagrams.

Finally, in chapter 6 the results of the studies are discussed. Directions for future research and implications for educational practice complete this thesis.

CHAPTER 2

DIAGRAMMATIC LITERACY IN SECONDARY SCIENCE EDUCATION*

Students in secondary Science education seem to have difficulties with understanding diagrams. The present study focused on explanatory factors that predict students' difficulties with process diagrams, i.e., diagrams that describe a process consisting of components that are related by arrows. From 18 compulsory national Biology exams of secondary school pre-university students all process diagram tasks (n = 64) were included in corpus. Features of the task, student, and diagram were related to the difficulty of that particular task, indicated by the cohort mean exam score. A hierarchical regression analysis showed main effects for (1) the cognitive task demand, (2) the familiarity of the components, and (3) the number of components in a diagram. All these main effects were in the expected direction. We also observed interactions. Within the category of tasks with a high cognitive demand, tasks about a diagram of which students have low prior content knowledge were more difficult than tasks about a diagram of which students have high prior content knowledge. Tasks with a high cognitive demand about a diagram with familiar arrows were, surprisingly, more difficult than tasks with a high cognitive demand about a diagram with unfamiliar arrows. This latter finding might be attributed to compensation for task difficulty by the large number of components in the diagrams involved. The final model explained 46 percent of the variance in exam scores. These results suggest that students have difficulties (1) with tasks that require a deeper understanding when the content is new, (2) with diagrams that use unfamiliar component conventions, and (3) with diagrams that have a small number of components and are therefore probably more abstract.

1. INTRODUCTION

Diagrams are effective learning tools (Winn, 1991). They can help build a mental model and can make abstract ideas more concrete by triggering learners to use their spatial skills. Research shows that diagrams support learners' self-explanation (Ainsworth & Loizou, 2003), inference generation, and information integration and that they do reduce comprehension errors (Butcher, 2006; Cromley, Snyder-Hogan, & Luciw-Dubas, 2010). Larkin and Simon's (1987) comparison of sentence processing and diagram processing models shows how clustering and placement of components in diagrams make it easier to find information and to use it effectively. Once the first piece of relevant information has been found in a diagram, it is very likely that the next piece will also be found. However, many studies report that students have difficulties with diagram interpretation (Bowen & Roth, 2002; Chittle-

---

* Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2013). Diagrammatic literacy in secondary science education. *Research in Science Education, 43*, 1785–1800. doi:10.1007/s11165-012-9331-0

borough & Treagust, 2008; Guthrie, Shelly, & Kimmerly, 1993; Kindfield, 1993; Mathai & Ramadas, 2009; Schönborn, Andersen, & Grayson, 2002).

Interpreting diagrams is a key competence for students in mastering many of the biological theories. Bezemer and Kress (2008) showed that images are increasingly prominent as carriers of meaning in students' textbooks. In Biology course books, process diagrams explain processes like protein synthesis, immunology, photosynthesis, cellular respiration, compound cycles, etc. (Bowen & Roth, 2002; Campbell & Reece, 2002). In the present study we will study these biological process diagrams, defined as dynamic diagrams that describe a biological process consisting of components that are related by arrows.

Students are often not explicitly trained in how to interpret diagrams. They should be made aware of the conventions of diagrammatic representations and the scope and limitations of each diagrammatic mode (Gilbert, 2005; Winn, 1991). Ainsworth (2006) states that learners who are first presented with a novel representation must understand how it encodes information and how it relates to the domain it represent. It makes sense to provide opportunities for students to actively investigate diagrams and to build up adequate strategies for interpretation. Designers of such a program cannot rely on information on the difficulties students have in tasks that require an interpretation of biological process diagrams, as a thorough investigation of students' difficulties on the interpretation of these specific types of diagrams has not been performed yet.

The present, exploratory, study fills this gap in knowledge, examining three factors (i.e., the task, student, and diagram), and the extent in which these explained variances in difficulty, indicated by the cohort mean exam scores of students on diagram tasks.

2. THEORETHICAL FRAMEWORK

In the next three sections we will present research on features of the task, student, and diagram, and their effect on learning and performance.

2.1 Task features

A classic way of classifying the level of cognitive demand of a task is by using Bloom's Taxonomy (Bloom, Krathwohl, & Masia, 1956). This taxonomy distinguished six categories: knowledge, comprehension, application, analysis, synthesis and evaluation. Krathwohl (2002) revised the original taxonomy into a two-dimensional framework, with the dimensions knowledge and cognitive processes. Crowe, Dirks, and Wenderoth (2008) developed a 'Blooming Biology Tool' that can be used to assess the level of difficulty on tasks for several domain specific skills.

Guthrie et al. (1993) investigated the cognitive processing of several displays, e.g. graphs, pictures and static diagrams, under two different task conditions, i.e., local and global searches. For local searches only a part of the display needs to be explored and the cognitive demand is specific and precise: Local searches demand

discrimination between categories. Global searches involve the interpretation of the entire display and found to be more difficult than local searches. The latter effect is in line with the cognitive load theory (Sweller, 1994). Winn (1993) concludes that different search goals lead to different search patterns while readers look at diagrams. Goals set search strategies and effective search strategies requires knowledge of content and of the conventions of diagrams.

We will incorporate the findings of the presented studies about task features into the present study's definition of cognitive task demand.

2.2 Student features

A student's familiarity with the design of a diagram might be an important factor for learning and understanding (Carlson, Chandler, & Sweller, 2003). In this context, familiarity refers to the likeliness that students have a cognitive schema available that supports the interpretation of the diagram. Previous research suggests that familiarity with the spatial arrangement of the components and with the component type of a diagram has a positive effect on performance on learning tasks. Winn (1982) reversed the order of an evolutionary dinosaur diagram which decreased performance. In another study (Winn, 1988), participants performed worse on recall tasks when components were presented as unfamiliar icons rather than as labeled boxes. Winn, Li, and Schill (1991) found that participants who are familiar with the terms and conventions of family tree diagrams solved kinship problems faster. They conclude that familiar diagrams do not require the computation of solutions from rules and that diagrams should be used with students who already possess some knowledge of the terms and conventions.

Winn and Sutherland (1989) examined the effects of varying the familiarity and number of elements in maps and diagrams on the recall of lists of elements and their location. They found no main effect for familiarity, but high-ability participants used more helpful strategies like chunking and story construction than low-ability learners when presented an unfamiliar diagram. They conclude that there is some evidence that differences in performance are related with students selection of helpful or non-helpful strategies.

Prior knowledge of the content and the processes presented in the diagram might also affect student performance (e.g., Cook, Wiebe, & Carter, 2008) . Hegarty and Just (1989) focused on similar graphical displays as Mayer and Gallini (1990) and found that individual differences in prior knowledge affected diagram comprehension. Compared to participants with a low prior knowledge, participants with a high prior knowledge were more capable of locating the relevant information in a diagram and extracting information more selectively. Winn (1993) argued that prior knowledge facilitates analysis of a diagram in two ways. First, knowledge schemata inform readers what to look for and where the information is most likely to be found. Secondly, prior knowledge provides a structure in which information found can be organized. Lowe (1996) found that the difference between interpretation processes, i.e., the search for information and the interpretation and construction of

weather maps, was affected by the level of prior content knowledge. An eye fixation study (Canham & Hegarty, 2010) complements these findings by concluding that information selection improves after instruction about relevant meteorological principles.

2.3 Diagram features

A biological process diagram consists of components that are related by arrows: Arrows can express many relations, such as pointing or connecting, sequence, change over time, path, or manner of movement or forces (Heiser & Tversky, 2006). The function of an arrow in a process diagram is usually conveyed by a label.

An early study by Holliday, Brunner, and Donais (1977) compared the effectiveness of two types of process diagrams that illustrated water, oxygen, nitrogen, and carbon dioxide cycles. In one diagram type, the components were text boxes with labels; in the other, they were iconic drawings with labels. In a number of retention and comprehension tasks the iconic drawings were found to lead to a significantly better performance, but that this effect was limited to low-ability students. Winn and Sutherland (1989) found that low-ability students remembered lists and locations of labeled elements better when they were shown as drawings rather than as squares. These results are supported by the dual coding theory of Paivio (1990) which suggests that the presentation of two modes of information, i.e., verbal and visual, promotes effective use of working memory.

The number of components and arrows in process diagrams can vary extensively. Cognitive load theory (Sweller, 1994) predicts a high intrinsic cognitive load if interaction between many elements, i.e., components, must be learned. Limited working memory makes schema acquisition, which is a major learning mechanism, difficult when multiple elements interact. Winn (1988) found that the number of components in a diagram had a significant effect on the recall of names and positions.

In their landmark study on diagram design principles, Mayer and Gallini (1990) found that students, mainly with low prior-knowledge levels, performed better on conceptual recall and problem solving questions when the illustration contained labels that describe the parts and steps of how mechanical devices such as a pump and a gear system works. Mayer and Moreno (1998) argue that when text and diagrams are separated, learners have to read some portion of the text and then maintain it in their working memory while reading the diagram. This places higher cognitive demand on working memory and increases the possibility that, because of working memory limitations, some information will be lost.

Winn (1991) presents a framework for learning from maps and diagrams which focuses on pre-attentive perceptual organization. In this framework, configuration and discrimination of components are fundamental when perceiving diagrams. The meaning of a diagram cannot be interpreted before the components have been configured and discriminated.

Configuration is the spatial relationship among components in a diagram. It determines which components appear to form clusters, the sequence in which components are processed, and which components receive the most attention. Clustering of components has been found to be effective for memorization tasks. In Biology, clustering often distinguishes certain components as functional groups or it is used to depict a specific location of a sub-process. Discrimination is the ease with which a component can be distinguished from another. In diagrams, discrimination can be supported by using many characteristics like color, shape, and size.

In the current study, the relationship was examined between features of the task, student, and diagram on the one hand and difficulty of a diagram task on the other hand. We formulated the following hypotheses:

1) Task features: Cognitive task demand does explain differences in task difficulty.
2) Student features: Diagram familiarity does explain differences in task difficulty, and prior content knowledge does explain differences in task difficulty.
3) Diagram features: The presented studies on diagram features usually focus on the effect of task achievement when a single element in a diagram is changed, e.g. label or no labels (Mayer & Gallini, 1990); text boxes or iconic drawing (Holliday, Brunner, & Donais, 1977). The present study included a large number of diagrams that can vary on many elements from each other. This large variability makes it difficult to predict the effect of any specific element or combinations of elements on task difficulty so no hypotheses were formulated for diagram features.

We also expected that difficulty of a diagram task can be explained by the interaction of cognitive task demand on the one hand and student features and diagram features on the other. It seems reasonable to expect that differences in diagram features have a larger effect on difficulty when the task is more cognitive demanding than when the task is less cognitive demanding (e.g., Mayer & Gallini, 1990; Carlson et al., 2003). We also find it reasonable to expect that the effect of cognitive demand is larger when prior knowledge is relatively low than when prior knowledge is relatively high (e.g., Canham & Hegarty, 2010), and that the effect of cognitive demand is larger when familiarity is relatively low than when familiarity is relatively high (e.g., Winn et al., 1991). In other words, when the task is more cognitive demanding, the effect of student and diagram variability is more pronounced.

3. METHOD

3.1 Data collection

Data from the Dutch compulsory national exams were collected by Cito – National Institute for Educational Measurement and can be obtained for research purposes. For the present study, the student's scores per task on 18 Dutch Biology national exams (Table 2.1) from the period 2001-2009 were obtained, i.e., there were two exams each year, a regular and a retake exam. The participants of the obtained Dutch Biology national exams were secondary school pre-university students with an average age of 18 years old.

The nine cohorts, i.e. 2001-2009, can be assumed to be similar with respect to student composition. For the last three years of their study in secondary education, the students chose Biology as one of their major topics within their exam program for which they received 480 hours of training. The biology curriculum in the Netherlands is well described on a national level. All students in the Netherlands follow a similar curriculum with respect to the content, although teaching approaches may and will vary of course.

Table 2.1. Summary of the national Dutch biology exams and data obtained from the Cito

Year	Exam	Tasks in exam	N/M size	Diagram tasks selected
2001	1	40	728/2145[a]	2
2001	2	45	27/27[a]	2
2002	1	40	1784/10865	7
2002	2	44	180/180	2
2003	1	38	1846/11073	5
2003	2	41	198/198	4
2004	1	41	1814/11893	4
2004	2	40	282/282	0
2005	1	39	1940/12444	8
2005	2	37	258/258	4
2006	1	36	6264/12804	6
2006	2	41	418/418	3
2007	1	35	7153/13798	6
2007	2	38	392/392	1
2008	1	37	8070/15288	3
2008	2	37	442/442	2
2009	1	35	10624/17539	3
2009	2	40	471/471	2
	Total	704	42891/110517	64

Note. N = sample size; M = population size; 1 = regular exam; 2 = retake exam.
[a]These samples are relatively small compared with the other regular and retake exams in the dataset because this was a transition year to the new exam program

For this research, student's scores per task from the regular and the retake Dutch Biology national exams 2001-2009 were used as data. The data obtained from the regular exams were large simple random samples (Hoyle, Harris, & Judd, 2002) of the scores per student per question; the data from the retake exam contained the entire population. The exams included a total number of 704 tasks of which 64 tasks required direct interpretation of a process diagram. All 64 tasks, selected by the first author (i.e. a part-time high school Biology teacher with 10 years' experience), were included in the present study.

Each task focused on a specific problem but the answer could require several steps: Each correct step was rewarded with a point. In the exams the information of the maximum score a student could achieve was presented to them in the left margin of the task: The maximum score of each of the 64 included tasks varied from 1 till 4. From the student scores on these 64 tasks we calculated the cohorts' mean score for

each task, linearly transformed to a range from 0-1. This score was used as the dependent variable as we understood it as indicating diagrammatic task difficulty, i.e., a low cohort mean score indicates a high diagrammatic task difficulty.

3.2 *Explanatory variables*

The explanatory variables are presented in Table 2.2. *Exam* was included in the analysis as a covariate as we wanted to control for the possible differences between the students in the regular exam and the students who participate in the retake exam. Students who register for the Biology retake exam do that for various reasons. They can have failed their regular Biology exam, but also can want to increase their grade for Biology to compensate for a low grade on another subject (e.g., Chemistry or French), or they need an excellent grade to be selected for fixed-number studies in higher education like Medicine.

3.3 *Task features*

The explanatory variable *Cognitive task demand* was based on the studies from Guthrie et al. (1993), Bloom et al. (1956), and Crowe et al. (2008). We defined two levels of *Cognitive task demand*: tasks with a ‘low’ and tasks with a ‘high’ cognitive demand.

We defined a ‘low’ *Cognitive task demand* when only a few elements needed to be explored. Once the relevant information was found in the diagram, finding a correct answer asked for little cognitive processing. This meant that this type of task only required information recall or understanding of a concept or terms. Examples of this task type in the dataset were summarizing the elements found, labeling elements, describing step-by-step a part of the process and/or some simple arithmetic like adding or subtracting amounts, such as the amount of energy transported to a compartment in an ecosystem.

A ‘high’ *Cognitive task demand* was usually a more global task; a larger part or the entire diagram needed to be explored. Once the selected information was found it must be processed in working memory and integrated (evaluated, inferred, compared, judged) with prior knowledge. The task might require a prediction of the most likely outcome given a new situation of the process or an interpretation of the data and a selection of the best conclusion. An example of a task with a ‘high’ cognitive demand from the regular Biology exam 2002 was: “Explain—solely on the basis of the diagram—that an increase of dead material can lead to an increase of the number of lemmings”. This category also included ‘productive’ tasks, i.e., tasks where students had to add arrows or components to an existing diagram or create a diagram from a set of given components.

The tasks were coded for Cognitive task demand by two expert and experienced Biology teachers (Cohen’s $\kappa = .87$, with a 95% confidence interval of $.63 < \kappa < 1.0$) and agreement was found on all items afterwards.

Table 2.2. The number of tasks (N) and task difficulty (cohort mean score and standard deviation) for the explanatory variables

Explanatory variable	*N*	Task difficulty	
		M	*SD*
Exam			
Regular[a]	44	.57	.17
Retake	20	.44	.26
Task features			
Cognitive task demand			
low[a]	44	.60	.19
high	20	.36	.17
Student features			
Familiarity diagram components			
familiar[a]	47	.56	.22
unfamiliar	17	.45	.17
Familiarity diagram arrows			
familiar[a]	58	.53	.22
unfamiliar	6	.48	.13
Familiarity spatial arrangement			
familiar[a]	52	.55	.22
unfamiliar	12	.43	.15
Prior content knowledge			
high[a]	47	.55	.19
low	17	.48	.27
Diagram features			
Arrow type			
movement	21	.48	.18
step[a]	18	.60	.20
transfer	18	.49	.26
feedback	7	.56	.18
Component type			
iconic[a]	17	.58	.19
symbolic	27	.49	.23
text boxes	20	.54	.20
Clustered			
yes[a]	21	.52	.20
no	43	.53	.22
Process labels			
labeled[a]	24	.54	.20
unlabeled	40	.52	.22
Sequence			
linear[a]	30	.53	.22
cyclic	34	.52	.21
Components		9.27[b]	5.61
Arrows		11.00[b]	7.07

Note. [a]*Reference group.* [b]*This is the mean number of components and arrows.*

3.4 Student features

Familiarity of diagrams referred to the familiarity with components, arrows and spatial arrangement of a diagram. In the national Dutch Biology exams, diagrams might have specific conventions for components, arrows or spatial arrangement which a student is not familiar with. In such a case a student has to invest in understanding these conventions to perform on a task. For example, components might be combined in an unusual way or arrows might have a different meaning than in a conventional diagram about the same topic. There are also many examples of diagrams that were cross topic, e.g. nervous system and blood circularly system combined, or that used multiple conventions for components and arrows in one diagram.

The spatial arrangement of components in a diagram can be unfamiliar in various ways: the orientation might be right to left instead of left to right, bottom to top instead of top to bottom, linear instead of circular (or vice versa), etc. It could also be that components are clustered in an unusual manner. The familiarity of the components, arrows and spatial arrangement of a diagram were coded by the same two expert and experienced Biology teachers. Interrater reliability (Cohen's κ) for the familiarity of components, arrows and spatial arrangement were .86, with a 95% confidence interval of $.60 < \kappa < 1.0$, .82, $.48 < \kappa < 1.0$, and .87, $.63 < \kappa < 1.0$, respectively.

Prior knowledge of the content was defined to be 'high' if the information presented in the diagram was part of the curriculum. Although the Dutch Biology curriculum for pre-university education is well defined, it might, however, be that some topics that were not in the curriculum were considered to be part of the prior knowledge of the students. This variable was coded by the two teachers as well (Cohen's $\kappa = .86$, with a 95% confidence interval of $.60 < \kappa < 1.0$).

3.5 Diagram features

Arrows in a biological process diagram might have different meanings depending on the context, such as 'transport', 'causality', 'transformation', or 'feedback' interaction (Figure 2.1). Arrows that focus on 'transport' represent the flow of energy or matter from one component to another component. The arrows can be labeled with the amount of energy or matter that is transported per time or with the represented processes. Arrows that represent 'causality' describe step by step processes; a process in one component has an effect on the next component. The arrows represent the process and could be labeled with the name of the process. A sense of movement could be created by clustering, e.g. from the nucleus to the cytoplasm (see the 'causality' panel in Figure 2.1).

The 'transformation' arrow type describes the transformation of the compartment itself, e.g. a chemical compound. The arrows represent the chemical reactions and were sometimes labeled with the process name (e.g. hydrolysis) or the responsible enzyme (e.g. amylase); sometimes the arrows also depicted movement from one location to another by using clustering.

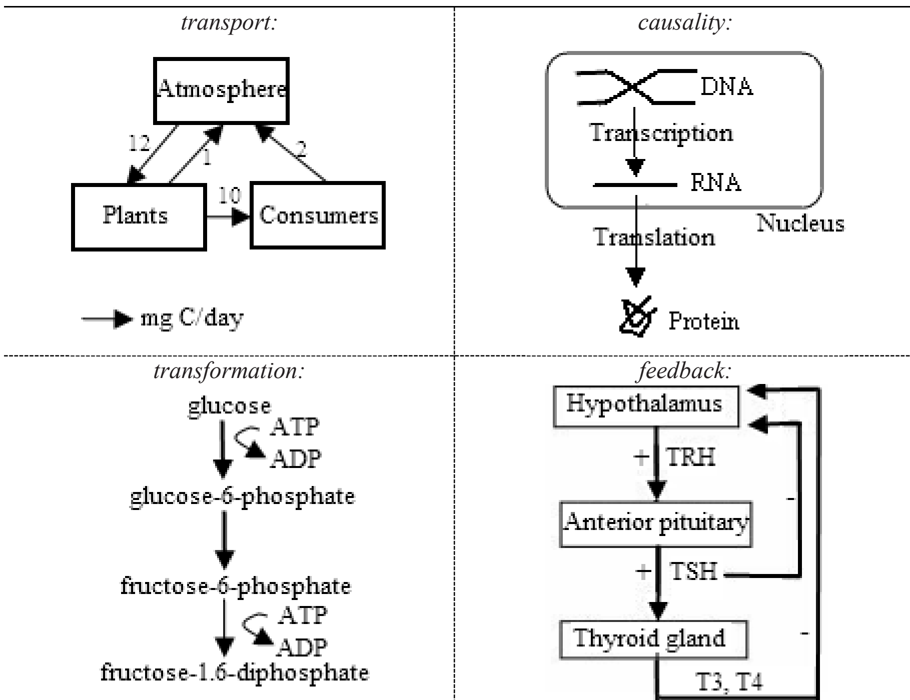


Figure 2.1. Arrow types and exemplary diagrams. The presented diagrams are exemplary and where not part of the actual dataset. Some examples of topics in the actual dataset with these arrow types are: transport: food webs, diffusion; causality: nervous system, immune system; transformation: assimilation/dissimilation; feedback: homeostasis, climate.

The 'feedback' arrow type was typically used in diagrams explaining hormonal processes. Compartments of hormone diagrams represented the glands or sometimes the hormone itself and the arrows represent production of the hormone and the type of feedback, i.e. inhibition or stimulation, of the hormone on the production of itself or other hormones. The arrows were often labeled by a plus or minus sign representing the direction of the feedback. Feedback arrows are also common in climate models, e.g. the effect of the carbon dioxide emission on sea level rise or photosynthesis.

Finally, the components of a diagram were classified as 'text boxes', 'symbolic' or 'iconic'. If some components were spatially separated from others to form a functional group or to depict a location then we considered clustering to be present in the diagram. Diagrams with arrows that were labeled directly or in indirectly by using a legend, e.g. a letter/number on the arrow corresponding to a legend item, were categorized as labeled. If a diagram could only be read in a single direction then the diagram was categorized as linear instead of cyclic. Diagram features were very distinct and were therefore only coded by the first author.

3.6 Data analysis

A hierarchical multiple regression analysis was performed with task difficulty as the dependent variable, features of the task, student, and diagram as explanatory variables and exam (regular vs. retake) as covariate. The categorical variables were converted using dummy coding; the reference groups (Table 2.2) were selected by several criteria. First, the group had to be well defined, and preferably be at the upper or lower boundary and contain a sufficient number of cases (Hardy, 1993). We also pre-analyzed if the selection of the reference group would leave out any interesting significant main effects and interaction effects in the regression model. These criteria led to a selection of reference groups that assured that no significant effects missed in the final regression model. The continuous variables, i.e. *Compartments* and *Arrows*, were centered around their means to reduce problems with multicollinearity (Frazier, Tix, & Barron, 2004).

The interaction terms were created by multiplying the explanatory variables. The variables entered the analysis in two steps. The first step included the covariate *Exam* and the main effect *Cognitive task demand* in the first block and all other main effects in the second block of the model (stepwise method). Only the significant main effects were kept in the regression model. The interaction between the *Cognitive task demand* and the other explanatory variables was first inspected visually by plotting the simple slopes of each combination. In the second step the most promising interaction terms and the accompanying main effects were added to the model one by one using the enter method, keeping only the significant interactions effects and their accompanying main effects in the final model. Significant interaction effects were analyzed post hoc by a *t*-test for simple slopes as suggested by Aiken and West (1991).

4. RESULTS

The results of the hierarchical multiple regression analysis with task difficulty as the dependent variable are summarized in Table 2.3. Model 1 includes the main effect from the variable *Cognitive task demand*. This variable explains 30% of the variance, $F(1, 62) = 14.18$, $p < .001$. In the second block of the analysis the significant main effects, *Familiarity diagram components* and *Components*, and the accompanying main effects of the significant interaction effects, *Familiarity diagram arrows* and *Prior content knowledge*, were added to the model. Model 2 explains 37% of the variance.

The significant interaction effect, *Cognitive task demand x Prior content knowledge*, was added to the third block of the analysis. Model 3 explains 43% of the variance, $F(6, 57) = 7.68$, $p < .001$. In the fourth block of the regression analysis the interaction effect *Cognitive task demand x Familiarity diagram arrows* was added to the final model which now explains 46% of the variance, $F(7, 56) = 7.59$, $p < 0.001$.

Table 2.3. Model summary of the multiple regression analysis with task difficulty (cohort mean score) as the dependent variable

Model	Variable(s) included	Adjusted R^2	F	p
1	Cognitive task demand	.295	14.184	< .001
2	Familiarity diagram components, Components, Familiarity diagram arrows, Prior content knowledge	.373	7.234	< .001
3	Cognitive task demand x Prior content knowledge	.426	7.682	< .001
4	Cognitive task demand x Familiarity diagram arrows	.456	7.589	< .001

The final regression model is presented in Table 2.4. The covariate *Exam* was not a significant predictor ($t = 1.39$, $p = .17$). The variable *Cognitive task demand*, with 'low' as the reference group, was found to be a significant predictor ($B = -.175$, $p = .003$) of task difficulty. This means that tasks with a 'high' cognitive demand showed lower mean exam scores and can be considered to be more difficult, compared to tasks with a 'low' cognitive demand.

Table 2.4. Final model of hierarchical regression analysis for variables predicting task difficulty (cohort mean score) on diagrammatic tasks (N = 64)

Variable	B	$SE\ B$	t	p
(Constant)	.578	.044	13.042	.000
Cognitive task demand	-.175	.056	-3.115	.003
Exam	.063	.045	1.391	.170
Familiarity diagram components	-.140	.052	-2.684	.010
Components	.009	.004	2.104	.040
Familiarity diagram arrows	-.063	.103	-.612	.543
Prior content knowledge	.094	.062	1.499	.140
Cognitive task demand x Prior content knowledge	-.248	.095	-2.617	.011
Cognitive task demand x Familiarity diagram arrows	.291	.145	2.008	.050

Note. Components and Arrows are centered at their means. Only unstandardized B is reported because β is not properly standardized in equations that include interaction terms and are thus not interpretable (Frazier et al., 2004).

The explanatory variables *Familiarity diagram components* ($B = -.140$, $p = .010$) and *Components* ($B = .009$, $p = .040$) also were significant predictors. Note that 'unfamiliar' is the reference group for the variable *Familiarity diagram components* and that *Components* is centered to its mean score. This means that tasks about diagrams with 'unfamiliar' components showed lower mean exam scores and can be considered to be more difficult, compared to tasks about diagrams with 'familiar' components. The positive effect for the number of *Components* showed that the more *Components* a diagram had, the higher the mean exam scores. This indicates that the task was less difficult if there were more *Components*.

The upper panel of Figure 2.2 presents the simple slopes for the significant interaction effect *Cognitive task demand x Prior content knowledge*, (B = -.248, p = .011). Post hoc probing of the interaction by a *t*-test for simple slopes shows that the interaction between a 'high' *Cognitive task demand* and *Prior content knowledge* is significant, $t(16) = 2.14$, $p < .05$. The interaction between a 'low' *Cognitive task demand* and *Prior content knowledge* is not significant, $t(40) = 1.32$, $p > .05$. This means that the combination of a task with a 'high' cognitive demand and 'low' *Prior content knowledge* showed a lower mean exam score, indicating that the task was more difficult than a combination of a task with a 'high' cognitive demand and 'high' *Prior content knowledge*.

The interaction effect *Cognitive task demand x Familiarity diagram arrows*, (B = .291, p = .050) was added to the regression model in the fourth step. Figure 2.2 (lower panel) presents the simple slopes for this interaction effect. Post hoc probing of the interaction by a *t*-test for simple slopes shows that the interaction between a 'high' *Cognitive task demand* and *Familiarity diagram arrows* is significant, $t(16)$ = 2.15, $p < .05$. This means that the combination of a 'high' *Cognitive task demand* and an 'unfamiliar' arrow type showed higher mean exam scores than the combination of a 'high' *Cognitive task demand* and a 'familiar' arrow type: This indicates, surprisingly, that tasks with a 'high' cognitive demand were less difficult with an 'unfamiliar' arrow type than with a 'familiar' arrow type. The interaction between tasks with a 'low' cognitive demand and *Familiarity diagram arrows* is not significant, $t(40) = .69$, $p > .05$.

5. DISCUSSION

The aim of this study was to examine variables that could explain the difficulty of a diagrammatic task. In this study, the cohorts' mean exam score on a diagrammatic task was used as an indicator of task difficulty. It was hypothesized that features of the task and student explained differences in task difficulty; no hypothesis was formulated for diagram features. We also hypothesized that difficulty of a diagram task can be explained by the interaction of cognitive task demand on the one hand and student features and diagram features on the other. In the result section, we reported significant main and interaction effects for features of the task, student, and diagram.

The *Cognitive task demand* explained most of the variance in task difficulty. The familiarity of the components was also related to task difficulty. Diagrams that contain 'familiar' components showed higher mean exam scores and were therefore seen as less difficult. We found no interaction effect between *Cognitive task demand* and *Familiarity diagram components*. This means that the effect of 'familiar' components on task difficulty is not different for tasks with a 'low' and a 'high' cognitive demand. This result seems to be consistent with research by Winn and Sutherland (1989) and Winn et al. (1991) who also focused on familiarity of components. Their experiments were performed in a controlled setting in which students got a limited amount of time to complete the tasks, with a limited variation in diagram presentation.

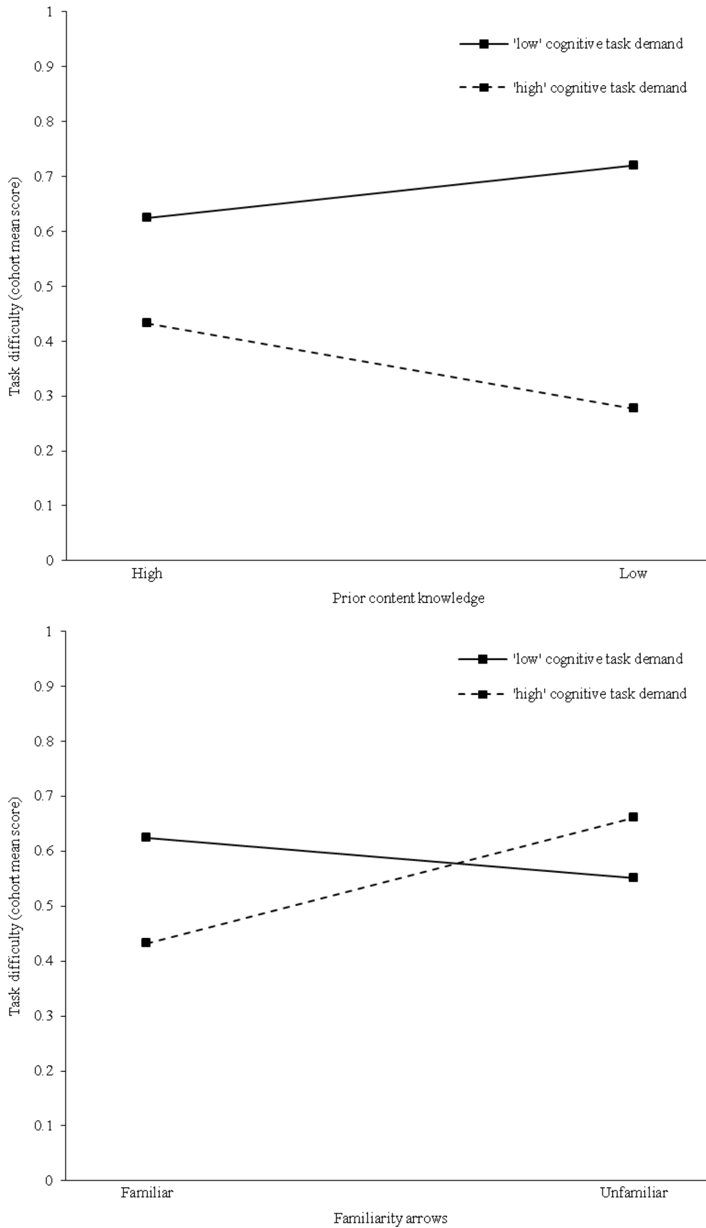


Figure 2.2. Plot of the significant interaction effect between Cognitive task demand and Prior content knowledge (top panel) and Cognitive task demand and Familiarity diagram arrows. The presented lines are the simple slopes and were calculated as suggested by Aiken and West (1991).

Moreover, they used memorization and response latencies as indicators for performance. The present study however demonstrated that in an ecological valid setting familiarity is also an important design issue for various diagram tasks. This conclusion is in line with the findings of Carlson et al. (2003) who suggest that diagrams, from a cognitive load perspective, facilitate understanding when they activate familiar schemas of the learner.

The number of *Components* was negatively related to the task difficulty; the more *Components* in a diagram the less difficult a task was. This suggests that adding more *Components* to a diagram makes it easier to interpret. This might be because the information in a diagram becomes more concrete if more *Components*, i.e., bits of information, are added: The diagram becomes less ambiguous and comprehension errors are less likely.

There was no interaction effect of *Cognitive task demand* and the number of *Components*, suggesting that this effect was not different for tasks with a 'low' and a 'high' cognitive demand. From a cognitive load (Sweller, 1994) point of view it might be expected that an increase in the number of *Components* might be more difficult with a 'high' *Cognitive task demand*, i.e., when the entire diagram needs to be examined. Not observing this effect in this corpus might be due to the fact that most diagrams are 'familiar' and that no real new information had to be learned. Note that this corpus is taken from central compulsory exams, in testing practice it is unusual to include learning of new items. There might be a third-order effect between *Cognitive task demand*, number of *Components* and *Prior content knowledge* but there was not enough data available to confirm this.

Post hoc analysis of the interaction between *Cognitive task demand* and *Prior content knowledge* revealed a significant difference in task difficulty for tasks with a 'high' cognitive demand and *Prior content knowledge* levels. There was no main effect for *Prior content knowledge* and there was no interaction between *Prior content knowledge* and a 'low' *Cognitive task demand*. This latter result is somewhat different as the findings by Lowe (1992) who reported that meteorologists selected more relevant information on weather maps than non-meteorologists. The introduction of a diagram about a new biological topic to students with a fair amount of biological background is probably not similar to presenting abstract and domain specific weather maps to non-meteorologists. It is suggested here that at least some of the students' knowledge about conventions in process diagrams which facilitate an analysis of the diagram might be easily transferred to biological diagrams about unknown topics.

The interaction effect between tasks with a 'high' cognitive demand and *Prior content knowledge* seems to confirm the conclusions of Winn (1993) that the absence of well-constructed knowledge schemata makes interpretation of a diagram more difficult. This finding raises some concerns for learning new scientific topics by diagram representation, especially when considering that images become increasingly important as carriers of meaning (Bezemer & Kress, 2008).

Post hoc analysis of the interaction between *Cognitive task demand* and *Familiarity diagram arrows* revealed a positive effect on task difficulty for tasks with a

'high' cognitive demand when the diagram has an 'unfamiliar' arrow type. This effect might seem a bit peculiar at first glance. But a closer look showed that two of the three diagrams with an 'unfamiliar' arrow type had many more *Components* (i.e., 23 and 24 components) than average (M = 9.27, SD = 5.61). This increase in concreteness by these number of *Components* might compensate for the 'unfamiliarity' of the arrows. No other main effects or interaction effects (i.e., besides the number of *Components*) where found with the explanatory variables which described the diagrams' features. This obviously does not imply that previous research on the effect of diagram features like *Component type* (Holliday et al., 1977), *Clustering* (Winn, 1991), *Process labels* (Mayer & Gallini, 1990), etc., is disapproved. It merely suggests that the heterogeneous composition of the diagrams in our dataset probably makes it more difficult to find any significant effect as many factors might interact. It must also be mentioned that many of the other studies were about diagrams which might were not similar to biological process diagrams, e.g. a bike pump and a car brake (Mayer & Gallini, 1990), a family tree diagram (Winn et al., 1991), or a weather map (Lowe, 1996; Canham & Hegarty, 2010).

The results from the present study provide relevant insights into the design of a training of students in the use of process diagrams in Biology. We conclude that a training program should: (1) include strategies for encoding diagrams with unfamiliar components, (2) focus on the interpretation of abstract (i.e., highly conceptual with a minimal amount of context information) diagrams, and (3) facilitate students in learning how to gain a deeper understanding of diagrams that contain new information. We suggest that a training on meta-cognition which involves self-explaining (Bielaczyc, Pirolli, & Brown, 1995) and thinking about diagrams on a meta level, e.g., what conventions are used, how diagrams can be constructed, what information (the designer of) the diagram is trying to communicate, etc., might be an effective approach. Another approach would be to show students several videos of novices who interpret diagrams successfully or unsuccessfully as a form of modelling. The students observe and have to compare and evaluate task behaviour; this kind of training has been done earlier within various domains, e.g., written composition (Raedts, Rijlaarsdam, Van Waes, & Daems, 2007), and has proven to be effective.

Finally, this study adds to previous research on the role of task demand (e.g., Guthrie et al., 1993), prior knowledge (e.g., Lowe, 1996; Cook et al., 2008), familiarity with conventions (e.g., Carlson et al., 2003), and design features (e.g., Mayer & Gallini, 1993; Canham & Hegarty, 2010) in learning from external representations. The present study reveals what factors contribute to the difficulty of the tasks, but not how these factors interact during processes of understanding. We still need a more thorough understanding of which problems are urgent and relevant when improving students diagram interpretation skills. Foci of further research on the interpretation of process diagram could be, for example, differences in students' search efficiency and solving strategies and the role of spatial ability.

CHAPTER 3

STUDENTS' ABILITY TO SOLVE PROCESS-DIAGRAM PROBLEMS IN SECONDARY BIOLOGY EDUCATION*

Process diagrams are important tools in Biology for explaining processes like protein synthesis, compound cycles, etc. The aim of the present study was to measure the ability to solve process-diagram problems in Biology and the relationship with prior knowledge, spatial ability and working memory. For this purpose, we developed a test that represents process diagrams and adjacent tasks used in secondary education Biology. Results show that the ability to solve process-diagram problems is correlated to prior knowledge, spatial abilities, and visuospatial working memory capacity. A difference in impact of spatial skills was demonstrated for the level of cognitive demand when solving process-diagram problems.

1. INTRODUCTION

Diagrams are important tools in science education. They allow us to communicate abstract information. Diagrams explain natural phenomena that cannot be directly observed: too small, too large, too slow, or too fast. Process diagrams form a distinct class of diagrams: they convey functional information about a dynamic process by the spatial configuration of components and arrows. In Biology, process diagrams explain processes like protein synthesis, immunology, photosynthesis, cellular respiration, compound cycles, etc. (e.g., Reece et al., 2010). In biology education, students are faced with process-diagram problems that require them to select and extract, to interpret and to infer the presented information.

Although diagrams aim to facilitate learning (Larkin & Simon, 1987; Winn, 1993), students have difficulties with diagram interpretation (e.g., Schönborn, Anderson, & Grayson, 2002). Previous studies found that prior knowledge (e.g., Cook, 2006), working memory and spatial skills (e.g., Hegarty & Sims, 1994), and task demand (e.g., Guthrie, Shelly, & Kimmerly, 1993) contribute to the interpretation process of scientific representations. The present study was focused on providing more insight into students' ability to solve process-diagram problems in Biology in secondary education.

---

* Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2015). Students' ability to solve process-diagram problems in secondary biology education. *Journal of Biological Education, 49*, 1–13. doi: 10.1080/00219266.2014.888363

2. THEORETICAL FRAMEWORK

Two frameworks are relevant for problem solving with diagrams: (1) the working memory model of text and picture comprehension of Schnotz and Bannert (2003) and (2) the cognitive load theory (Sweller, 1994).

2.1 Working memory

In Schnotz and Bannert's model (2003), text and diagrams are processed through verbal and visual systems in working memory to construct an integrated mental model. Prior knowledge has a selective and organizational function. Students with little prior knowledge have more difficulties in creating effective mental models (Mayer & Moreno, 2003). The construction of a mental model draws on cognitive resources of the visuospatial sketchpad (Sims & Hegarty 1997). Students with a high spatial ability can devote more resources on building referential connections between the visual and verbal mental model than low spatial ability learners (Mayer & Sims, 1994).

2.2 Cognitive load theory

The cognitive load theory (Sweller, 1994) assumes a limited working memory storage capacity and unlimited long-term memory storage capacity. Intrinsic cognitive load is high when materials include many interacting elements. Working memory limits then make it difficult to assimilate the presented information. In such a case, long-term memory expands the processing abilities of working memory by the storage of information into schemas, i.e., cognitive constructs that incorporate multiple elements of information into a single element. When knowledge schemas are available they can be brought to working memory as chunks and thereby reduce cognitive load.

2.3 Problem solving

Kindfield (1993) concluded that the use of representations in reasoning and problem solving co-evolves with domain expertise. Experts possess schemas that contain declarative and procedural knowledge that is used for *problem solving* processes (Chi, Feltovich, & Glaser 1981). Larkin, McDermott, Simon, and Simon (1980) found that the availability of schemata's facilitated efficient search in diagrams. It also guides the interpretation of a problem and the formulation of a solution (Chi et al., 1981).

2.4 Prior knowledge

Cook (2006) showed that prior knowledge is one of the most determining factors for success in learning from representations: Domain knowledge affects information selection, encoding, interpretation, and inferencing from diagrams.

Prior knowledge is important for *selecting* task-relevant information in a diagram. Novices focus on surface features of a domain-specific diagram, whereas experts attend to more relevant content (Canham & Hegarty, 2010; Chi et al., 1981; Cook, Carter, & Wiebe 2008). For instance, Cook and colleagues (2008) compared the interpretation process of students with low and with high prior knowledge of cell transport diagrams, i.e., diffusion and osmosis. Low-prior knowledge students focused less on relevant features, e.g., a concentration gradient or an active transport zone, when these features were not specifically emphasized in the diagram. When task-relevant information is found it must be further *encoded* to construct an integrated mental model (Schnotz & Bannert, 2003).

Prior knowledge also affects *interpretation* and *inference* processes *after* the presented information is encoded. Kragten, Admiraal, and Rijlaarsdam (2013a) found that absence of domain knowledge impaired the interpretation of process diagrams when cognitive task demand was high, but not when cognitive task demand was low.

2.5 Spatial ability and working memory

Spatial ability and working memory relate to students' problem-solving ability in scientific diagrams (e.g., Bodner & McMillen, 1986), especially when it requires spatial transformation processes (Hegarty & Sims, 1994), visualisation (Kozhevnikov, Hegarty, & Mayer, 2002), and mental model construction (e.g., Mayer & Sims, 1994).

Various studies report that spatial ability and Chemistry problem solving relate (see Wu & Shah, 2003, for an extensive review about this issue), both in spatial and in non-spatial higher-order cognitive tasks (e.g., Bodner & McMillen, 1986; Pribyl & Bodner, 1987). Wu and Shah (2003) conclude that understanding both types of tasks, spatial and non-spatial, required a similar ability to dis-embed and restructure problems.

Hegarty and Sims (1994) found that high spatial ability and performance on tasks involving the mental animation of a mechanical system relate. They suggest that poor performing participants with a low spatial ability might process spatial transformation inaccurately or have a visuospatial sketchpad with a smaller capacity. Kozhevnikov, Hegarty and Mayer (2002) presented graphs of motion to high and low spatial ability participants. They asked them to visualize and interpret the motion of an object. High-spatial ability participants interpreted the graph as an abstract schematic representation and generated a correct description of the object's motion. Low-spatial ability participants tended to interpret the graph literally as a pictorial illustration of a situation. In addition, Kozhevnikov, Motes, and Hegarty (2007)

found that low-spatial ability participants had problems solving kinematics problems when they had to combine two motion vectors or switch their frames of reference.

Previous research mostly focusses on Physics and Chemistry and uses a small number of representations and tasks. The present study measures the ability to solve process-diagram problems in Biology and the relationship with prior knowledge, spatial ability and working memory. For this purpose, a test with process-diagram problems was designed. In the method section, we formulated several hypotheses about the relation of performance on the process-diagram test with prior knowledge, spatial ability, and working memory.

3. METHOD

3.1 Participants

The participants were 42 secondary school pre-university students from a high school in the Netherlands (mean age 18 years, 22 females). The students participated voluntarily. The last three years of their study in secondary education they chose Biology as a major topic within their exam program for which they received 480 hours of education.

3.2 Data collection

Data collection was spread over two days within a two week period. The tests were planned just before the students' final national exams and were administered under school time in a classroom at their school.

3.3 Process-diagram test

To provide evidence whether the process-diagram test contains a representative sample of process diagrams and tasks, we will describe the construction process, the included process diagrams, and justify the tasks included. In the result section, we will report on the homogeneity and descriptive statistics.

3.3.1 Construction

The process-diagram test was designed in two stages. First, the first and the second author (respectively a part-time high school Biology teacher with 10 years' experience and an expert in the construction of national exams) designed a first version of the process-diagram test and the scoring model. Two external national exam experts and another high-school Biology teacher evaluated this first version; they confirmed face validity. The suggestions for improvement from the external experts made us to revise the final version.

3.3.2 Process diagrams

We included a total of 28 diagrams (Table 3.1), selected from previous national Biology exams, biology text books (e.g., Campbell & Reese, 2002), and the Internet in the test. We redesigned most of the diagrams to be understood without any additional instructional, explanatory and/or contextual text.
The process-diagram test aims to contain a good reflection of process diagrams used in secondary-education Biology. Therefore we selected four biological topics, i.e., ecology, protein synthesis, dissimilation, and hormones. The diagrams we selected include a variety of components (range = 1-30), arrows (range = 2-29), and conventions (i.e., from abstract text boxes to less abstract iconic pictures). Diagrams used for instruction were not included in the process-diagram test as we found it important that students did not see any of the diagrams included in the process-diagram test before.

3.3.3 Tasks

The process-diagram test consists of 97 tasks. Students' ability to solve process-diagram problems will be measured by their performance on these tasks. All tasks were scored as correct or incorrect.

Each topic of the process-diagram test contains tasks with a low cognitive demand and a high cognitive demand. Table 3.2 presents some examples. We categorized the tasks based on Guthrie et al. (1993), Crowe et al. (2008), Kragten, Admiraal, and Rijlaarsdam (2013a), and the cognitive load theory (Sweller, 1994).

Tasks with a 'low cognitive demand' require only a few elements to be explored and/or element interactivity is low. Once the relevant information is selected and encoded, formulating a correct answer requires little cognitive processing: the information could be easily read from the diagram. These tasks ask, for instance, for summarizing the elements found, describing a part of the process step-by-step, and/or some simple calculations like adding or subtracting amounts. For instance, to answer the first low cognitive task from Table 3.2, a student can calculate the increase or decrease per compartment easily (i.e., adding the incoming arrows and subtracting the outgoing arrows). The student calculates each compartment independently, so element interactivity is low.

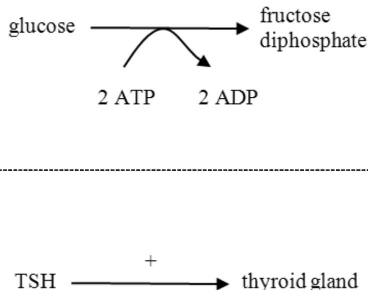
A 'high cognitive task demand' is usually a more global task (Guthrie et al., 1993); a large part or the entire diagram needs to be explored and the components interact. Once the selected information is found, a mental model must be built in working memory (Buckley, 2000; Hegarty & Just, 1993) and integrated (evaluated, inferred, compared, judged) with prior knowledge.

We expected that the scores on these two tasks types differ significantly, as this indicates the validity of these concepts in the process-diagram test. Furthermore, we expected that scores on tasks with a low cognitive demand and a high cognitive demand are correlated because both task types were predicted to rely on prior knowledge and selecting and encoding the information in the presented diagram.

Table 3.1. Diagrams included in the process-diagram test

Topic	Diagrams	Tasks		Examples
		LCTD	HCTD	
Ecology	Carbon cycle on earth	6	2	
	Food web in a fresh lake	2	0	
	Carbon cycle in an American lake[a]	3	3	NO_3^- (soil) —denitrification→ N_2 (atmosphere)
	Flow of energy tropical rainforest	-	2	
	Nitrogen cycle on earth	7	-	
	Phosphorous cycle in a Dutch fresh lake	4	-	
	Nutrient cycle in an ecosystem	6	-	
	Balancing the nitrogen cycle in Dutch agriculture	-	3	plants —2 g/C/h→ consumers
	A global climate model[b]	1	2	
	Nitrogen cycle in traditional Chinese agriculture	3	-	
	Total	*32*	*12*	
Protein synthesis	Infection with a retrovirus	8	1	nucleus: DNA —transcription→ RNA —translation→ protein
	The lytic and lysogenic cycle of a bacteriophage	-	1	
	Translation at a ribosome	5	1	
	Tryptophan synthesis and feedback	1	4	
	Total	*14*	*7*	

Table 3.1. Continued

Dissimilation	Decarboxylation and citric acid cycle	-	4	
	Anaerobic dissimilation of glucose	1	-	
	Dissimilation and the formation of ATP	2	2	
	Oxidative phosphorylation	1	1	
	Dissimilation of glucose by two bacteria[a]	2	2	
	Glycolysis	3	-	
	Total	*9*	*9*	
Hormones	Hormonal regulation of sperm production	4	-	TSH —+→ thyroid gland
	Negative feedback after injection with hormones	-	1	
	Feedback and hormonal effects	-	1	
	A theoretical model of hormonal regulation	2	-	
	Pituitary gland, ovaries and uterus[a]	1	-	
	Hormonal regulation of growth in a human	1	1	
	Types of feedback loops[a]	1	-	
	Indigestion hormones of the stomach	-	2	
	Total	*9*	*5*	

Note. LCTD = low cognitive task demand; HCTD = high cognitive task demand

[a]A depiction with multiple diagrams that a student, for instance, had to compare. [b]The arrows of this climate model represent a feedback mechanism, i.e., see hormones.

Table 3.2. Examples of tasks with a high and low cognitive demand from the process-diagram test

Low cognitive task demand	
	1. There are compartments in which the amount of carbon decreases. Give the name and the amount of decrease of these compartments. (Carbon cycle on earth)
	2. Describe each step (1-8) of the infection with a retrovirus. (Infection with a retrovirus)
High cognitive task demand	
	1. Paul states: "If the combustion of fossil fuel remains 5 Gigatons a year then it will increase by 50 Gigatons in 10 years". Reason why this statement is wrong. (Carbon cycle on earth)
	2. Explain how the loss of half products, e.g., α-ketoglutarate, during the citric acid cycle can be compensated.(Decarboxylation and citric acid cycle)

Note. Between parentheses is the name of the associated diagram in Table 3.1.

3.4 Prior knowledge, spatial ability and working memory tests

The tests on prior knowledge, spatial ability and working memory and their hypothesized relationship with low or high cognitive demand tasks of the process-diagram test are presented in Table 3.4.

3.4.1 Prior knowledge

Students' prior knowledge about the topics in the process-diagram test was measured by a test with 56 open and closed questions. The prior knowledge test consists of questions asking for the recall of basic concepts, e.g., 'What is the definition of an ecosystem?', and tasks asking for the understanding of processes, e.g., 'What is the role of a producer in an ecosystem?' We hypothesize that prior knowledge relates positively to both low and high cognitive task demand. Achievement on a task with a low cognitive demand relies on searching and encoding information, facilitated by domain specific knowledge (Winn, 1993). The presence of knowledge schemas facilitates achievement on a task with a high cognitive demand because such schema's keep cognitive load low (Mayer & Moreno 2003).

3.4.2 Spatial ability

In the present study, a number of spatial ability tests were selected from the Ekstroms' kit of factor referenced cognitive tests (Ekstrom et al., 1976). These tests were used in previous research on learning Science and interpreting scientific diagrams (e.g., Hegarty & Sims, 1994; Kozhevnikov et al., 2002; Kozhevnikov et al., 2007), and dual coding working memory models (e.g., Mayer & Sims, 1994). For parsimony reasons we will not describe these tests in full detail because they have been discussed extensively in previous literature.

For spatial orientation and visualization, we included the Card Rotation Test, Cube Comparisons Test, Form Board Test, Paper Folding Test, and Surface Development Test. These tests require the manipulation of a figure's spatial orientation; for visualisation the figure must first be restructured.

The interpretation of process diagrams requires a specific set of procedural knowledge. For instance, although the main theme of an ecological diagram might be carbon flux (i.e., movement of carbon per unit of time), mentally visualizing the flow of carbon would not be a very effective strategy. More likely is that a participant would encode the diagram into a more static mental model and a propositional causal model. Then the participant may explore for problem solution (Schnotz & Bannert, 2003) in a piecemeal manner (Hegarty, 1992) by applying rules and conventions.

Therefore, we hypothesized that the tests on visualisation and spatial orientation factors were uncorrelated to both low and high cognitive demand tasks of the process-diagram test. Indeed process-diagram tasks do not require rotation or actually visualizing the movement of components: Most studies that found correlations between visualisation and/or spatial operation factors and interpreting scientific diagrams focussed on tasks that require mental operations (e.g., Hegarty & Sims, 1994; Kozhevnikov et al., 2002; Mayer & Sims, 1994).

The Choose a Path Test, i.e., a marker test for the spatial scanning aptitude factor (Ekstrom et al., 1976), was also administered. In this test, each item consists of a diagram with a network of lines; participants must find a line that connects two components among a complex field of dead ends. Scores on the Choose a Path Test were expected to be influenced by students' ability to configure and discriminate the presented elements, i.e., a crucial step when people search for information in a diagram (Winn, 1993). We hypothesized that achievement on tasks from the process-diagram test with a low cognitive demand will positively correlate with scores on the Choose a Path Test because these tasks focussed primarily on selecting the correct information. Scores on tasks with a high cognitive demand will not correlate to the Choose a Path Test scores because these tasks require skills like making inferences, in addition to selecting and encoding information.

3.4.3 Working memory

Miyake et al. (2001) concluded that simple storage-oriented tasks in the visuospatial domain are good predictors for the amount of storage in the visuospatial sketchpad and the closely connected central executive, i.e., the regulating and controlling system of working memory (Baddeley, 1986).

The Shape Memory Test (Ekstrom et al., 1976) measures the ability to remember a group of shapes and their positions in relation to each other. The shapes are abstract forms that one cannot easy encode in any other modality than visual. Students with smaller visual working memory capacity could experience cognitive overload when the cognitive task demand is high. For tasks with a high cognitive demand, students need to build and explore a mental model that draws on the capacity of

visuospatial memory (Sims & Hegarty, 1997). For this, we expect that visual working memory correlates to high cognitive task demand. For low cognitive tasks, there is no need to build complex mental models because the task does not demand this strategy, i.e., students approach a diagram in a goal-based manner (Winn, 1993).

3.5 Data analysis

First, we calculated descriptive statistics for the process-diagram test, the prior knowledge test, and the spatial ability and working memory tests. The process-diagram test and the prior knowledge test were also tested for internal reliability indicated by KR-20. Then, we used correlations to show the relationships between process-diagram test, on the one hand, and prior knowledge, spatial ability and working memory, on the other hand.

4. RESULTS

4.1 Students ability to solve process-diagram problems

Table 3.3 presents the descriptive statistics for the process-diagram test, the prior knowledge test, the spatial ability tests, and the working memory test. The average score on 64 tasks (*M* = 41.14, *SD* = 8.54) of the process-diagram test with low cognitive demand was 64% correct (range = 30%-88%). The average score on 33 tasks (*M* = 14.17, *SD* = 6.03) of the process-diagram test with high cognitive demand was 43% correct (range = 9%-79%).

Table 3.3. Descriptive statistics for the process-diagram test, spatial ability tests and working memory test, and the prior knowledge test

Variable	Test	Scoring items	Min	Max	*M*	*SD*
Low cognitive task demand[a]	Process-diagram	64	19	56	41.14	8.54
High cognitive task demand[b]		33	3	26	14.17	6.03
Spatial ability:						
Spatial orientation	Card Rotation	80	32	80	63.00	12.29
	Cube Comparisons	21	1	19	11.33	3.58
Spatial scanning	Choose a Path	16	1	16	9.90	4.65
Visualisation	Form Board	24	3	20	11.52	4.39
	Paper Folding	10	-1	10	5.81	2.80
	Surface Development	30	-5	30	22.38	8.49
Working memory	Shape Memory	16	3	15	9.79	3.67
Prior knowledge	Prior knowledge[c]	56	26	49	38.81	6.28

Note. Min = minimum score of a student; Max = maximum score of a student. The prescribed scoring procedure from Ekstrom et al. (1976) was adopted
[a]KR-20 = .85; [b]KR-20 = .82; [c]KR-20 = .78

For the process-diagram test, internal reliability indicated by KR-20 was .85 for tasks with a low cognitive demand and .82 for tasks with a high cognitive demand. Figure 3.1 presents the boxplot and a scatterplot for students' scores as percentages of correct answers on tasks of the process-diagram test with low and high cognitive demand. A paired samples *t*-test showed that students successfully completed significantly more tasks with a low cognitive demand (M = .64, SD = .13) than tasks with high cognitive demand (M = .43, SD = .18; $t(41) = 10.00$, $p < .001$, $d = 1.34$). Tasks of the process-diagram test with low cognitive demand and tasks with a high cognitive demand correlated significantly, $r = .66$, $p < .01$.

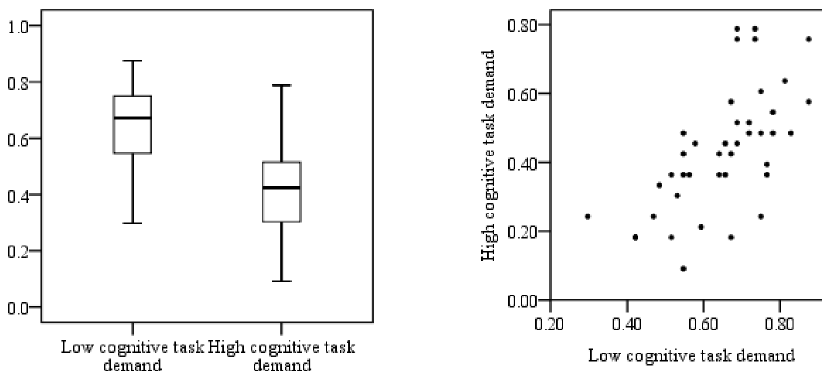


Figure 3.1. Boxplot and scatterplot for students' scores on tasks of the process-diagram test with low and high cognitive demand. The scores are presented as percentages of correct answers.

The scores on the prior knowledge test (M = 38.81, SD = 6.28) were relatively high with an average score of 69% correct answers, and ranged from 46% to 88% correct. KR-20 for the prior knowledge test was .78 after removal of two items.

4.2 The relationship between the scores on the process-diagram test and included explanatory tests

As hypothesized, scores on tasks of the process-diagram test with low cognitive demand correlates significantly with scores on the prior knowledge test (r = .46) and the Choose a Path Test (r = .43); the significant correlation with the Surface Development Test (r = .53) was not hypothesized* (Table 3.4). We found no significant correlations, as hypothesized, between low cognitive task scores and the Card Rotation Test, the Cube Comparison Test, the Form Board Test, the Paper Folding Test, and the Shape Memory Test.

* The Surface Development Test correlated strongly, $r = .62$, $p < .01$ level (2-tailed), with the Choose a Path test.

High cognitive task scores from the process-diagram test correlated, as hypothesized, significantly to the prior knowledge test (*r* = .38) and the Shape Memory Test (*r* = .41); the significant correlation with the Surface Development Test (*r* = .43) was not hypothesized. We found no significant correlations, as hypothesized, between high cognitive task scores and the Choose a Path Test, the Card Rotation Test, the Cube Comparison Test, the Form Board Test, and the Paper Folding Test.

Table 3.4. Predictions and correlations between the process-diagram test and spatial ability, working memory and prior knowledge

Variable	Test	LCTD		HCTD	
		Prediction	*r*	Prediction	*r*
Prior knowledge	Prior knowledge	+	**.46****	+	**.38***
Spatial ability:					
Spatial scanning	Choose a Path	+	**.41****	-	**.18**
Spatial orientation	Card Rotation	-	**.19**	-	**-.04**
	Cube Comparisons	-	**.07**	-	**.12**
Visualization	Form Board	-	**.14**	-	**.14**
	Paper Folding	-	**.12**	-	**-.05**
	Surface Development	-	.53**	-	.43**
Working memory	Shape Memory	-	**.20**	+	**.41****

Note. Predicted correlations are presented by a plus (+; correlated) or minus (-; uncorrelated) sign. Found correlations (i.e., correlated and uncorrelated) that were hypothesized are printed ***bold****. LCTD = Low cognitive task demand; HCTD = High cognitive task demand*
**p < .05 (2-tailed). **p < .01 level (2-tailed)*

5. DISCUSSION

The present study measured the ability to solve process-diagram problems in Biology and the relationship with prior knowledge, spatial ability and working memory. The process-diagram test developed in this study contains a valid representation of process diagrams and adjacent tasks used in secondary education Biology. The test consists of 97 tasks (64 low and 33 high cognitive demand) and 28 diagrams. The mean scores on tasks with a low and high cognitive demand differed significantly; the internal homogeneity of both subtests was high. Therefore we conclude that task difficulty was operationalized reliably and validly. Both subtests correlated. It seems that similar skills and knowledge accommodate achievement on both task types.

As hypothesized, scores on the prior knowledge test correlated positively with tasks with a low and a high cognitive demand. We expected that prior knowledge would correlate to a low cognitive task demand because it facilitates search (Winn, 1993) and with a high cognitive demand because knowledge schemata keep cognitive load low. The correlation between the prior knowledge test and tasks with a high cognitive demand from the process-diagram test was moderate (though significant).

Scores on the Choose a Path Test also positively correlated, as hypothesized, with low cognitive demand tasks. We assume the Choose a Path Test to be a measure for the ability to search for information in a complex spatial diagram.

We hypothesised that tests from the visualisation and spatial orientation factors would be uncorrelated to scores on tasks from the process-diagram test with low and high cognitive demand. The latter was confirmed except for scores on the Surface Development Test: The Surface Development Test correlated with scores on both low and high cognitive demand tasks. The strong correlation between the Surface Development Test and the Choose a Path Test might explain this unexpected finding. Presumably both tests tap, to some extent, the same ability, i.e., configuring elements in a complex spatial field.

Performance on tasks of the process-diagram test with a high cognitive demand was expected to correlate to the scores on the Visual Memory Test. The Visual Memory Test was expected to be a measure for the capacity of the visuospatial sketchpad available for constructing a runnable mental model. A moderate correlation was found between the Visual Memory Test scores and task scores from the process-diagram test with high cognitive demand.

We conclude that the ability to solve process-diagrams problems involves the presence of prior knowledge, spatial abilities, and visuospatial working memory capacity. This study thereby adds to a large body of previous research on the role of these factors in learning from external representations. The correlations we found are, however, not fully congruous with previous studies (e.g., Hegarty & Sims, 1994; Kozhevnikov et al., 2002; Mayer & Sims, 1994) and show that solving process diagrams problems with a low and a high cognitive demand both require different spatial skills.

A limitation of the present study is the specific focus on biological process diagrams. We choose these types of diagrams because of the significance in teaching and learning Biology. However, hesitate to generalize the findings to other types of diagrams (e.g., tree or anatomical diagrams) or other scientific domains. Furthermore, we only found moderate correlations, suggesting that other factors (e.g., strategy use) might also be important in process-diagram problem solving.

All in all, we think this study might help the Biology education community. The study stresses that prior knowledge must be present (and activated) when students are presented process diagrams. Students who study process diagrams, teachers who use process diagrams for teaching biological processes, and instructional designers who incorporate process diagrams in study material should anticipate to this. Furthermore, scores on tasks with a high cognitive demand were below average. We suggest that specific training on solving these type of problems and the interpretation of process diagrams in general might be needed. Finally, this study shows that even within a homogenous group (i.e., pre-university students with an extensive Biology training) variance in spatial ability factors account for individual differences in solving process-diagram problems. These students might particularly benefit from a training program that includes a more strategic approach to interpret process diagrams.

CHAPTER 4

STUDENTS' LEARNING ACTIVITIES WHILE STUDYING BIOLOGICAL PROCESS DIAGRAMS*

Process diagrams describe how a system functions (e.g., photosynthesis) and are an important type of representation in Biology education. In the present study, we examined students' learning activities while studying process diagrams, related to their resulting comprehension of these diagrams. Each student completed three learning tasks. Verbal data and eye-tracking data were collected as indications of students' learning activities. For the verbal data, we applied a fine-grained coding scheme to optimally describe students' learning activities. For the eye-tracking data, we used fixation time and transitions between areas of interest in the process diagrams as indices of learning activities. Various learning activities while studying process diagrams were found that distinguished between more and less successful students. Results showed that between-student variance in comprehension score was highly predicted by meaning making of the process arrows (80%) and fixation time in the main area (65%). Students employed successful learning activities consistently across learning tasks. Furthermore, compared to unsuccessful students, successful students used a more coherent approach of interrelated learning activities for comprehending process diagrams.

1. INTRODUCTION

Graphical representations are becoming increasingly prominent as carriers of meaning (Bezemer & Kress, 2008; Roth & McGinn, 1998). In current science textbooks, diagrams are more-and-more becoming instructional entities that can be studied, to some degree, independently from the text (e.g., Reece et al., 2010). Process diagrams are important and abundantly present in the diagram category. They describe how a system functions (e.g., photosynthesis, biogeochemical cycles) through the use of components that are connected by arrows. These arrows indicate transformation, movement, sequence, etc. (Heiser & Tversky, 2006). Several studies demonstrate that students have difficulties interpreting such diagrams (e.g., Chittleborough & Treagust, 2008; Kriz & Hegarty, 2007; Schönborn, Anderson, & Grayson, 2002). Hence, a fine-grained analysis of learning activities that contribute to the diagram comprehension process can support the design of an evidence-based training program to foster learning from such diagrams.

Two popular process-tracing techniques to obtain an online record of students' learning activities are the think-aloud protocol (e.g., concurrent, Ericsson & Simon,

---

* Kragten, M., Admiraal, W., Rijlaarsdam, G. (2015). Students' learning activities while studying biological process diagrams. *International Journal of Science Education, 37*, 1915–1937 doi:10.1080/09500693.2015.1057775

1993; cued retrospective, Van Gog, Paas, Merriënboer, & Witte, 2005) and eye tracking (Holmqvist et al., 2013).

1.1 Learning activities

Learning is an active process of knowledge construction; students build an internal mental model to comprehend a diagram's content by employing learning activities, e.g., encoding, inference (Hegarty, 2005). When students study diagrams autonomously, they must regulate the occurrence of these learning activities. They must employ cognitive as well as metacognitive learning activities and use conceptual and procedural domain knowledge to achieve learning goals (Boekaerts, 1997).

1.1.1 Cognitive learning activities

Various studies have described cognitive (and metacognitive) learning activities students used while studying texts (Pressley, 2000; Pressley & Afflerbach, 1995) and texts with diagrams (Butcher, 2006; Azevedo & Cromley, 2004; Cromley, Snyder-Hogan and Luciw-Dubas, 2010). Cromley et al. (2010) collected verbal protocols from first year biology majors reading an 8-page passage from their textbook which included seven diagrams. A wide range of cognitive learning activities were distinguished, including activating prior knowledge, paraphrasing, summarizing, and inference. The authors examined whether students employed different cognitive learning activities when studying diagrams vs. full text. They found that when studying diagrams the students made more inferences than in texts. Two variables were positively related to scores that indicated the elaborateness of the mental model: (1) the variety of the three types of inferences (i.e., inferences, elaborations, and hypotheses), and (2) the greater use of inferences. The latter finding is in line with many other studies reporting inferencing as crucial for the comprehension of graphical representations (Chi, 2000; Cromley et al., 2013; Hegarty, 2005; Kriz & Hegarty, 2007).

1.1.2 Metacognitive learning activities

Metacognitive learning activities regulate the cognitive learning activities. They are powerful predictors of general academic achievement (Wang, Haertel, & Walberg, 1990) and science achievement (e.g., Akyol, Sungur, & Tekkaya, 2010). Meijer, Veenman and Van Hout-Wolters (2006) set up a taxonomy of metacognitive learning activities for the interpretation of think-aloud protocols of secondary school students. This taxonomy was constructed for analyzing protocols of students studying History texts and solving Physics problems. The taxonomy contains six main categories: orientating, planning, executing, monitoring, evaluating, and elaboration. Science reading research shows that proficient readers engage in orienting activities like reading the title and subheadings and planning activities concerning decisions about how to navigate through the instructional material (Pressley, 2000; Pressley &

Afflerbach, 1995; Veenman, 2012). Furthermore, proficient readers reread difficult or important parts and they generate and answer questions (Veenman, 2012).

1.1.3 Domain knowledge

The importance of domain knowledge for the interpretation of scientific graphical representations is well documented (Cook, 2006; Canham & Hegarty, 2010; Cook, Carter & Wiebe, 2008; Kriz & Hegarty, 2007). For instance, Kriz and Hegarty (2007) evaluated learning from animated displays (i.c., a flushing cistern). They found that participants with high domain knowledge (i.c., engineering students with no specific prior knowledge about how the flushing cistern worked) were more likely to construct a correct mental model (with correctness indicated by reported number of steps in the causal chain of the flushing cistern mechanism) than participants with low specific domain knowledge (i.c., humanities and social studies students). Interestingly, a majority of the students with high domain knowledge initially constructed an incorrect mental model. However, in contrast to the participants with low domain knowledge, these students were able to revise their mental models (with correctness indicated by correct answers on comprehension questions).

1.2 Learning activities and eye-movements

The eye-mind hypothesis claims a direct relationship between fixation durations and on-going cognitive processes, where longer fixation durations indicate more extensive processing (Just & Carpenter, 1976; She & Chen, 2009). A crucial and difficult step in eye-tracking data is to determine which learning activities take place during eye movements. For instance, transitions (i.e., shifting focus from one location to another) could refer to either active integration of several parts of the representation (Schwonke, Berthold & Renkl et al., 2009; Mason, Pluchino, & Tornatora, 2013) or random ineffective searching behavior. Furthermore, students' attention might shift to specific parts of a representation as due to prior knowledge or by an effect induced by the display (Canham & Hegarty, 2010; Kriz & Hegarty, 2007).

The influence of prior knowledge on eye movements has been examined within several domains. Cook et al. (2008) collected eye-tracking data to examine how prior knowledge influenced students' eye movements and interpretation of a graphical representation of cellular transport mechanisms. They found that low prior knowledge students tended to focus on more salient features (e.g., colored proteins), whereas high prior knowledge students tended to focus on more thematically relevant content.

Van Gog, Paas and Van Merriënboer (2005) examined differences in eye movements between high and low expertise participants while performing a troubleshooting task about an electrical circuit. The level of expertise was indicated by a combined score of task performance, self-reported mental effort, and standardized mean score of total fixation time. Three phases of the troubleshooting task were defined: 1) problem orientation, 2) problem formulation and action decision, and 3) action

evaluation and next action decision. They found that participants with high expertise spent relatively more time on the orientation phase and on evaluating their actions and deciding on their next action than participants with low expertise.

Several studies examined the relationship between eye-tracking measures, e.g., fixation durations and transitions, and learning outcomes of scientific graphical representations (with or without explanatory text). Mason et al. (2013) examined how students learn from a science text in two conditions: with concrete or abstract illustrations. These researchers related a variety of eye-tracking measures to immediate and delayed post-tests on factual knowledge and on transfer. Most significant positive correlations were found for both post-tests on factual knowledge in the condition with abstract illustrations. Positive eye-movement variables were, among others, fixation duration on the illustration, first-pass fixation duration on the illustration and transitions from the end text segment to the illustration. They concluded that for the text with the abstract illustration, the quality of the learning performance is associated with higher fixation duration and more attempts to integrate the graphical and verbal information.

Schwonke et al. (2009) presented their participants worked-out examples of probability problems that included multiple representations, i.e., a diagram, text, and an equation area. Participants had to relate and integrate these representations to build an elaborate mental model of probability theory. The researchers correlated eye movement measures (mean fixation time, cumulative fixation duration, mean fixation duration, transition frequency) with participants' conceptual understanding and transfer performance. No relations with transfer performance were found. However, conceptual understanding correlated significantly with the mean and cumulative fixation duration on the diagram.

1.3 Focus and rationale of the present study

In the present study we aim to provide an in-depth analysis of which learning activities significantly predict students' comprehension level while studying a specific type of graphical representation, i.e. the process diagram. Learning from graphical representations has been the interest of research for the last few decades in various fields, for instance in contributions to multimedia theory (e.g., Mayer, 2001), cognitive load theory (e.g., Carlson, Chandler & Sweller, 2003), and the role of prior knowledge and inferences (e.g., Canham & Hegarty, 2010; Cook, 2006; Cook et al., 2008; Kriz & Hegarty, 2007; Schwonke et al., 2009). Some diagram types, such as mechanical diagrams (e.g., Hegarty & Just, 1993; Kriz & Hegarty, 2007) or evolutionary diagrams (e.g., Catley, Novick, & Shade, 2010) have been studied extensively. However, research on process diagrams is limited. Körner (2005) demonstrated that students benefit from specific training on a specific graphical representation, i.e., the hierarchical graph. A fine-grained analysis of how the more successful students learn from process diagrams might facilitate the design of a specific training for this diagram type.

1.4 Research questions

The present study has two perspectives: (1) we monitored students' learning activities while they successively studied three different process diagrams and (2) we examined learning activities by triangulating eye-tracking and verbal data. The first perspective will give us an overview of learning activities that students employed while studying process diagrams. This will allow a detailed and robust analysis of which learning activities significantly predict students' level of comprehension of process diagrams. Such an analysis will enable us to determine whether these activities can be regarded as indicators of a more general strategic approach. The analysis of a series of tasks per student will enable us to determine whether students employed learning activities consistently across tasks.

The second perspective will serve three purposes. First, using alternative methods for assessing learning activities can validate these methods. Second, using two alternative methods might allow us to identify learning activities that otherwise would remain undetected. Third, relating eye-tracking and verbal data can tell us whether and to what extent relatively easy to collect eye-tracking data can indicate the occurrence learning activities. This might be informative for researchers who are interested in obtaining eye-tracking data as an alternative for verbal reports as the latter are very labor intensive to analyze.

The following research questions were defined:

1) Which learning activities distinguish between relatively more and less successful students studying process diagrams? We answer this question by examining which learning activities are employed and when, after which we relate the frequency of these learning activities to the level of comprehension. Furthermore, we will examine whether students employed these significant learning activities consistently across tasks.
2) Are learning activities that significantly predict the comprehension of process diagrams related?

2. METHOD

2.1 Participants

Forty-two students from two classes in a regular secondary school in the northwest of the Netherlands were invited to participate in the study. A total of 32 students volunteered: 10 students declined the invitation mainly because they were not able to schedule an appointment due to other commitments during the data collection period. The students were finishing their last year of pre-university upper secondary education. They had chosen Biology as a major topic within their exam program with a study load of 480 hours during three years of upper secondary education. The students were in high-performing classes with a focus on Science, i.e., all students had also chosen Chemistry and Math and most had also chosen Physics; they are likely to pursue an academic career. Finally, after the calibration procedure (see section 'Eye-tracking apparatus and materials'), we had to exclude three datasets. So 29

students (14 female, M_{age} = 18.3 years, age range: 17–19 years) participated in the present study.

2.2 Procedure for Learning Tasks

Three learning tasks were conducted consecutively in a quiet room during a single session of approximately one hour with each individual participant. The students were first acquainted with the cued retrospective think-aloud procedure (Van Gog, Paas, Merriënboer, & Witte, 2005)—which followed the first two learning tasks—by a warm-up session with a small process diagram. The students were then allowed to ask questions about the procedure; the experimenter could provide some guidance. The learning tasks, conducted by the first author (i.e., the experimenter), were then given to the students while their eye movements were monitored.

The first learning task started after the calibration procedure (see section 'Apparatus and materials') proved successful. The students received the following instruction:

> "You will be presented with a process diagram. Try to understand as much as you can. The maximum time allowed is 4 minutes. If you are ready sooner you can stop by pressing the left mouse button."

Next, the students had access to the process diagram. Half the students started with diagram 1 (Figure 4.1a), the other half with diagram 2 (Figure 4.1b). When students completed this learning task, the instruction for the cued retrospective think-aloud procedure followed:

> "You will be presented with an animation of where you were looking while you were learning from the diagram. Try to tell as much as you can about what you were thinking. Tell anything that comes to mind, act like you are alone and nobody is listening and keep talking. The animation will be played at half speed and you can pause and continue whenever you like."

Next, students were shown a replay of their eye movements using a spotlight. The animation ran at 0.5x normal speed. Students could pause and continue the animation whenever necessary by pressing the space bar. The procedure of the first learning task was repeated for the second learning task; when student were presented diagram 2 in the first learning task they were presented diagram 1 in the second learning task and vice versa.

In the third learning task, students were presented with diagram 3 (Figure 4.1c) and were again instructed to try to learn as much as they could from this diagram. In addition, we informed them they had to complete a test about what they have learned from diagram 3; No think-aloud protocol was collected.

In sum, we collected eye-tracking data for three learning tasks and verbal data for the first two of them. Students' comprehension was directly inferred from the verbal data from learning task 1 and learning task 2; comprehension of learning task 3 was measured by a test (see section 'Measuring comprehension').

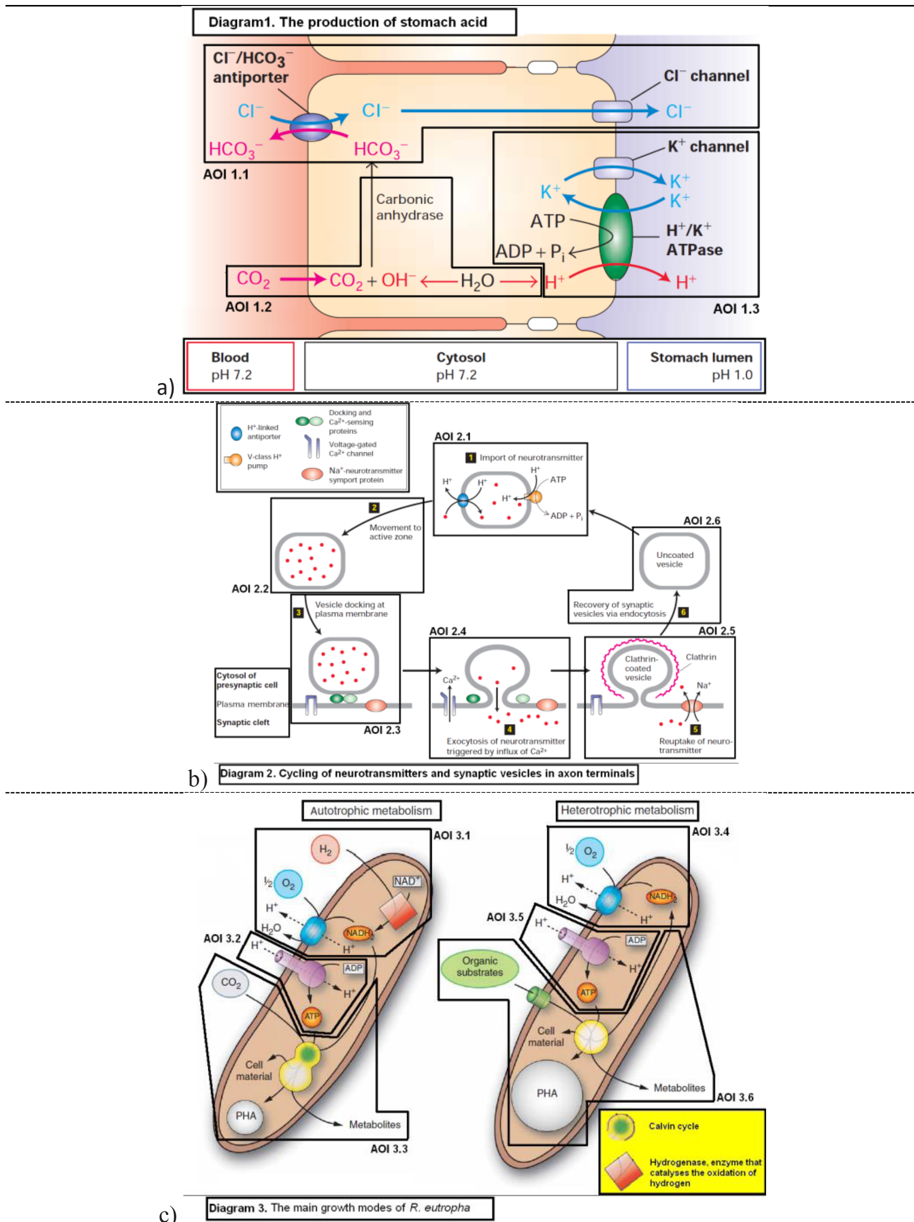


Figure 4.1. Three process diagrams used for the learning tasks. Diagrams used in the study were translated to Dutch. Areas of interest are indicated by thick black lines. The 15 areas of interest of the main area are numbered, i.e., Diagram 1 (1.1 till1.3), Diagram 2 (2.1 till 2.6), and Diagram 3 (3.1 till 3.6).

2.3 Selection of Process Diagrams

In this study we focus on *learning* processes. Therefore, the topics of the diagrams were new and not part of the past curriculum. Diagram 1 and 2 were slightly adjusted versions of diagrams from an international textbook for university students (Lodish et al., 2012). Diagram 3 was an adjusted version of a diagram from a scientific journal (Pohlmann et al., 2006). The process diagrams (Figure 4.1) were graphical representations with few verbal elements, e.g., a title, a legend, and labels; there was no support from explanatory text. The diagrams were translated into Dutch. Students were not familiar with the processes presented in the diagrams but should have enough skills and knowledge to be able to comprehend the content; this was confirmed by their Biology teacher. With regard to diagram 1, students were taught in previous years about transport mechanisms (e.g., active and passive transport) and the composition of stomach acid (i.e., hydrogen chloride). But they were not taught how these processes relate to the process of the production of stomach acid. Concerning diagram 2, students were taught in previous years about how neurons work, but not about the cyclic process of regaining the neurotransmitters and the synaptic vesicle. As to diagram 3, students were taught in previous years about biochemical processes like assimilation and dissimilation, but not with hydrogen as an energy source or within bacteria that switch between metabolic modes.

2.4 Learning activities from verbal reports

We established a coding scheme (Table 4.1) for the cued retrospective think-aloud protocols by continuously switching between a theoretical and data-driven point of view. We started with a broad range of previously defined learning activities from studies that focused on learning from text and diagrams (e.g., Butcher, 2006; Azevedo & Cromley, 2004; Cromley et al., 2010; Meijer et al., 2006; Van Gog, Paas, & Merriënboer, 2005).

The final coding scheme distinguishes between an orientation phase and an elaboration phase. We defined three main categories of activities, i.e., *cognitive*, *metacognitive* and *diagram learning activities*.

The orientation phase contains all activities that occur before a student starts actually studying the main area (i.e., the numbered areas of interest in Figure 4.1) of the process diagram. Within this orientation phase we found both *diagram learning activities* and *cognitive learning activities*. The category *diagram learning activities* in this phase contains activities using diagrams meta-information like *reading the title*, *reading the labels regarding the level of organization,* and *localizing the legend items in the main area*. The category *cognitive learning activities* in this orientation phase deal with the activation of prior knowledge (Meijer et al., 2006).

The main phase begins when students started describing the process as depicted in the main area. The category *cognitive learning activities* within this phase contains five activities: 1) *giving meaning to a process arrow* (cf., paraphrases, Butcher, 2006; descriptive statements, Lowe, 1999; paraphrasing, Meijer et al., 2006), 2) *inference* (cf., integration inference, Butcher, 2006; knowledge inference, Chi, 2000;

inference, Cromley et al., 2010; causal statements, Lowe, 1999), 3) *relating prior knowledge* (cf., background knowledge, Cromley et al., 2009), 4) *alternative hypothesis* (cf. hypothesizing, Azevedo & Cromley, 2004; Meijer et al., 2006), and 5) *comparing elements across areas of interest* (cf., connecting parts of text by reasoning, Meijer, 2006). *Giving meaning to a process arrow* is defined as correctly describing movement, transformation, or a next step in the diagram as depicted by arrows (Kragten et al., 2013a). Inferences are, as mentioned in the theoretical framework, essential for learning. We defined inferences as statements that include the relation between processes that are not literally displayed (cf., Chi et al., 1994; Chi, 2000).

Students might use basic prior knowledge to facilitate learning from the process diagram. *Relating prior knowledge* defines connections students' made to background knowledge without integrating this into a causal statement. This means that students recognized a part of the process. Some students compared elements across areas of interest (AOIs). We interpreted this as an attempt to construct a global representation (Hegarty & Just, 1993) and coded it as *comparing elements across areas of interest*.

The category *metacognitive learning activities* in the main phase contains 1) *rereading parts of the diagram* (cf., rereads text in diagrams, Cromley et al., 2009; rereading, Meijer et al., 2006), and 2) *self-questioning* (cf., Azevedo & Cromley, 2004; Cromley et al., 2010). We interpreted *self-questioning* as a metacognitive learning activity and not as a cognitive learning activity (cf. Cromley et al., 2010), while students seemed to come up with questions when they are in cognitive disequilibrium. This suggests that students were monitoring their understanding. Initially, there were more subcategories within the main category *metacognitive learning activities* (e.g., evaluation of knowledge). However, the number of observations in these subcategories remained very small and did not show any significant relation with comprehension. For parsimony reasons we did not include these items in the present study.

The category *diagram learning activities* contains the same items as the orientation phase.

2.5 Eye-tracking apparatus and materials

Eye movements were captured using a custom built corneal reflection binocular eye tracker. The eye-tracking software, i.e., ITU Gaze Tracker 2.1b, ran on a dual core 2.5 GHz HP Pavilion laptop with a sampling rate of 60 Hz. The stimulus monitor (i.e., a 22" screen with a resolution of 1920x1080) was connected through a VGA cable with the laptop; an external microphone and keyboard were also connected. This setup enabled the experimenter to perform the necessary actions during the sessions, e.g., starting the calibration procedure, updating the database with eye-tracking data, and starting the measurements. The students sat in front of the screen with the position of their head fixed at 60 cm distance by a forehead rest.

Table 4.1. Coding scheme for the cued retrospective think aloud protocols and descriptive statistics

Phase → *Category* → Learning activity → "Example"	Short term	M_s (*SD*)	
		Learning task 1	Learning task 2
Orientation phase			
Cognitive learning activities			
Activating prior knowledge	Orientate Prior	0.11 (0.31)	0.14 (0.35)
"and when I was reading stomach acid I thought of Chemistry and H_3O^+ etc."			
Diagram learning activities			
Reading the title	Orientate Title	0.79 (0.42)	0.28 (0.45)
"I read the title, the production of stomach acid"			
Reading the labels regarding the organizational level	Orientate Level	0.50 (0.51)	0.07 (0.26)
"blood, cytosol and stomach"			
Localizing legend items in the main area			
"H^+ linked antiporter, where is it in the diagram, oh there at the first step"	Orientate Legend	-	1.66 (1.86)
Main phase			
Cognitive learning activities			
Giving meaning to a process arrow	Meaning Arrow	8.14 (3.32)	5.69 (3.47)
"H^+ is transported to the stomach"			
Inference	Inference	2.04 (1.95)	2.34 (2.00)
"ATP is used to transport H^+ to the stomach"			
Relating prior knowledge	Relate Prior	0.64 (0.87)	0.48 (0.95)
"that is active transport because ATP is being used"			

Table 4.1. Continued

Alternative hypothesis “Chloride en Calcium Carbonate are being exchanged because otherwise the cell gets charged”	Alt. Hypothesis	0.75 (0.89)	1.76 (2.53)
Comparing elements across areas of interest “I’m checking what goes from the blood to the cytosol, what goes to the stomach and back to the blood”	Compare	0.68 (0.98)	1.07 (1.13)
Metacognitive learning activities			
Self-questioning ”but why is ATP used? Is that for H^+ of K^+”	Self-Questioning	3.61 (3.51)	1.93 (2.19)
Rereading parts of the diagram “now I’m going to check everything for the second time”	Meta Reread	1.64 (1.10)	1.10 (0.98)
Diagram learning activities			
Reading the title “so now I’m checking the title”	Read Title	0.64 (0.73)	0.48 (0.57)
Reading the labels regarding the organizational level “the pH of the blood and cytosol is 7.2, the pH of the stomach is 1.0”	Read Level	1.36 (1.28)	1.07 (0.70)
Using the legend “Clathrin, I don’t know what it is so I look in the legend”	Use Legend	-	3.10 (2.23)

Note. M_s *= Mean number of utterances per student.*

Calibration was performed before the start of the learning tasks by displaying 9 white circles that shrank to a smaller circle on a black background. The calibration procedure was considered successful when the residuals for both eyes were < .5° of the visual angle. The learning tasks were designed using OGAMA 4.2, i.e., an open source software project which also supports capturing eye-tracking data send from ITU Gaze Tracker via UDP. Only fixations that lasted 100 milliseconds (or longer) within a 30 pixels diameter were analyzed.

2.6 Learning activities from eye-tracking data

Several Areas Of Interest (AOIs) were defined for each diagram according a functional criterion. i.e., the AOIs enclose more or less different sub-processes of the entire process depicted in the diagram (see Figure 4.1). Therefore the sizes of the AOIs in the main areas varied. The title, legend (diagram 2 and 3 only), and the level of organization, were also defined as AOIs. We assume that fixation time in the AOIs indicated ongoing learning activities; transitions between the AOIs indicated integration activities (see Table 4.2).

Table 4.2. Description of eye-tracking measures

Eye-tracking measures	Explanation	*M* (*SD*)		
		Learning task 1	Learning task 2	Learning task 3[b]
Fixation time	Total time spent (s) in:			
Main	the AOIs of the main area	100.30 (40.45)	91.38 (36.37)	158.61 (62.03)
Title	the AOI where the title is located	2.36 (1.22)	3.91 (3.80)	6.80 (4.95)
Level	the AOI where the level of organization is located	8.37 (4.82)	6.01 (3.95)	7.55 (4.99)
Legend	the AOI where the legend is located	[a]	23.77 (12.30)	9.18 (5.82)
Transitions	Number of transitions:			
Main	Between AOIs of the main area	42.86 (20.66)	43.86 (21.67)	115.24 (46.43)
Title	from the main area to the AOI where the title is located	2.03 (1.50)	1.41 (1.35)	2.03 (1.59)
Level	from the main area to the AOI where the level of organization is located	11.14 (6.84)	8.45 (7.63)	7.45 (5.44)
Legend	from the main area to the AOI where the legend is located	[a]	19.97 (10.24)	13.28 (7.62)

Note. [a]*Diagram 1 contains no legend.* [b]*Maximum time in learning task 3 was 5 minutes, i.e., 1 minute more than in the other two learning tasks.*

2.7 Measuring comprehension

For the first and second learning task, we inferred the level of comprehension of the depicted process directly from the inferences students made in their verbal reports (cf., Chi et al, 1994). The rationale for not using an additional test is that the students interacted twice with the diagrams; in the learning phase and during the cued retrospective think-aloud protocols. Verbalizing their thoughts might have helped students to strengthen their understanding and this might endanger the validity of an additional test. Students' score on inference was calculated as the sum of the unique number of inferences uttered per learning task.

For the third learning task—after which no verbal data was collected—the students completed a comprehension test. The test consisted of reconstructing the processes of the two main growth modes (i.e., autotrophic and heterotrophic) of *E. eutrophia* by drawing. Asking students to reconstruct a biological process by drawing is commonly used to measure students' understanding of a specific topic (e.g., Quillin & Thomas, 2015; She & Chen, 2009). We did not expect students to use a memorization strategy as they were instructed to 'understand as much as they can' and they were not informed before the learning task on the type of test they had to perform.

The comprehension test had 28 items (M = 13.0; SD = 7.0; 1 = correct; 0 = incorrect); internal reliability indicated by KR-20 was .90. Items consisted of the combination of a component and the associated arrow leading away from the component in the right direction. Two completed examples of a part of the test are presented in Figure 4.2. Note that the students had to reconstruct the autotrophic and the heterotrophic mode; Figure 4.2 only shows two examples of the autotrophic mode. The completed test on the left side of Figure 4.2 is more elaborate (i.e., especially the part of the process where the Calvin cycle takes place) although there are also some mistakes and elements missing. The completed test on the left side scored 9 points and the test on the right side 5 points.

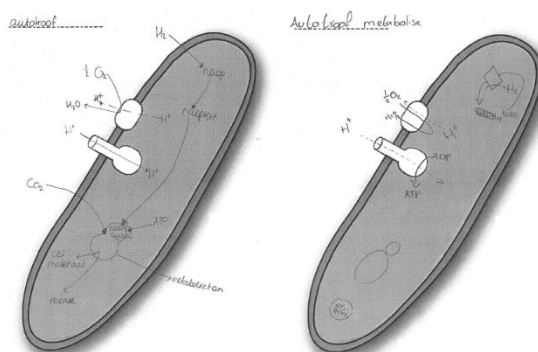


Figure 4.2. Examples of the completed comprehension test of two students (i.e., only the autotrophic part of the test is shown).

2.8 Data analyses

First, all variables were normalized on the basis of the mean and standard deviation of the specific learning task. The rationale behind this normalization was that we use a within-students design where the responses to different learning tasks can be considered as responses to different tests (i.e., a multivariate model). Normalization at the level of the learning task allowed us to focus on differences in student behavior and comprehension score instead of differences evoked by the content or design of the diagrams. For example, a more informative diagram would likely produce more utterances. After normalization, all variables can be interpreted as student's behavior or comprehension score compared to the other students within the particular learning task. The latter allowed us to also analyze between-student and within-student between-diagram differences in behavior and comprehension score across learning tasks.

The first research question was answered by applying the two-level multilevel model presented in Equation 1 (i.e., student level = j and diagram level = i) with students' comprehension score as the response variable (y_{ij}) and learning activity variables as fixed effect explanatory variables.

$$y_{ij} = \beta_{0ij} + \beta_1 \cdot x_{1ij} \qquad (1)$$

$$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$$

The random part includes the between-students residuals (u_{0j}) and the within-students between-diagrams residuals (e_{0ij}). The fixed part includes β_0, i.e., the grand mean, and β_1, i.e., the regression coefficient of an added explanatory variable. Note that β_{0ij} is the intercept-only model if $\beta_1 \cdot x_{1ij}$ is omitted. For each prediction of the response we extended the intercept-only model with one explanatory variable ($\beta_1 \cdot x_{1ij}$) at a time. We did not extend the model to contain multiple explanatory variables. This would place too many restrictions on a model because not all explanatory variables could be obtained from the three learning tasks (see Table 4.1 and Table 4.2). We tested whether the intercept-only model was improved by adding a fixed effect variable by calculating the χ^2 likelihood-ratio (Hox, 2010) with one degree of freedom (i.e., one explanatory variable added). The statistical significance of the fixed effect (β_1) was tested using a two-sided Wald z-test (Hox, 2010). Furthermore, we calculated the amount of variance explained per level, i.e., student (σ_u^2) and diagram (σ_e^2), after adding the fixed effect to the intercept-only model (Hox, 2010, p.71). The amount of variance explained per level might decrease, or even increase (Snijder & Bosker, 1994), depending on the distribution of the explanatory variable within the levels (Hox, 2010). Therefore, we calculated the intraclass correlation (Hox, 2010) of each explanatory variable which provides the distribution of the variance on the student level and diagram level. In short, we alternately added a fixed effect variable to the intercept-only model and then tested the model for signif-

icant improvement, the significance of the fixed effect, and the change in explained variance per level.

The second research question was answered by correlating the normalized scores of the learning activities. Only those learning activities that were shown to be related to the comprehension scores were included. The correlation coefficients provided insight into possible dependencies between successful learning activities, indicating whether students use a coherent set of succesful learning activities across the three learning tasks.

3. RESULTS

3.1 Learning activities related to comprehension of process diagrams

In Table 4.3, we present the regression coefficients of the models (see Equation 1) with a single fixed learning activity variable as the explanatory variable and comprehension score as the response variable.

From the variables in the orientation phase, Orientate Legend (β_1 = .42, p = .01) and Orientate Prior (β_1 = .68, p = .04) were significant predictors of comprehension score. Student variance explained by adding Orientate Legend as a fixed effect to the intercept-only model is 17.4%, χ^2 (1, N = 29) = 5.53, p < .05. Adding Orientate Prior to the intercept-only model explained 6.1% of student and 6.9% of diagram level variance, χ^2 (1, N = 58) = 6.44, p < .05. The intraclass correlation (ICC) of Orientate Prior was 50.9% on the student ($\rho_{student}$) and 49.1% on the diagram level ($\rho_{diagram}$).

From the variables in the main phase, Meaning Arrow (β_1 = .59, p < .001), Self-Questioning (β_1 = .28, p = .02), Read Title (β_1 = .21, p = .05), and Read Level (β_1 = .30, p = .01) appeared to be significant predictors of comprehension score. Adding Meaning Arrow to the intercept-only model explained 79.6% of student and 1.3% of diagram level variance, χ^2 (1, N = 58) = 23.63, p < .001. The intraclass correlation of Meaning Arrow was 41.1% on the student and 58.9% on the diagram level. The variable Self-Questioning explained 4.8% of the variance on the student and 10.49% on the diagram level compared to the intercept-only model, χ^2 (1, N = 58) = 5.13, p < .05. The intraclass correlation of Self-Questioning was 44.1% on the student and 55.9% on the diagram level. The intercept-only model was not improved by adding the variable Read Title, χ^2 (1, N = 58) = 3.20, p > .05. Adding the variable Read Level to the intercept-only model explained 21.8% of the variance on student and 2.4% on diagram level, χ^2 (1, N = 58) = 5.67, p < .05. Read Level had an intraclass correlation of 37.0% on the student and 63.0% at the diagram level.

Table 4.3. Regression coefficients of single variable fixed effects models with comprehension score as dependent variable and intraclass correlations of the explanatory variables

Phase / Variable	β_1 *(SE)*	*p*	σ^2explained *(%)*[a]		ICC (%)	
			student	diagram	$\rho_{student}$	$\rho_{diagram}$
Orientation phase						
Orientate Title*	.35 (.21)	.11	-.6[b]	6.1	47.2	52.8
Orientate Level	.20 (.25)	.43	9.3	-4.3	47.8	52.2
Orientate Legend*	**.42 (.17)**	**.01**	17.4		100.0	
Orientate Prior*	**.68 (.32)**	**.04**	6.1	6.9	50.9	49.1
Main phase						
Meaning Arrow***	**.59 (.10)**	**< .001**	79.6	1.3	41.1	58.9
Relate Prior	.21 (.12)	.08	11.9	1.5	23.2	76.8
Alt. Hypotheses	-.05 (.11)	.66	-4.2	2.6	1.7	98.3
Compare	.15 (.12)	.23	4.2	1.5	37.9	62.1
Self-Questioning*	**.28 (.12)**	**.02**	4.8	10.4	44.1	55.9
Meta Reread	.02 (.13)	.85	3.0	-1.5	39.9	60.1
Read Title[c]	**.21 (.10)**	**.05**	-22.4	18.4	0.0	100.0
Read Level*	**.30 (.12)**	**.01**	21.8	2.4	37.0	63.0
Use Legend	.25 (.18)	.17	6.2		100.0	
Eye-tracking measures						
Fixation Time Main***	**.51 (.09)**	**< .001**	64.8	6.7	55.5	44.5
Fixation Time Title	.15 (.10)	.14	-4.3	4.4	36.0	64.0
Fixation Time Level	.14 (.10)	.15	9.6	.0	26.0	74.0
Fixation Time Legend	.13 (.12)	.28	7.2	.0	2.4	97.6
Transitions Main**	**.29 (.11)**	**.01**	26.7	.8	54.1	45.9
Transitions Title	.02 (.09)	.82	-.7	.2	10.7	89.3
Transitions Level	-.04 (.09)	.65	-4.5	1.7	15.1	84.9
Transitions Legend	.21 (.13)	.10	8.1	3.0	25.0	75.0

Note. Significant regression coefficients (Wald z-test) are in boldface.β_1 = regression coefficient of the explanatory variable (see Equation 1). ICC = Intraclass correlation of the explanatory variable; an indication of the proportion of variance at the student level and diagram level (Hox, 2010, p.15).

[a]Variance explained (Hox, 2010, p.71) after adding a single fixed effect variable compared to the intercept-only model. Student level and diagram level variance of the intercept-only model for comprehension score (i.e., the response variable) per combination of diagrams is: diagram 1, 2 and 3 ($\rho_{student}$ = 46%; $\rho_{diagram}$ = 54%), diagram 1 and 2 ($\rho_{student}$ = 52%; $\rho_{diagram}$ = 48%), diagram 2 and 3 ($\rho_{student}$ = 37%; $\rho_{diagram}$ = 63%) [b]Note that variance on individual levels can also increase (e.g., Snijder & Bosker, 1994. [c]Regression coefficient is significant (Wald z-test) but likelihood-ratio test shows that the intercept-only model is not improved by adding this variable. Read Title will not be used in following analysis.

**$p < .05$, χ^2 likelihood-ratio test with df = 1 **$p < .01$, χ^2 likelihood-ratio test with df = 1 ***$p < .001$, χ^2 likelihood-ratio test with df = 1.*

From the eye-tracking measures, the variables Fixation Time Main (β_1 = .51, p < .001) and Transitions Main (β_1 = .29, p = .01) appeared to be significant predictors of comprehension score. Adding Fixation Time Main as a fixed effect to the intercept-only model explained 64.8% of the variance on student and 6.7% on diagram level, χ^2 (1, N = 87) = 23.23, p < .001. The intraclass correlation of Fixation Time Main was 55.5% at the student and 44.5% at the diagram level. The variable Transitions Main explained 26.7% of the variance on the student and .8% on the diagram level compared to the intercept-only model, χ^2 (1, N = 87) = 6.68, p < .01. The intraclass correlation of Transitions Main was 54.1% at the student and 45.9% at the diagram level.

3.2 Relationships between learning activities that significantly predict comprehension

Table 4.4 presents the correlation matrix of learning activities that significantly correlated to comprehension score. Note that all significant correlations were positive. Meaning Arrow was significantly correlated with Orientate Legend (r = .50), Self-Questioning (r = .28), and Read Level (r = .38). Furthermore, Fixation Time Main was significantly correlated with Orientate Legend (r = .36), Orientate Prior (r = .27), Meaning Arrow (r = .58), Self-Questioning (r = .48), and Read Level (r = .47). Transitions Main was significantly correlated with Orientate Prior (r = .43). Finally, Fixation Time Main was correlated to Transitions Main (r = .63).

Table 4.4. Correlation matrix of students' learning activities that are related to the comprehension score

	1	2	3	4	5	6
1. Orientate Legend						
2. Orientate Prior	-.03					
3. Meaning Arrow	.50**	-.04				
4. Self-Questioning	.01	.01	.28*			
5. Read Level	.24	.07	.38**	.17		
6. Fixation Time Main	.36*	.27*	.58***	.48***	.47***	
7. Transitions Main	.28	.43**	.17	.24	.23	.63***

*p < .05 **p < .01 ***p < .001

4. DISCUSSION

4.1 Learning activities related to comprehension of process diagrams

4.1.1 Orientation phase

We found that the frequency of using the legend and activating prior knowledge in the orientation phase was positively related to students' level of comprehension. The finding that an elaborate orientation phase is related to increased performance is in line with previous research (Pressley & Afflerbach, 1995; Van Gog, Paas, & Merriënboer, 2005).

The variable Orientate Legend, obtained from verbal protocols of learning task 2, explained a medium amount (Cohen, 1988) of student-level variance in comprehension score. It might be plausible to suggest that orientating on the legend at an early stage in the learning process reduces workload in the main phase enabling resources to be used for learning.

Activating prior knowledge in the orientation phase explained a small amount of student-level and diagram-level variance of comprehension score. As mentioned in the theoretical framework, the role of prior knowledge is important for the interpretation of scientific graphical representations (e.g., Cook, 2006). As indicated by the intraclass correlation, a fair number of students who activate prior knowledge in the orientation phase do this for both learning tasks.

4.1.2 Main phase

Four predictors were significant for comprehension in the main phase: giving meaning to a process arrow, self-questioning, reading the title, and reading the organizational levels. Reading the title will not be discussed as adding this variable did not improve the intercept-only model (i.e., as indicated by the likelihood-ratio test).

No less than 80% of student level variance in comprehension score was explained by the number of process arrows a student gave meaning to. Giving meaning to the process arrows seems like a basic activity. When studying process diagrams, arrows are the key signifiers. However, within our group—i.e., students with an intermediate level of expertise—it is hard to imagine why some students employ this behavior so minimally. Intraclass correlations suggest that the number of process arrows students give meaning to is mostly between-student behavior, i.e., within-student between-diagram behavior is consistent across learning tasks.

Self-questioning explained a relatively small amount of student-level and diagram-level variance in comprehension score. The intraclass correlations suggest that a fair amount of variance of self-questioning can be considered as within-student behavior. The significant effect of self-questioning on comprehension score is in line with previous studies (e.g., King, 1989).

Reading the organization levels explained a moderate amount of student-level variance and a small amount of diagram-level variance. Coping with the different levels of organization of biological processes is an important and difficult aspect of

learning Biology (Knippels, 2002; Verhoef, 2003). We suggest that students who actively integrate the hierarchical level of organization in which the process takes place into their visual representation, are likely to understand more of the context, which will, in turn, facilitate understanding. Intraclass correlations suggest that reading the levels of organization is mostly within-student between-diagram behavior i.e., student behavior differs across learning tasks. The labels with the level of organization seemed to be equally informative in the diagrams (i.e., unlike the legends in diagram 2 and 3), suggesting that this is an effect of differences in size, position, or saliency (Cook et al., 2008; Lowe, 1999).

4.1.3 Eye-tracking measures

The eye-tracking measures Fixation Time Main and Transitions Main were significant predictors of comprehension score; Fixation Time Main explained 65% of the between-student variance of comprehension score (cf., Mason et al., 2013; Schwonke et al., 2009). These findings support previous findings where fixation time is associated with ongoing cognitive processes (Hannus & Hyönä, 1999; Just & Carpenter, 1976; She & Chen, 2009) and where transitions are associated with integration processes (Mason et al., 2013; Schwonke et al., 2009). Fixation Time Main might negatively influence comprehension of unsuccessful students' because they are satisfied with their learning progress in an early phase as they are not able to detect gaps in their knowledge (Chi, 2000). Intraclass correlations of Fixation Time Main and Transitions Main (i.e., measured on all three learning tasks) suggest that a fair amount of variance of both variables can be attributed to between-student differences. The latter suggests that students behave consistently across the learning tasks with regard to ongoing cognitive and integrative processes..

4.2 Relationships between learning activities that significantly predict comprehension score

4.2.1 Relationships between variables from verbal protocols.

The significant correlations between on the one hand Meaning Arrow and on the other hand Orientate Legend, Self-Questioning, Read Level, and Fixation Time Main suggest that giving meaning to process arrows indicates a strategic and in-depth approach for interpreting process diagrams, i.e., planned behavior (Pressley & Afflerbach, 1995). The significant relations between Self-Questioning, Meaning Arrow, and comprehension score, may reflect a sequential path of reasoning found throughout the collected verbal protocols. When students gave meaning to one or several process arrows they often uttered questions that focused on the meaning of these arrows. Attempts to answer these questions might lead to inferences and thereby to a higher comprehension score.

4.2.2 Relationships between variables from verbal protocols and eye-tracking measures

The relationship between fixation time and learning activities supports the idea that increased fixation time indicates the occurrence of in-depth cognitive processing. The high correlation between Fixation Time Main and Transitions Main suggests that students with longer total fixation time in the main area revisited more different parts in the main area.

4.3 Limitations

The target group of the present study might be considered a limitation. The learning activities employed by this target group might be specific for students with their level of expertise and experience with studying process diagram. For this, we recommend caution in extending the results to other target groups such as lower grade students and university students. However, the choice for this target group in the present study was deliberate. These students might soon be faced with even more difficult process diagrams in higher education (Dos Santos & Galembeck, 2015). They might face some serious challenges in their study and future careers if their skills for learning from process diagrams are insufficient. Another limitation might be the specific focus on biological process diagrams. Despite this focus, we expect that our results extend to learning from process diagrams in nearby scientific domains (e.g., chemistry, physics) and even to some more distant domains like geography.

4.4 Conclusion and implications for education

In sum, the present study adds a fined-grained analysis of learning from a specific diagram type, i.e., the process diagram, to previous research on learning from graphical representations. We found various learning activities that distinguished more and less successful students while learning from process diagrams.

Some distinct findings were that successful students were more likely to employ learning activities such as using the legend in the orientation phase; in the main phase successful students more often give meaning to process arrows and read the organizational levels. The latter two findings were not found in earlier research. The importance of employing orientating activities is in line with previous research (Pressley & Afflerbach, 1995; Pressley, 2000, Van Gog, Paas, & Merriënboer, 2005). However, the present study adds to existing insights that applying a specific diagram activity, i.e., using the legend in the orientation phase, is important as well.

In accordance with previous studies, we also found that successful students were more likely to employ learning activities such as activating prior knowledge (Pressley & Afflerbach, 1995; Pressley, 2000), self-questioning (e.g., Azevedo & Cromley, 2004) and spending more time in the main area of the process diagram (cf., Mason et al., 2013; Schwonke et al., 2009; She & Chen, 2009).

We also found that students employed successful learning activities consistently across learning tasks: the learning approaches seemed to be stable. The present study thereby contributes to research that focused on the generalization of (meta)cognitive activities across tasks or domains (e.g., Meijer et al., 2006; Veenman, 2012). Furthermore, we conclude that successful students use a more coherent approach of interrelated learning activities for comprehending process diagrams than unsuccessful students.

The present study provides relevant insights into the topics of a training that specifically focusses on learning from process diagrams. The training could teach students when and how to employ cognitive and metacognitive strategies found to be characteristic for successful students in the present study.

CHAPTER 5

EFFECT OF A MULTIPLE-STRATEGY TRAINING WITH A STEPWISE WORKING-ROUTINE ON LEARNING FROM PROCESS DIAGRAMS

The present study evaluated the effect of multiple-strategy training on learning from process diagrams. The training focussed on a stepwise working-routine that included when and where to employ cognitive and metacognitive learning strategies and on affective strategies to invest effort in the implementation of this stepwise working-routine. The study followed an experimental pretest-posttest design. Students (N = 180) were randomly assigned to the experimental or the control condition. Structured equation modeling was applied to examine the direct and indirect effects, through invested mental effort and perceived task difficulty, on learning from process diagrams. We observed an indirect positive significant effect of multiple-strategy training, through invested mental effort, on learning from process diagrams compared to the control group. We observed no significant direct effect of multiple-strategy training on learning from process diagrams compared to the control group.

1. INTRODUCTION

The adage "A diagram is (sometimes) worth ten thousand words" from Larkin and Simon (1987) implies that diagrams have the potential to communicate efficiently. There is, however, ample evidence that students have difficulties with learning from diagrams (e.g., Chittleborough & Treagust, 2008; Cromley et al., 2013b; Kriz & Hegarty, 2007; Schönborn, Anderson, & Grayson, 2002). Although there is vast amount of research on conditions for learning from diagrams (e.g., Canham & Hegarty, 2010; Kriz & Hegarty, 2007) limited empirical research has been done on the effects of training students to learn from diagrams. The few exceptions are studies by Cromley et al. (2013a; 2013b). Students' difficulties in learning from diagrams might be explained by insufficient learning strategies and an incoherent and unstable task execution (Kragten, Admiraal, & Rijlaarsdam, in press). The purpose of this study was to evaluate whether learning from diagrams could be facilitated by a multiple-strategy training (Alexander, Graham, & Harris, 1998; Dole, Nokes, & Drits, 2009), including a stepwise working-routine (cf. Souvignier & Mokhlesgerami, 2006).

1.1 Learning strategies

Learning strategies are mental routines or procedures that facilitate learning performance and are potentially conscious and controllable activities (Alexander et al., 1998; Dole et al., 2009; Pressley & Harris, 2006). In their meta-analysis, Donker, De Boer, Kostons, Dignath Van Ewijk, & Van Der Werf (2014) examined the effect of

learning strategy instruction on academic performance, with a focus on improving self-regulated learning (e.g., Boekaerts, 1997; Schunk & Zimmerman, 1997). The authors included studies that aimed at improving cognitive, metacognitive and management (i.e., management of effort, peers and others, and the environment) strategic skills, as well as studies that aimed at improving motivation and metacognitive knowledge. Strategies that focused on improving metacognitive knowledge (i.e., when and where should the strategy be used), and on elaboration, planning, and task value, proved to be most effective. In an earlier meta-analysis, Hattie, Biggs, and Purdie (1996) concluded that best results are obtained when strategy training is metacognitive.

Studies that focused on learning from diagrams identified inference and (activating) prior knowledge as key elements (Canham & Hegarty, 2010; Cook, Carter, & Wiebe, 2008; Cromley, Snyder-Hogan, & Luciw-Dubas, 2009; Cromley et al., 2013a; Kriz & Hegarty, 2007). Prior knowledge affects learning from diagrams because it influences how information is selected and encoded (Canham & Hegarty, 2010; Kriz & Hegarty, 2007); Novices tend to focus on surface features whereas experts focus on thematically relevant information (Cook et al., 2008; Lowe, 1999). Furthermore, learners with more prior knowledge are likely to integrate the encoded graphical representation into a correct mental model (Kriz & Hegarty, 2007; Schnotz & Bannert, 2003). Inferential processes play also a critical role in learning from diagrams (Kriz & Hegarty, 2007). Cromley et al. (2013a) compared the effectiveness of three learning conditions on diagram comprehension. Although students in all conditions improved overall comprehension of diagrams, only students in the two learning conditions that fostered inferential processes improved their scores also on inferential comprehension questions.

1.2 Self-regulated learning

Self-regulated learners effectively employ cognitive and metacognitive learning strategies to achieve their goals. For academic achievement, however, learners also must be motivated to invest mental effort in the employment of learning strategies (Boekaerts, 1997; Paas, Tuovinen, Van Merriënboer, & Darabi, 2005; Pintrich & De Groot, 1990). Hence, a positive relationship likely exists between learner involvement, invested mental effort, and performance.

Azevedo and Cromley (2004) evaluated the effect of training in self-regulated learning on students' learning in a hypermedia learning environment (i.e., nonlinear information presented in multiple modalities). Students in the experimental group received, individually, a 30 minute explanation about several key concepts of self-regulated learning, e.g., planning, monitoring, and strategies; students in the control group did not receive this information. Next, during 45 minutes students studied the human circulatory system in a hypermedia learning environment. Students were instructed to think aloud during the latter phase. Results indicated that the experimental group, compared to the control group, improved their mental models more (students were required to write an essay and draw a diagram) and performed better

on a labelling task (i.e., naming parts of the heart). Verbal protocol data showed that students in the experimental group employed more strategies that indicated self-regulated learning than the control group.

Souvignier & Mokhlesgerami (2006) evaluated the effect of three training programs that aimed at fostering reading strategies and reading comprehension. The training programs of 20 lessons of 45 minutes each were developed within the self-regulated learning framework. They differed with respect to which aspects of strategy instruction were taught: strategy knowledge, cognitive self-regulation, or motivational self-regulation. The training program that incorporated all three aspects started with four lessons on motivational self-regulation (e.g., setting realistic goals), continued with a core strategy program of 12 lessons, and ended with four lessons on cognitive self-regulation. The core strategy program was designed using a 2 (cognitive and metacognitive strategies) x 2 (organization and elaboration) strategy scheme to cover all aspects of reading strategies. In the lessons on cognitive self-regulation, students were trained how to use strategies in a flexible way and offered a working-routine (a “reading plan”) to structure the process of self-regulation during the reading process. Results from the retention test were most promising for the students in the programs that incorporated all three aspects. Only this latter condition outperformed the control group, i.e., students in regular classes, on understanding and application of reading strategies.

1.3 Strategy instruction models

Dole et al. (2009) reported studies distinguishing single-strategy instruction and multiple-strategy instruction models. The present study adapts a multiple-strategy instruction model. Popular multiple-strategy instruction models are reciprocal teaching (Palincsar & Brown, 1984), direct explanation (Duffy et al., 1987), self-regulated strategy development (Harris & Graham, 1996), and transactional strategy instruction (Pressley et al., 1992). After the original studies, these models further developed, which led to some crossover between the models. For instance, in the original work of Palincsar & Brown (1984) on reciprocal teaching, strategies were not explicitly explained to students. However, later studies included explicit teaching of the strategies at the beginning of the training (Rosenshine & Meister, 1994), which is in line with the direct explanation model (Duffy et al., 1987).

The delivery method of most multiple-strategy instruction models includes four steps: 1) explaining the strategies , 2) modeling the strategies, 3) guided practice with feedback while learning to use the strategies, and 4) learners performing the task independently. These steps reflect the social cognitive theoretical model of becoming a self-regulated learner (Schunk & Zimmerman, 1997; Zito, Adkins, Gavins, Harris, & Graham, 2007), in which self-regulated learning is developed through four phases: observational, imitative, self-controlled, and self-regulated. Below we explain the steps of multiple-strategy instruction models in more detail.

With respect to explaining the strategy it is clear that this should be direct and explicit (Pressley, Borkwski, & Schneider, 1989). A learner must be made aware of

what the strategy is about and how to use it. It should also be clear where and when the strategy applies.

With respect to modeling, positive effects have been found on self-efficacy, motivation and academic achievement (Schunk & Zimmerman, 1997). Traditionally, thoughts and motives are verbalized, e.g., by an expert, while performing a task to demonstrate behavior. A recent development in modeling is using eye movement to guide learners' attention during instruction. Jarodzka et al. (2013) used a domain-expert's eye movements and verbal explanation to instruct learners how to perform a perceptual task, i.e., classifying fish locomotion. In this study, there were two experimental conditions of superimposing eye movements on a video, i.e., as a spotlight or as a dot. Students in the spotlight condition (i.e., other information was blurred) increased visual search efficiency. Students in the dot condition (i.e., other information was not blurred) increased interpretation performance. Mason, Pluchino, and Tornatora, (2015) studied the effect of using eye movement modeling on integrative processing and learning from an illustrated text. Lower secondary students in the experimental condition were shown eye movements (without verbal explanation) of a graduate student who modeled how to process an illustrated text. Eye movements modeling had a positive effect on the amount of integrative processing of the illustrated text and—as a result—on recall and transfer posttests, compared to the control condition.

Strategies have to be practiced to become automatized (Pressley et al., 1989). Guided practice can take on a variety of forms but usually consists of teacher or peer feedback on strategy use and behavior and scaffolding. Typically there is a gradual shift in responsibility for the employment of the desired behavior from the teacher to the learner.

Explaining, modeling and guided practice are traditionally delivered by the teacher or through teacher-student interaction. However, computer-based multimedia learning environments are becoming an acceptable alternative to conventional lessons (Lowe & Schnotz, 2008; Kay, 2012). It is, on the other hand, known that students have difficulties with self-regulating their learning with computer-based multimedia environment (Winter, Greene, & Costich, 2008). Delen, Liew, and Willson (2014) examined whether self-regulated learning can be supported by offering learners an enhanced video learning environment. This environment (i.e., the experimental condition) differed from the common video environment (i.e., the control condition) in the function to note-taking, seeking supplemental sources and self-evaluation through practice questions. Students of the experimental condition outperformed students of the control condition on a recall test. The authors attributed the latter result to the effect of the enhanced learning environment on students' engagement and activated self-regulated learning behavior.

1.4 Focus of the present study and research questions

We designed a multiple-strategy training based on the self-regulated strategy development model (Harris & Graham, 1996) to facilitate learning from process dia-

grams. The training specifically focussed on learning from process diagrams; a distinct, but abundantly present, type of diagram that is used in modern science text books to explain processes like photosynthesis, immunology, and protein synthesis (e.g., Reece et al., 2010). Focussing on process diagrams allowed the development of a multiple-strategy training that included specific declarative (e.g., conventions) and procedural knowledge (e.g., where to start, how and when to use the legend). Previous research showed that both types of knowledge distinguished successful learners from less successful learners for this type of diagram (Kragten et al., in press). It has been demonstrated that students benefit from detailed instruction on a specific type of representation (e.g., Körner, 2005). Furthermore, the training was designed for a specific group of students, i.e., pre-university upper secondary school students with an intermediate level of expertise in science. These students are likely to pursue a career in university Science in which they are soon faced with even more challenging process diagrams (Dos Santos & Galembeck, 2015). Hence, facilitating students to learn from process diagrams is an important issue.

The present study's first research question is defined as: "What is the effect of multiple-strategy training on learning from process diagrams?" We expect learners in the multiple-strategy training condition to outperform the control group in learning results.

Two additional research questions are formulated. First, we assume that the training group increases their invested mental effort more than the control group, because the multiple-strategy training is expected to have a positive effect on motivation and/or self-efficacy and thereby increasing willingness to invest mental effort. We also assume that increased invested mental effort is positively related to learning from process diagrams. Therefore, the present study's second research question is defined as: "What is the indirect effect of multiple-strategy training, through invested mental effort, on learning from process diagrams?" We expect that the multiple-strategy training has a positive significant effect on learning results through invested mental effort.

Second, we assume that the training group perceives less task difficulty; the training offers them procedural knowledge and learning strategies that allows them to manage difficulties. An alternative hypothesis could be that perceived task difficulty is inherent to the content of the information presented by the diagram and that there is no effect of multiple-strategy training on perceived task difficulty. In any case, we assume that perceived task difficulty is negatively related to learning from process diagrams, i.e., the more difficult the task is perceived the less is learned. Robinson (2001), for instance, found that students' task performance (i.e., giving verbal directions using a map) was related to perceived task difficulty. Therefore, the present study's third research question is formulated as: "What is the indirect effect of multiple-strategy training, through perceived task difficulty, on learning from process diagrams?"

2. METHOD

2.1 Participants

Participants were 180 students (118 female) from 8 classes from 8 secondary schools in the northwest of the Netherlands. All students were in their last year of pre-university secondary education. They had chosen Biology as a major topic within their exam program with a study load of 480 hours during three years of upper secondary education. The students were in high-performing classes with a focus on Science (i.e., all students also had chosen Chemistry and Math and almost all had chosen Physics).

2.2 Strategy instruction learning environment

The present study follows a pretest-posttest design. A computer-based 'enhanced' multimedia learning environment was designed to deliver the instructions for the pretest, intervention, and posttest phase (see Figure 5.1). The multimedia environment is 'enhanced' because it contains interactive elements, e.g., built-in stops, tasks and feedback, and user controls (e.g., stop, play, pause, and replay). Students were also provided paper-and-pencil workbooks.

Figure 5.1. Screenshot of the computer-based enhanced multimedia learning environment (original was presented to the training group in Dutch).

This enhanced multimedia environment was designed following the general guidelines for the design of multimedia instructional material (Mayer & Moreno, 2002). Instruction was provided in multiple modalities (i.e., the modality principle)—verbal explanations, minimal text and (animated) illustrations—and were timely coordinated (i.e., temporal contiguity principle). The verbal explanations were in an informal style to enhance engagement (i.e., the personalization principle). The left area of Figure 5.1 continuously presented information that supported the verbal explanation. Parts of the information in the left area that were currently discussed were highlighted (i.e., the signalling principle). In the right area of the learning environment there was always a picture of the lecturer to simulate social presence. The right area also contained two text areas. The text area below the picture of the lecturer signalled the phase of the training, e.g., explaining the strategy; the text area above the lecturer contained keywords. The amount of text was kept to a minimum to avoid redundancy and split-attention effects.

A first version of the enhanced multimedia learning environment and the workbooks was tested in a pilot study with 12 students. The students provided feedback about the content and functionality of the environment and their suggestions were incorporated into a new version. This new version was tested by a high school biology teacher and suggestions were again incorporated into the final version.

2.3 Procedure

The pretest, intervention, and posttest were conducted in one single session in a classroom with computers in students' school. Students were randomly assigned to the experimental condition or the control condition when entering the classroom and were assigned a workplace. Each workplace had an USB stick with the multimedia learning environment, a headphone, general instructions, and the workbooks for the pretest, intervention (training or control tasks), and the posttest. Workplaces were separated by demountable walls to assure privacy. The general instructions were read aloud by the experimenter, i.e., the first author. Next, students' were instructed to start the multimedia learning environment. From that point on, students were led through the pretest, intervention, and posttest phase by instructions in the multimedia learning environment and the workbooks.

2.4 Pretest, intervention, and posttest phase

2.4.1 Pretest phase

The pretest phase consisted of a learning task followed by a test. The learning task was delivered by the multimedia learning environment. Students were presented the following instruction:

> "You will be presented a process diagram. Try to understand as much as you can. Study for about 5 minutes. When you are ready you can stop by pressing the spacebar. You will have to complete a test about what you have learned. The process diagram will not be available while you complete the test."

After instruction, students were presented a process diagram of the absorption and secretion of bile acid and cholesterol in the entero-hepatic cycle (Figure 5.2). Students were supposed to be familiar with parts and function of the entero-hepatic circulation; the basics of the entero-hepatic circulation (i.e., bile is produced by the liver, stored by the gallbladder, excreted to the intestines, and then reabsorbed) had been taught in previous years. Students also were supposed to be familiar with basic chemical reactions and transport processes (e.g., active transport, endocytosis). The detailed biochemical reactions and transport processes of the presented process diagram of the entero-hepatic cycle were not taught previously. Students' task was to learn this new information.

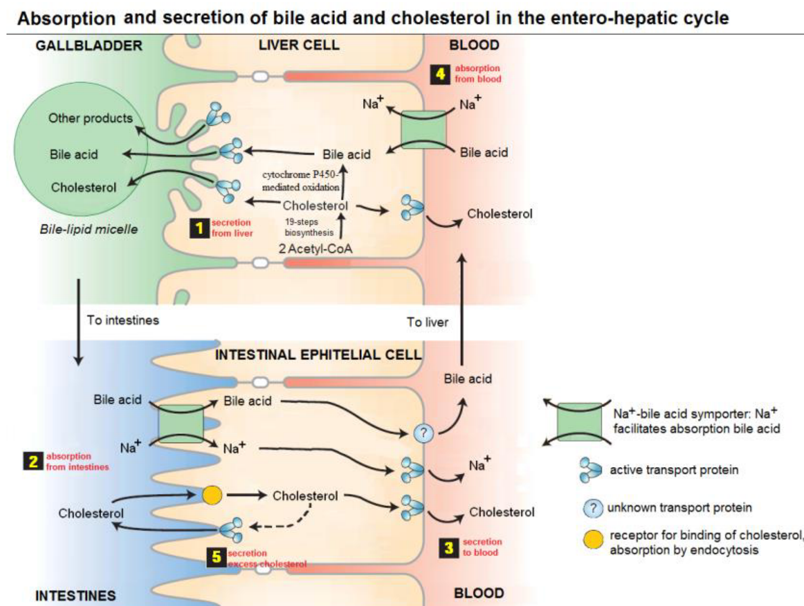


Figure 5.2. Process diagram of the pretest learning task (original diagram was presented to the students in Dutch).

When students finished learning, they pressed the spacebar and got the information how much time they spent on learning from the process diagram; students had to report this on the front page of the pretest workbook. Due to technical problems (e.g., the macro that kept time was not allowed to run by the particular system) and human errors (e.g., students forgot to report the time) we only obtained complete sets (pretest and posttest) of time on learning task data for 110 students.

Then students were instructed to open their pretest workbook. First, students had to rate their invested mental effort and perceived task difficulty on a 7-point rating scale. Self-rating scales on invested mental effort and perceived task difficulty have been used extensively in cognitive load studies (Van Gog & Paas, 2008; for per-

ceived task difficulty see also Robinson, 2001). Students' invested mental effort was assessed by asking: "How much effort did you invest in understanding the diagram?" with 1 as very little and 7 as very much. Perceived task difficulty was assessed by asking: "How difficult was it to comprehend the diagram?" with 1 as very easy and 7 as very difficult. Next, students had to finish the pretest; they were not allowed to adjust their ratings of invested mental effort and perceived task difficulty when they started the pretest. The pretest contained eight multiple-choice items that tested students' knowledge (e.g., "Which transport mechanism is responsible for the absorption of cholesterol in the small intestine?") and comprehension (e.g., "Which statement correctly describes the operation of the Na+/bile salt symporter?") of the process diagrams. Each correct item was rewarded with one point.

2.4.2 *Intervention phase*

Students in the experimental condition were presented a multiple-strategy training based on cognitive strategy instruction (Harris & Graham, 1996). The training aimed at teaching students a stepwise working-routine that included when and where to employ cognitive and metacognitive learning strategies. The training consisted of five phases: 1) emphasizing that diagrams are important and useful, 2) explaining the strategic approach, 3) providing a model of the strategic approach, and 4) practicing with the strategic approach, and 5) providing feedback on practicing.

Phase 1. Diagrams are important and useful. This phase started by presenting students a familiar process diagram from their Biology textbook and introducing the subject of the training (i.e., learning from process diagrams). The latter aimed at motivating students for using the strategic approach they are about to learn by showing them a concrete example of a task they often have to perform.

Before students are willing to use a strategy, they have to be convinced that the strategy is effortful, i.e., extra invested mental effort enhances performance, and essential, i.e., one cannot reach competence in a domain without the strategy (Alexander et al., 1998). The continuing information presented in this phase aimed at the latter prerequisites. For this, we informed students that research indicated that diagrams are effective learning tools by briefly explaining, in simple wordings, the multimedia theory (Mayer, 2005), the dual coding principle (Paivio, 1986), and the visual argument (Larkin & Simon, 1987). Furthermore, we stressed that the ability to interpret diagrams is essential in higher education and their future careers. Finally, we informed students that the strategy they are about to learn is based on insights from previous research.

Phase 2. Explaining the strategic approach. In this phase, students were directly informed about the strategic approach for learning from process diagram. The strategic approach was based on findings from research that focussed on strategic text reading (e.g., Pressley & Harris, 2006), inferring (Kriz & Hegarty, 2007), metacognition (e.g., Meijer, Veenman, & Van Hout-Wolters, 2006), and the interpretation of process diagrams (Kragten et al., in press). The working routine of the strategic ap-

proach consisted of four major steps: 1) orientating, 2) finding a starting position, 3) elaborating, and 4) summarizing.

The first step of the working routine (i.e., orientating) consisted of employing activities regarding the meta-information in the process diagram (i.e., title, labels with organisational information, legend). Students were instructed to read the title first and to activate relevant prior knowledge. Next, students were instructed to find and read all labels with information about the organisational level and again to activate relevant prior knowledge. For instance, when a label indicated “liver cell” students might already have some knowledge about processes that take place in that particular cell type. Finally, students had to read the legend items, activate prior knowledge, and find the legend items in the main area of the process diagram. The latter activity was found in a previous study to distinct successful learners from less successful learners (Kragten et al., in press).

The second step was finding a starting position in the main content area of the diagram. In some process diagrams the starting position is clearly indicated. In other diagrams the starting position is not clear and students had to decide where to start reading. For the latter case, students were taught to find a component from where they can go through a large part of the process without going against the flow (i.e., against the direction of the arrows) of the diagram.

The third step of the working routine (i.e., elaborating) consisted of activities that facilitated the students to comprehend the information presented in the main content area of the process diagram. Students were taught to give meaning to every arrow in the content area of the diagram. The multimedia learning environment contained a brief instruction on the most common conventions of arrows in process diagrams (Kragten et al., 2013a; 2015). Knowledge of conventions (Cromley et al., 2013b) and arrows that give meaning to processes (Kragten et al., in press) are important determinants for success in learning from (process) diagrams. Students were also taught to continuously activate prior knowledge, ask questions and make inferences as they proceeded through the main content area of the diagram.

The fourth step of the working routine focussed on the importance summarizing the process as a final stage in learning from a process diagram. Students were instructed to summarize the entire content of the diagram in a few sentences.

Phase 3. Example of the strategic approach. In phase 3, a video was shown where the lecturer modeled the strategic approach (Figure 5.3). In this video, the lecturer was thinking aloud while he was using the strategic approach to learn from a process diagram. The video showed a spotlight to indicate the lecturer’s eye movements. The spotlight was used to guide students’ attention to the part of the diagram that was currently discussed. As discussed earlier, using eye movement to guide learners’ attention facilitates search activities and fosters learning processes (Jarodzka et al., 2013; Mason et al., 2015). The part of the diagram that was not highlighted was slightly darkened but not blurred so that students remained visual access to the entire process diagram. During the video, students had to note the time when the expert passed to a next step of working routine (in Figure 5.3 at 3:47 the lecturer starts with step 3, i.e., elaboration). The latter task aimed at rehearsing the

stepwise working-routine but also at keeping students engaged (Szpunar, Jing, & Schacter, 2014).

Phase 4 and 5. Practicing with the strategic approach and feedback. In phase 4, students practiced using the four steps of the strategy with a new diagram, i.e., the production of stomach acid. The workbook contained tasks to practice each step. For instance, to practice the first step, students were instructed to read the title and report relevant prior knowledge in their workbook. Students were provided feedback by the multimedia learning environment after the tasks of each step; students could request this feedback at their own pace by pressing the spacebar when they finished the tasks of a step.

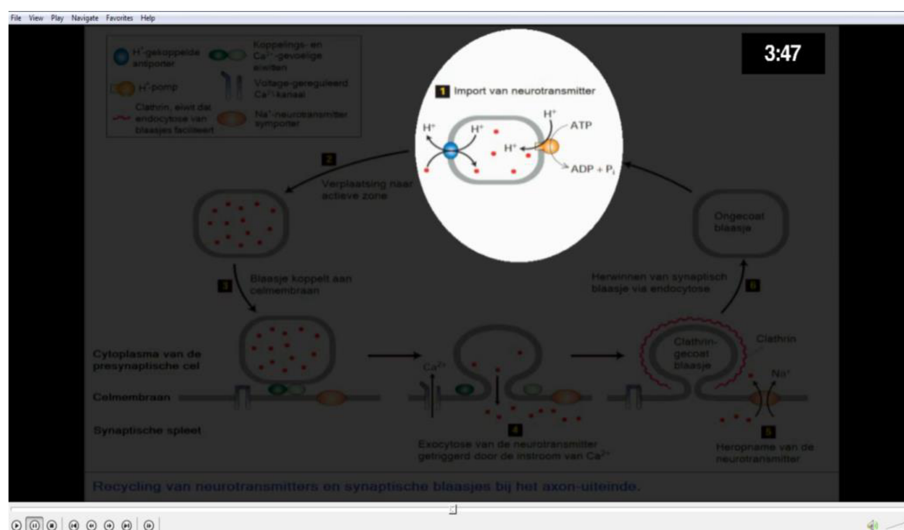


Figure 5.3. Screenshot of the modeling video. The diagram presents the process of the recycling of neurotransmitters and synaptic vesicles at the axon terminal (original version).

At the end of the training, students were instructed to use the strategy they just learned during the posttest phase.

Students in the control condition were presented five process diagrams by the multimedia learning environment. Students had to make a total of nine tasks in their workbook about these process diagrams; the tasks tested if student were able to comprehend the process diagrams. When students finished the tasks, they received feedback (i.e., the correct answers and some further explanations) by the multimedia environment; students in the control condition did not receive any strategic information. Two process diagrams in the control condition were identical to diagrams presented to the students in the experimental condition. The control condition contained three extra process diagrams (with tasks). The latter was based on evaluation of the pilot study and meant to assure that students in both conditions spent an approximate equal amount of time (± 75 min) in the intervention phase.

2.4.3 Posttest phase.

The posttest phase, like the pretest phase, consisted of a learning task followed by a test. The posttest phase started with the same instruction as the pretest phase. After the instruction, students were presented a process diagram that depicted the two main growth modes (autotrophic and heterotrophic) of *Ralstonia eutrophia* (adapted from Pohlmann et al., 2006). Students' task was to study this information. Again, like in the pretest phase, students had to report how much time they spent on learning from the process diagram.

After the learning task, students first had to rate their invested mental effort and perceived task difficulty before they could start with the posttest. The posttest consisted of four subtests. The first two subtests of the posttest consisted of reconstructing the processes of the two main growth modes of *R. eutrophia* by redrawing the components and the arrows. Correct items consisted of the combination of a chemical component and the associated arrow leading away from the chemical component in the right direction (e.g., an arrow depicting passive transport of carbon dioxide into the cell at the autotrophic mode). The subtests that required students to reconstruct the processes of the autotrophic and heterotrophic mode had 13 and 11 scoring items, respectively; each correct item was rewarded with one point. Asking learners to reconstruct a process by drawing is a common procedure to measure students' understanding on a specific topic (Quillin & Thomas, 2015; She & Chen, 2009). The third subtest of the posttest consisted of drawing and naming sub-processes (e.g., the Calvin cycle) and facilitative enzymes (e.g., ATP-synthases). The latter subtest had 5 scoring items. Correct items consisted of correctly naming and drawing a sub-process or an enzyme; each correct item was rewarded with one point. The fourth subtest of the posttest consisted of 10 scoring items; three multiple-choice questions and two open ended questions (answers were scored on 7 elements). Again, each correct item of the latter subtest was rewarded with one point.

2.5 Data analysis

First, we analysed internal reliability, indicated by KR-20, of the pretest and the subtests of the posttest. Once reliable scales were established for the subtests of the posttest, we further examined if these subtests could be presented by a single latent variable within a structured equation model (Figure 4), i.e., the posttest score. Next, descriptive statistics of measures of the pretest and posttest phase were calculated. Furthermore, we conducted a one-way ANCOVA to determine if there was a statistically significant difference between conditions on time on learning task in the posttest phase, controlling for time on learning task in the pretest. We did not include the latter variables into the structured equation model because of missing data as discussed earlier. The outcome of the ANCOVA was used as supportive evidence.

To answer the research questions, a structured equation model (Figure 4) with invested mental effort and perceived task difficulty as mediators, i.e., a multiple mediator model (Preacher & Hayes, 2008), was constructed to evaluate the zero-order (i.e., the direct effect of condition on posttest score without the indirect effects), di-

rect and indirect effect of condition on posttest score. The posttest score was controlled for by the items of the pretest. The effect of experimental condition on invested mental effort in the posttest phase was controlled for invested mental effort in the pretest phase. This allowed us to evaluate the effect of experimental condition on mental effort gain and the effect of mental effort gain on posttest score. The same procedure applied to perceived task difficulty. When both effects of an indirect path were significant, we used the product-of-coefficients approach (Sobel, 1982) to test if the specific indirect effect (Preacher & Hayes, 2008) was significant.

3. RESULTS

3.1 Descriptive statistics

The eight items of the pretest did not converge into a single scale so it was decided to include them separately (*prei1* till *prei8* in Figure 5.4) to the structured equation model to control for posttest score (i.e., latent variable *Postscore* in Figure 5.4). The subtests of the posttest did converge into single scales. Internal reliability, as indicated by KR-20, of the subtests that required students to reconstruct processes of the autotrophic mode and heterotrophic mode was .84 and .82, respectively. Internal reliability, as indicated by KR-20, of the subtest that required students to place and name sub-processes and enzymes was .62. The latter subtest's internal reliability could not be improved by removing items. The subtest that contained multiple choice questions and open ended questions had 10 scoring items and internal reliability, as indicated by KR-20, was .72 after deleting three scoring items. The descriptive statistics of the measures of the pretest and the posttest phase are presented in Table 5.1.

We examined whether the subtest scores of the posttest could be represented by a single latent variable, i.e., the posttest score (i.e., latent variable *Postscore* in Figure 5.4). First, we adjusted the regression weights of *Postscore* for internal reliability of the subtests; the more reliable a subtest is the more it contributes to latent variable *Postscore*.

The correlations (i.e., the standardized regression weights after adjusting for internal reliability) between *Postscore* and scores for reconstructing the autotrophic mode (*postreconauto*), reconstructing the heterotrophic mode (*postreconhetero*), drawing and naming the sub-processes and enzymes (*postenzsub*), and questions (*postquestions*) are $r = .93$, $r = .91$, $r = .79$, and $r = .59$, respectively, and all are significant with $p < .001$. Because all correlations are high we concluded that one single latent variable represents the subtest scores adequately.

There is a significant difference between conditions on time on learning task in the posttest phase controlling for time on learning task in the pretest ($F(2,107) = 10.74$, $p = .001$). This means that students in the experimental condition spent relatively more time on the learning task of the posttest phase, compared to students in the control condition.

Table 5.1. Descriptive statistics of the measures of the pretest and the posttest phase

Phase	Control group					Training group				
Variable (*short term*)	Obs	*M*	*SD*	Min	Max	Obs	*M*	*SD*	Min	Max
Pretest phase										
Task difficulty (*pretd*)	88	4.36	1.25	2	7	92	4.48	1.13	2	7
Mental effort (*preme*)	88	4.68	1.20	2	7	92	4.88	1.18	2	7
Pretest										
Item 1 (*prei1*)	88	.56	.50	0	1	92	.42	.50	0	1
Item 2 (*prei2*)	88	.55	.50	0	1	92	.55	.50	0	1
Item 3 (*prei3*)	88	.44	.50	0	1	92	.37	.49	0	1
Item 4 (*prei4*)	88	.59	.49	0	1	92	.58	.50	0	1
Item 5 (*prei5*)	88	.42	.50	0	1	92	.48	.50	0	1
Item 6 (*prei6*)	88	.82	.39	0	1	92	.64	.48	0	1
Item 7 (*prei7*)	88	.53	.50	0	1	92	.42	.50	0	1
Item 8 (*prei8*)	88	.24	.43	0	1	92	.17	.38	0	1
Learning time (s)	71	190.35	96.34	36	483	72	170.79	77.04	51	409

Table 5.1. Continued

Posttest phase										
Task difficulty (*postd*)	88	3.84	1.34	1	7	92	4.03	1.39	2	7
Mental effort (*postme*)	88	4.61	1.45	1	7	92	5.25	1.16	2	7
Posttest										
Reconstruct autotrophic mode (*postreconauto*)[a]	88	4.92	3.50	0	13	91	5.40	3.79	0	13
Reconstruct heterotrophic mode (*postreconhete*)[b]	88	4.00	3.01	0	11	91	4.23	3.19	0	11
Draw/name enzymes/sub-processes (*postenzsub*)[c]	88	0.80	1.06	0	5	91	0.89	1.13	0	4
Questions (*postquestion*)[d]	88	2.47	1.77	0	7	92	2.92	2.12	0	7
Learning time (s)	55	217.11	101.56	94	678	63	268.13	103.56	88	620

Note. Obs = observations; Min = minimal score of students; Max = maximal score of students
[a]*KR-20 = .84;* [b]*KR-20 = .82;* [c]*KR-20 = .62;* [d]*KR-20 = .72 after deleting three items*

3.2 Effect of multiple-strategy training on learning from process diagrams

Figure 5.4 presents the unstandardized results of the structured equation model. Posttest score (*Postscore*) was controlled for by items of the pretest; items *prei1* (*B* = 1.08, *p* = .005), *prei5* (*B* = 0.79, *p* = .050), and *prei7* (*B* = -0.139, *p* < .001) significantly predict posttest score (*Postscore*). There is no significant zero-order effect (not shown in Figure 5.4) of condition (*cond*) on posttest score (*Postscore*; *B* = 0.34, *p* > .05) and no significant direct effect of condition (*cond*) on posttest score (*Postscore*; *B* = 0.25, *p* > .05). However, the positive regression coefficient of the zero-order effect might suggest there is a trend in the direction of the research hypothesis.

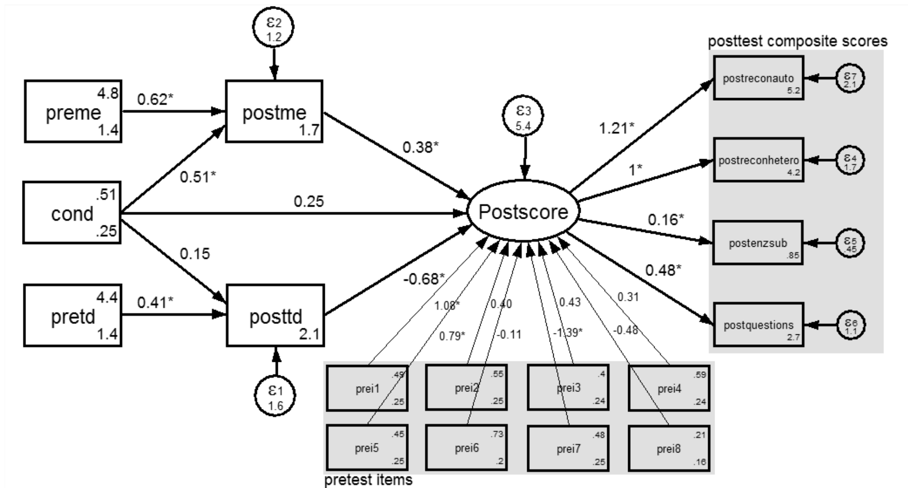


Figure 5.4. Structured equation model.
Note. Unstandardized regression weights are reported. The unstandardized regression weight between Postscore and postreconhetero is constrained to 1. Means of measured extraneous variables are displayed in the upper right corner and variance is displayed in the lower right corner. Intercepts of endogenous measured variables are displayed in the lower right corner. Error variances of errors are reported inside the circle below the naming label.
cond = condition (dummy variable, experimental condition = 1); Postscore; posttest score; preme = pretest mental effort score; postme = posttest mental effort score; pretd = pretest task difficulty score; posttd = posttest task difficulty score; postreconauto = posttest reconstruct autotrophic mode score; postreconhetero = posttest reconstruct heterotrophic mode score; postenzsub = draw and name enzymes and sub-processes score; postquestions = posttest questions score; prei1 to prei8 = pretest items score.
**p < .05.*

3.3 Indirect effect of invested mental effort on learning from process diagrams

There is a significant positive effect of condition (*cond*) on posttest invested mental effort score (*postme*; $B = 0.51$, $p = .002$) controlling for pretest invested mental effort score (*preme*). There is also a significant positive effect of posttest invested mental effort score (*postme*) on posttest score (*Postscore*; $B = 0.38$, $p = .009$). The indirect positive effect of condition (*cond*) on posttest score (*Postscore*) through posttest invested mental effort score (*postme*) is significant ($B = 0.19$, Sobel test, $z = 2.00$, $p = .045$).

3.4 Indirect effect of perceived task difficulty on learning from process diagrams

There is no effect of condition (*cond*) on posttest perceived task difficulty (*posttd*; $B = 0.15$, $p > .05$) controlling for pretest perceived task difficulty (*prettd*). There is a significant negative effect of posttest perceived task difficulty (*posttd*) on posttest score (*Postscore*; $B = -0.68$, $p < .001$).

4. DISCUSSION

4.1 Effect of multiple-strategy training on learning from process diagrams

We did not observe a zero-order nor a direct effect of multiple-strategy training on learning from process diagrams. A significant zero-order effect, however, is not necessary for an indirect effect (Shrout & Bolger, 2002; Zhao, Lynch, & Chen, 2010); it is possible that there are mediators and suppressors that counterbalance each other (Preacher & Hayes, 2008). The structured equation model of the present study does not contain a suppressor but there are several possible explanations for the total effect to be non-significant. For instance, for some students the training might have increased their lack of confidence in learning from process diagrams; this might be because the training confirmed that they are not successful in learning from process diagrams. The increased lack of confidence could have a negative effect on performance on the posttest and thereby suppress the total effect.

One might consider that not finding a direct effect is due to the design of the multiple-strategy training. An obvious design issue might be the intensity of the multiple-strategy training. The training consisted of one single session of approximately 75 minutes and this might not be enough to have a profound effect on students' ability to employ learning strategies. Interventions aimed at improving strategies are usually much larger with programs consisting of multiple lessons up to programs that last an entire year (Rosenshine & Meister, 1994; Souvignier & Mokhlesgerami, 2006). However, we were dedicated to develop a relatively small training because it would provide a more realistic opportunity for teachers to use the training in classroom. We expected a short but intensive training to be effective because the target group consists of high-performing students with a substantial amount of prior knowledge and experience in learning in Science. Another design issue might be the effectiveness of practicing and feedback in the training. Usually,

specific feedback and guidance is delivered by a teacher until the student is capable of performing the task independently. In the present study, feedback and guidance were provided by the enhanced multimedia learning environment. The feedback was not specific and there was no check whether students were capable of employing the strategic approach without guidance. The direct or indirect effect of the training might increase when the issues mentioned above are addressed in further research.

4.2 Indirect effect of invested mental effort on learning from process diagrams

An indirect effect of condition has been found on learning from process diagrams through invested mental effort. Students in the experimental condition invested more mental effort in the learning task of the posttest phase which had a positive effect on posttest score. The former result might suggest that students were willing to invest more mental effort because of increased strategy belief (Boekaerts, 1997), i.e., a motivational aspect of self-regulated learning. As discussed in the theoretical framework, students must be convinced that a strategy is essential and effortful before they are willing to use it (Alexander et al., 1998). We included several elements in the training that aimed at motivating students to invest more mental effort. For instance, we emphasized why diagrams are important and that elements of the training are based on findings from scientific research. Increased mental effort of students in the experimental condition could also be explained by increased self-efficacy in learning from process diagrams. Students' self-efficacy might have increased because the training facilitated how to learn from process diagrams following a strategic approach. Increased self-efficacy again might have increased students' willingness to invest more mental effort. From a cognitive load perspective (Sweller, 1994) it might be suggested that students' increased mental effort was germane, i.e., it facilitated learning; students who invested more mental effort learned more.

Students in the experimental condition also seemed to spend more time on the learning task in the posttest phase. This suggests that students employed the stepwise strategic approach; it is reasonable to expect that using the approach takes more time. The latter could be considered as supportive evidence for strategy belief (Boekaerts, 1997), i.e., students were willing to invest more time. Due to missing data we did not further explore how invested mental effort and more time on the learning task coincided.

4.3 Indirect effect of perceived task difficulty on learning from process diagrams

No indirect effect of condition was found on learning from process diagrams through perceived task difficulty. Students in the experimental condition did not perceive the posttest task to be more or less difficult than students in the control condition. This suggests that the strategic approach did not increase the complexity of the task by placing extra cognitive demands, neither did it make the diagram easier to understand, compared to the regular instruction condition. The significantly

negative relation of perceived task difficulty with learning from process diagrams was as expected; the more difficult the task was perceived, the lower the score on the posttest (cf. Robinson, 2001).

4.4 Limitations and recommendations for further research

While some limitations have already been discussed we here address a more general issue. We designed a training that used multiple instructional techniques (e.g., modeling, practicing, stepwise working-routine) and taught multiple strategies (e.g., activating prior knowledge, asking questions) to students in the experimental condition. We deliberately designed the training like this because we aimed at a maximum effect by including many elements that were found to be effective for learning. Furthermore, our goal was to design a training that could easily be used by teachers in classroom situations. The enhanced multimedia learning environment would allow teachers to train their students in learning from process diagrams without the need of becoming experts in strategy training. The backdraft of this choice of design is that it is impossible to determine which aspect of the training contributes to an effect or which aspect has to be improved.

We therefore recommend further research to examine the effect of the motivational, cognitive, and metacognitive aspects of the training (cf. Souvignier & Mokhlesgerami, 2006). Further research might also benefit from including more possible mediating variables like self-efficacy, strategy belief, and application of strategies as indicators of an indirect effect. In accordance with Zhao et al. (2010), including possible mediating variables is also a more general recommendation to researchers involved in intervention studies; Including mediators shows which variables are influenced by the intervention and how these variables contribute to the expected effect.

4.5 Conclusion and implication for education

In sum, we found that a multiple-strategy training, in a multimedia learning environment, affects learning from process diagrams positively via increased mental effort. The enhanced multimedia learning environment can easily be used by teachers as a medium for delivering the multiple-strategy training to their students; there were no technical problems running the learning environment on any of the computers of the schools that participated in the present study (except for the time on task macro). The latter and the evidence for its effectiveness suggest that the multiple-strategy training for learning from process diagrams has promising educational purposes.

CHAPTER 6

DISCUSSION

In this thesis, we aimed at getting deeper insight into students' difficulties comprehending process diagrams. These insights were used to inform the design of an intervention. In this concluding chapter, we first present the main results of the four separate studies of this thesis and then some reflections upon the studies and directions for future research. Finally, implications for educational practice are presented.

1. MAIN RESULTS

In the first two studies (chapter 2 and 3), we focused on factors that influence students' difficulties with process-diagram problem solving tasks. Two levels of process-diagram problem solving tasks were defined: tasks with a low and tasks with a high cognitive demand. The first study showed that cognitive task demand predicted students' difficulties with process-diagram problem solving tasks. Furthermore, students have difficulties with diagrams that use unfamiliar component conventions and that have a small number of components. Process-diagram tasks with a high cognitive demand when the content of the diagram was new proved to be more difficult than tasks with a high cognitive demand when the diagram did not contain new information.

Difficulties with process-diagram problem solving tasks are also predicted by student characteristics, i.e., prior knowledge, spatial abilities and working memory capacity (study 2, chapter 3). Students' performance on tasks with a high and a low cognitive demand were positively related to prior knowledge. Furthermore, scores on a spatial ability test that measured the ability to search for information in a complex spatial field were related to tasks with a low cognitive demand. The latter result was as hypothesized because we expected tasks with a low cognitive demand to rely on searching and encoding the presented information. Mental visualization and rotation abilities (i.e., visualization and spatial orientation factors, Ekstrom et al., 1976) were, as hypothesized, not related (with exception of one of five tests for these factors) to tasks with a low or a high cognitive demand (cf. Wu & Shah, 2003, for Chemistry problem solving). Students' scores on a visual working memory test were positively related to tasks with a high cognitive demand. The latter result was expected because tasks with a high cognitive demand require the formation of an elaborate mental model in working memory.

In the third study (chapter 4), we focused on learning activities that distinguished more and less successful students while studying process diagrams. In line with previous research (Meijer, Veenman, & Van Hout-Wolters, 2006; Pressley, 2000; Pressley & Afflerbach, 1995), we distinguished between an orientation and a main phase. In the orientation phase, successful students more often used the legend and activated prior knowledge. In the main phase, successful students more often: 1) gave meaning to process arrows, 2) questioned themselves, and 3) read the organizational levels of the diagrams. Results from measuring eye-movements showed that successful students also spend more time in the main area of the process diagram and shifted their focus more between different areas of interest. Successful students used a more coherent approach of interrelated learning activities and employed learning activities consistently across learning tasks.

The fourth study showed that students' learning from process diagrams can be enhanced by an intervention. Multiple-strategy training had an indirect positive significant effect on the experimental group, through invested mental effort, on learning from process diagrams, compared to the control group.

2. REFLECTIONS AND DIRECTIONS FOR FUTURE RESEARCH

2.1 Reflection on theory

Students construct a mental model in working memory when they process an external representation. For this thesis, multimedia theories (Mayer, 2001; Schnotz & Bannert, 2003), theories that focus on learning from graphical representations (e.g., Hegarty, 2005), and cognitive load theory (Sweller, 1988; Sweller, 1994; Sweller, Van Merriënboer, & Paas, 1998) provided a framework on how students construct a mental model of an external representation, i.c., process diagram. The construction of a mental model depends on (the interaction of) design features of the external representation (e.g., Canham & Hegarty, 2010; Mayer & Gallini, 1990), cognitive processes (Hegarty, 2005; Kriz & Hegarty, 2007), and abilities (Hegarty & Sims, 1994; Mayer & Sims, 1994; Sweller et al., 1998). Mental model construction is also influenced by the task; different tasks lead to different mental models (Schnotz & Bannert, 2003; Winn, 1991). The present thesis benefitted from the afore-mentioned theories; they served as a base to get deeper insight into students' difficulties with process diagrams. Such a deeper insight was also needed for the design of the intervention study (study 4, chapter 5). For this intervention study we could not rely on previous research—different external representations have different cognitive demands (Ainsworth, 2006)—and a thorough investigation on students' difficulties with process diagrams had not been performed yet. This thesis adds to previous research by showing how students' difficulties with a specific type of external representations, i.c., process diagrams, depends on design features (e.g., number of components), cognitive processes (e.g., activating prior knowledge; using the legend), abilities (e.g., working memory capacity), and task (cognitive task demand).

Research on self-regulated learning (Boekaerts, 1997) and (meta)cognition (Veenman, 2012) also supported the present thesis. Self-regulated learning theory provided a framework for students' learning activities while studying process diagrams. We also based the intervention of the final study (study 4, chapter 5) on self-regulated learning by using the cognitive strategy instruction model by Harris and Graham (1996), i.e., a model developed for supporting students in writing. We demonstrated that the latter model can also be used to facilitate students' learning from an external representation, i.c., the process diagram. Previous research on meta(cognition) identified many activities students employ, e.g., studying text (e.g., Pressley & Afflerbach, 1995) or solving problems (Meijer, Veenman, & Van Hout-Wolters, 2006). This thesis benefitted from these studies and adds some distinct activities (e.g., using the legend of a process diagram in the orientation phase) that can be distinguished when students study process diagrams.

2.2 Definition and measurement issues

A reoccurring issue in the four studies of the present thesis concerns how we defined and measured variables. In this section, we will reflect on some of these issues, discuss the effect these issues might have had on the results, and suggest some directions for future research.

In the first two studies (chapter 2 and 3), we defined two levels of task demand: tasks with a low and high cognitive task demand. In the first study, inter-rater agreement was high ($\kappa = .87$). In the second study, task demand explained 30% variance of task difficulty and internal reliability of low (KR-20 = .85); high (KR-20 = .82) cognitive task demand indicated a well-defined construct. Although two levels of task demand were functional with regard to this thesis' research questions, a more elaborated categorization of process-diagram problem solving tasks might be useful and could help the Biology education community more. Teachers might, for instance, benefit from such a categorization as a tool for selecting and designing tasks for an exam. Studies that focus on using Bloom's Taxonomy (Bloom, Krathwohl, & Masia, 1956) in Biology might give guidance to such a categorization: Crowe, Dirks, and Wenderoth (2008) developed an instrument for classifying tasks for several biological topics; Quillin and Thomas (2015) developed an instrument for classifying drawing tasks within Biology.

In the first study (chapter 2), we obtained items from Dutch Biology national exams. An advantage was that we were able to work with a large dataset ($N = 42891$) and could select tasks from many items (704). A disadvantage was that we were not able to measure effects of student variables directly. Instead, we had to rely on inter-rater agreement by experts. Inter-rater agreement (Cohen's' κ) showed that these constructs were measured reliable; students' prior knowledge and familiarity with the components, arrows, and spatial arrangement of the process diagrams were .87, .86, .82, .87, respectively. Not measuring the latter variables directly meant that within cohort between students variance—cohort mean exam scores were used as the dependent variable—was not addressed. It is, however, clear that students' prior

knowledge and familiarity with the components, arrows, and spatial arrangement, varies within cohorts. Future studies might consider a more fine-grained approach by directly measuring students' familiarity with the components, arrows, and spatial arrangement (prior knowledge is directly measured in the second study). Furthermore, contrary to our expectations, multiple regression analysis showed no significant effect (when controlled for other factors, e.g., cognitive task demand) for prior knowledge on task difficulty. We cannot rule out that not finding a significant effect might partly be due to not considering between student variance in prior knowledge.

In the second study (chapter 3), we measured students' prior knowledge on four secondary school Biology topics (ecology, protein synthesis, dissimilation, and hormones). All prior knowledge items (56) were aggregated into a single scale and internal reliability (KR-20 = .78) was confirmed. Tasks with low or high cognitive demand (KR-20 = .85, .82, respectively) were also based on the four topics and we assumed performance on these tasks to be invariant across topics. We examined whether prior knowledge scores were related to scores on tasks with low and high cognitive demand. This meant that we did not assume that there were between topic differences in the relation between tasks with a low and high cognitive demand and prior knowledge. We also did not assume that there were between topic differences in the relation between scores on tasks with a low and high cognitive demand and scores on the spatial ability and working memory tests. One could argue that the latter relations might vary per topic of the process diagram. For instance, the role of prior knowledge on task performance might be more distinct when the design features of process diagrams of a specific topic are more conceptual and abstract (see study 1, chapter 2) compared to other topics. Future studies might provide insight whether the relations examined in the second study vary per topic.

In the fourth study (chapter 5), we assessed with a pretest how much students comprehended of a process diagram they had just studied. The items (8) of the pretest did not converge into a single scale, i.e., internal homogeneity as indicated by KR-20 could not be confirmed. There might be several reasons for this. For instance, the pretest (and the posttest) assessed how much a student was able to comprehend from a process diagram. It is plausible different students focused on different aspects of the process diagram in the pretest and that this compromised internal homogeneity. Another possibility is that student's prior knowledge might be fragmented; knowing one item is not per se a precondition for knowing the other item. Learning the content of the process diagram of the pretest relied on basic prior knowledge of the topic (i.e., entero-hepatic cycle) and on prior knowledge about (biochemical) transformations and transport (e.g., diffusion, active and passive transport). When different students had different prior knowledge this might also have compromised internal homogeneity of the pretest. We must mention here that although the topics of the process diagrams of the pretest and posttest differed, we did not expect that this influenced internal validity. The topics of the process diagrams of the pretest and the posttest were both within the domain of biochemistry and relied mainly on (biochemical) transformations and transport processes. The latter processes were not

instructed in the intervention phase; hence, the influence of prior knowledge was kept constant.

2.3 Ecological validity and generalization

All participants in all four studies were students who were part of the target group of the present thesis, i.c., pre-university upper secondary school students with an intermediate level of expertise in Biology. All tests were administered at students' school, e.g., in classroom, computer lab, exam room. In the first two studies, parts of the data were collected during a national exam (first study, chapter 2) and a regular school exam (second study, chapter 3). Furthermore, we used authentic tasks students regularly encounter in secondary Biology education. Process diagrams were carefully selected (e.g., from Biology textbooks used in higher education) and adjusted for the target group. Biology teachers advised us on many issues on a regular basis.

There might be some concerns about the generalizability of the second and third study because students are from two classes from a single school (but not the same two classes in both studies). However, the school is a regular secondary school in the Netherlands and we have no reasons to expect that our samples are not representative. Another issue with respect to generalization is this thesis' target group, i.c., pre-university upper secondary school students with an intermediate level of expertise in Biology. The results might be specific for students with their level of expertise and experience with learning and problem solving with process diagrams. For this, we recommend to be careful to extend the results to other target groups such as lower grade students and university students.

Furthermore, this thesis focused on a specific type of diagram, i.c., the process diagram. Process diagrams used in this thesis were from a single domain, i.c., Biology, but we used many different process diagrams with a variety of biological topics and design features. For this, we believe our findings extent to process diagrams from other domains (e.g., chemistry, physics, geography, etc.). However, results cannot be easily extended to other type of diagrams (e.g., tree diagrams, phase diagrams, Venn diagrams) because of substantial differences in design features.

A final remark can be made with respect to the generalizability of the research design of the intervention in the fourth study (chapter 5). Future intervention studies might also benefit from including possible mediating variables in their design (see also Zhao, Lynch, & Chen, 2010). As mentioned in the fourth study: "Including mediators shows which variables are influenced by the intervention and how these variables contribute to the expected effect" (p. 77).

2.4 Unit of analysis

In the third study, we identified learning activities that distinguished more and less successful students while studying process diagrams. We performed correlational analyses to examine how these learning activities are interrelated. Although the

analysis offers insight into the coherence of students' approach while studying process diagrams, learning activities are not single events but belong to a string of activities that contain students' means and goals. Correlational analyses do not consider the order and the functional relationship of the learning activities. Future studies might consider an analysis that takes the order and the functional relationship of the learning activities into consideration, e.g., path analysis.

2.5 Consistency

In this section, we will discuss how the first three studies of this thesis have informed the final intervention study (study four, chapter 5). We will discuss which paths were followed and which were not and some suggestions for future research will be provided.

In the concluding section of the first study (chapter 2) we formulated three suggestions for the design of an intervention. One of the suggestions of the first study was followed-up in the final intervention study, i.e., the suggestion to "facilitate students in learning how to gain a deeper understanding of diagrams that contain new information". This suggestion was based on the finding that students have difficulties with tasks with high cognitive demand when the content is new. The other two suggestions of the first study had no follow-up in the present thesis. One suggestion was to teach students strategies for encoding diagrams with unfamiliar components and the other suggestion was to focus on problems solving with abstract process diagrams. The suggestion of the second study (chapter 3) to "design a training that focuses on solving tasks with high cognitive demand" was also not followed-up in the final intervention study of this thesis.

We did not include these suggestions for practical reasons and due to progressive insight. Including these suggestions would largely increase the size of the intervention. As mentioned in the final study, we were dedicated to develop a short training to provide teachers the opportunity to use the intervention in classroom. Furthermore, the intervention focused on *learning* from process diagrams and not on *problem solving*; the findings of the first two studies were related to students 'difficulties with process-diagram problem solving tasks. However, we assume that there is an overlap in activities students have to employ when *solving tasks* with a high cognitive demand and when *learning* from process diagrams. Tasks with a high cognitive demand are usually more global and, as in learning, students have to search, encode, integrate and infer (evaluate, compare, judge) the presented information (Hegarty, 2005). Hence, facilitating students *how* to learn from process diagrams might also help them to solve problems with process diagrams. Future studies might, however, incorporate the suggestions that were not followed-up in the final intervention of this thesis into the design of a training.

Furthermore, in the second study (chapter 3), we identified spatial ability factors that were related to process-diagram problem solving. In the concluding section of that chapter we suggested that learning strategies might be especially important for students with low spatial skills. This suggestion had no follow-up in the fourth study

(chapter 5). The decision not to include spatial ability tests in the pretest phase of the fourth study was also based on a practical reason: administering spatial ability tests cost extra time that was not available. However, in future studies we recommend to include spatial factors as possible moderating factors for the effect of an intervention that facilitates students' understanding of process diagrams. Students with low spatial skills might benefit more from such an intervention: The intervention might learn them strategies that compensate for their low spatial skills.

Finally, the third study (chapter 4) is strongly connected to the fourth study (chapter 5). The third study identified learning activities that distinguished more and less successful students while studying process diagrams. We assumed that students would benefit from facilitating them to employ these learning activities. Therefore, we included these learning activities into the design of the intervention.

3. PRACTICAL IMPLICATIONS

This thesis presents a fined-grained analysis of students' difficulties comprehending process diagrams. We think this thesis might help stakeholders of the Biology education community (e.g., students, teachers, teacher educators, instructional designers) to make informed decisions with regard to process diagrams when they teach, select or design learning materials, design exams, etc.

First, teachers should be aware of students' difficulties when dealing with process diagrams. Process diagrams are important and abundantly present in Biology education and teachers should know that students do not become diagrammatically literate by just studying it. We argue that students need more support when they have to solve process-diagram problems, e.g., as a task in their textbook, or when they have to learn from a process diagram, e.g., from their textbook or during instruction. The present thesis identified several conditions (e.g., cognitive task demand, familiarity with the components) that influence students' difficulties; teachers should consider these conditions when their students have to learn or solve problems with process diagrams. Especially, we want to advise teachers to demonstrate (more often) how they learn or solve problems with process diagrams by acting as a model. Teachers could act as a model by thinking-aloud to demonstrate how and when they activate prior knowledge, make inferences, use meta-information (e.g., title and the legend), ask questions, etc.

Second, teachers should be aware that individual differences, e.g., in prior knowledge, spatial ability, working memory capacity, meta(cognitive) learning activities, between students—even within homogenous groups—have an effect on students when they have to comprehend process diagrams. Teachers could use these findings in classroom. For instance, teachers might activate students' prior knowledge before process diagrams are instructed or give extra support to students with lower abilities.

Finally, this thesis might be interesting for teacher educators and instructional designers. Teacher educators might inform prospective teachers about students' difficulties with process diagrams. Such a lesson might be a good starting point for

introducing themes like (meta)cognitive strategies, observational learning, and self-regulated learning. Instructional designers might benefit from this thesis in various ways. It might help them by making more informed decisions when selecting or designing process diagrams (and adjacent tasks) for educational materials. Furthermore, we hope it helps them to design educational material with more attention to *how* to comprehend process diagrams. For instance, the educational material might include instructions and tasks that help students to develop a more strategic approach (see fourth study, chapter 5). Students could also benefit from tasks where they have to explain the content of a process diagram to each other. Observing, comparing, and evaluating task behaviour (i.e., observational learning) of peers has proven to be effective in the domain of writing (Braaksma, 2002; Raedts, Rijlaarsdam, Van Waes, & Daems, 2007).

REFERENCES

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newsbury Park, CA: Sage Publications.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16,* 183–198. doi:10.1016/j.learninstruc.2006.03.001

Ainsworth, S., & Loizou, A. Th. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science, 27*, 669–681. doi:10.1207/s15516709cog2704_5

Akyol, G., Sungur, S., & Tekkaya, C. (2011). The contribution of cognitive and metacognitive strategy use to students' science achievement. *Educational Research and Evaluation, 16*, 1–21. doi:10.1080/13803611003672348

Alexander, P. A., Graham, S., & Harris, K. R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review, 10*, 129–154. doi:10.1023/A:1022185502996

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*, 523–535. doi:10.1037/0022-0663.96.3.523

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

Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Writing Communications, 25*, 166–195. doi:10.1177/0741088307313177

Bielaczyc, K., Pirolli, P., & Brown, A. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem-solving. *Cognition and Instruction, 13*, 221–252. doi:10.1207/s1532690xci1302_3

Bloom, B. S., Krathwohl, D. R., & Masia, B. B. (1956). *Taxonomy of educational objectives: The classification of educational Goals*. New York, NY: D. McKay.

Bodner, G. M., & McMillen, T. L. B. (1986). Cognitive restructuring as an early stage in problem solving. *Journal of Research in Science Teaching, 23*, 727–737. doi:10.1002/tea.3660230807

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction, 7*, 161–186. doi:10.1016/S0959-4752(96)00015-1

Bowen, G. M., & Roth W. (2002). Why students may not learn to interpret scientific inscriptions. *Research in Science Education, 32,* 303–327. doi:10.1023/a:1020833231966

Braaksma, M. A. H. (2002). Observational learning in argumentative writing. Unpublished dissertation. University of Amsterdam.

Buckley, B. C. (2000). Interactive multimedia and model-based learning in biology. *International Journal of Science Education*, *22*(9), 895–935. doi:10.1080/095006900416848

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology, 98*, 182–197. doi:10.1037/0022-0663.98.1.182

Campbell, N. A., & Reece, J. B. (2002). *Biology* – 6th ed. San Francisco. CA: Pearson Education.

Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction, 20*, 155–166. doi:10.1016/j.learninstruc.2009.02.014

Carlson, R., Chandler, P., & Sweller, J. (2003). Learning and understanding science instructional material. *Journal of Educational Psychology, 95*, 629–640. doi:10.1037/0022-0663.95.3.629

Catley, K. M., Novick, L. R., & Shade, C. K. (2010). Interpreting evolutionary diagrams: When topology and process conflict. *Journal of Research in Science Teaching*, *47*, 861–882. doi: 10.1002/tea.20384

Chi, M., (2000). Self-explaining expository texts: The dual process of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in Instructional Psychology* (Vol. 5, pp. 161–238). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representations of physics problems by experts and novices. *Cognitive Science,* 5, 121–152. doi:10.1207/s15516709cog0502_2

Chittleborough, G., & Treagust, D. (2008). Correct interpretation of chemical diagram requires transforming from one level of representation to another. *Research in Science Education, 38*, 463–482. doi:10.1007/s11165-007-9059-4

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load. *Science Education, 90*, 1073–1091. doi:10.1002/sce.20164

Cook, M., Carter, G, & Wiebe, E. N. (2008). The interpretation of cellular transport graphics by students with low and high prior knowledge. *International Journal of Science Education, 30*, 239–261. doi:10.1080/09500690601187168

Cook, M., Wiebe, E. N., & Carter, G. (2008). The influence of prior knowledge on viewing and interpreting graphics with macroscopic and molecular representations. *Science Education, 92*, 848–867. doi:10.1002/sce.20262

Cromley, J. G., & Azevedo, R. (2006). Self-report of reading comprehension strategies: What are we measuring? *Metacognition and Learning, 1*, 229–247. doi:10.1007/s11409-006-9002-5

Cromley, J. G., Bergey, B. W., Fitzhugh, S., Newcombe, N., Wills, T. W., Shipley, T. F., Tanaka, J. C. (2013a). Effects of three diagram instruction methods on transfer of diagram comprehension skills: The critical role of inference while learning. *Learning and Instruction, 26*, 45–58. doi:10.1016/j.learninstruc.2013.01.003

Cromley, J. G., Perez, T. C., Fitzhugh, S. L., Newcombe, N. S., Wills, T. W., & Tanaka, J. C. (2013b). Improving students' diagram comprehension with classroom instruction. *The Journal of Experimental Education, 81*, 511–537. doi: 10.1080/00220973.2012.745465

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology, 35*, 59–74. doi: 10.1016/j.cedpsych.2009.10.002

Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in bloom: Implementing Bloom's Taxonomy to enhance student learning in biology. *Life Sciences Education, 7,* 368–381. doi:10.1187/cbe.08-05-0024

Delen, E., Liew, J., & Willson, V. (2014). Effects of interactivity and instructional scaffolding on learning: Self-regulation in online video-based environments. *Computers & Education, 78*, 312–320. doi:10.1016/j.compedu.2014.06.018

Duffy, G. G., Roehler, L. R., Sivan, E., Rackliffe, G., Book, C., Meloth, M., Vavrus, L. G., Wesselman, R., Putnam, J., & Bassiri, D. (1987). Effects of explaining the reasoning associated with using reading strategies. *Reading Research Quarterly, 22*, 347–368. doi:10.2307/747973

Dole, J. A., Nokes, J. D., & Drits, D. (2009). Cognitive strategy instruction. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 347–372). New York, NY: Taylor & Francis. doi:10.4324/9781315759609

Donker, A. S., De Boer, H., Kostons, D., Dignath Van Ewijk, C. C., & Van Der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review, 11*, 1–26. doi:10.1016/j.edurev.2013.11.002

Dos Santos, V. J., & Galembeck, E. (2015). Metabolic pathways visualization skills development by undergraduate students. *Biochemistry and Molecular Biology Education*. doi:10.1002/bmb.20858

Ekstrom, R. B., French, J. W, Harman, H. H., & Dermen, D. (1976). *Kit of factor referenced cognitive tests*. Princeton: Educational Testing Service.

Ericsson, K., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data (2nd ed.)*. Boston: MIT Press.

Frazier, P., Tix, A., & Barron, K. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology, 51,* 115–134. doi:10.1037/0022-0167.51.1.115

Gilbert, J. K. (2005). Visualization: A metacognitive skill in science education. In Gilbert, J. K. (Ed.), *Visualization in science education* (pp. 9–27). Dordrecht, the Netherlands: Springer.

Guthrie J. T., Shelley, W., & Kimmerly, N. (1993). Searching documents: Cognitive processes and deficits in understanding graphs, tables, and illustrations. *Contemporary Educational Psychology, 18,* 186–221. doi:10.1006/ceps.1993.1017

Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology, 24*, 95–123. doi: 10.1006/ceps.1998.0987

Hardy, M. A. (1993). *Regression with dummy variables*. Newsbury Park, CA: Sage Publications.

Harris, K. R., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation.* Cambridge, MA: Brookline.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, *66*, 99–136. doi:10.3102/00346543066002099

Heiser J., & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagram. *Cognitive Science, 30*, 581–592. doi:10.1207/s15516709cog0000_70

Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1084–1102. doi:10.1037/0278-7393.18.5.1084

Hegarty, M. (2005). Multimedia learning about physical systems. In: R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 447–465). Cambridge: Cambridge University Press.

Hegarty, M., & Just, M. A. (1989). Understanding machines from text and diagrams. In H. Mandl & J. Levin (Eds.), *Knowledge acquisition from text and picture* (pp. 171–194). Amsterdam, the Netherlands: Elsevier.

Hegarty, M., & Just, M. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory & Language, 32*, 717–742. doi:10.1006/jmla.1993.1036

Hegarty, M., & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory and Cognition,* 22, 411–430. doi:10.3758/bf03200867

Holliday, W. G., Brunner, L. L., & Donais, E. L. (1977). Differential cognitive and affective responses to flow diagrams in science. *Journal of Research in Science Teaching, 14*, 129–138. doi:10.1002/tea.3660140205

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures.* Oxford, UK: Oxford University Press.

Hoyle, R. H., Harris, M. J., & Judd, C. M. (2002). *Research methods in social relations.* USA, Wadsworth: Thompson Learning.

Hox, J. (2010). *Multilevel Analysis: Techniques and Applications* (2nd Edition). New York, NY: Routledge.

Jarodzka, H., Van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see: Guiding students' attention via a Model's eye movements fosters learning. *Learning and Instruction*, *25*, 62–70. doi:10.1016/j.learninstruc.2012.11.004

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*, 441–480. doi:10.1016/0010-0285(76)90015-3

Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior, 28*, 820–831. doi:10.1016/j.chb.2012.01.011

Kindfield, A. C. H. (1993). Biology diagrams: Tools to think with. *Journal of the Learning Sciences, 3*, 1–36. doi:10.1207/s15327809jls0301_1

King, A. (1989). Effects of self-questioning training on college students' comprehension of lectures. *Contemporary Educational Psychology, 14*, 366-381. doi:10.1016/0361-476X(89)90022-2

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*, 212–218. doi:10.1207/s15430421tip4104_2

Körner, C. (2005). Concepts and misconceptions in comprehension of hierarchical graphs. *Learning and Instruction*, *15*, 281–296. doi:10.1016/j.learninstruc.2005.07.003

Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002). Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction, 20*, 47–77. doi: 10.1207/s1532690xci2001_3

Kozhevnikov, M., Motes, A., & Hegarty, M. (2007). Spatial visualization in physics problem solving. *Cognitive Science, 31*, 549–579. doi:10.1080/15326900701399897

Knippels, M. C. P. J. (2002). Coping with the abstract and complex nature of genetics in biology education. The yo-yo learning and teaching strategy. Utrecht, the Netherlands: CD-ß Press.

Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2013a). Diagrammatic literacy in secondary science education. *Research in Science Education*, *43*, 1785–1800. doi:10.1007/s11165-012-9331

Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2013b). Geletterdheid in diagrammen in de bètavakken. *Tijdschrift voor Taalbeheersing*, *35*, 63–81. doi:10.5117/TVT2013.1.KRAG

Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2015). Students' ability to solve process-diagram problems in secondary Biology education. *Journal of Biological Education, 49*, 1–13. doi: 10.1080/00219266.2014.888363

Kragten, M., Admiraal, W., Rijlaarsdam, G. (2015). Students' learning activities while studying biological process diagrams. *International Journal of Science Education, 37*, 1915–1937. doi:10.1080/09500693.2015.1057775

Kriz, S., & Hegarty, M. (2007). Top-down and bottom-up influences on learning from animations. *International Journal of Human-Computer Studies, 65*, 911–930. doi:10.1016/j.ijhcs.2007.06.005

Larkin, J. H., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science, 208*, 1335–1342. doi:10.1126/science.208.4450.1335

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth 10,000 words. *Cognitive Science, 11*, 65–99. doi:10.1111/j.1551-6708.1987.tb00863.x

Lodish, H., Berk, A., Kaiser, C. A., Krieger, M., Bretscher, A., Ploegh, H., Amon, A. & Scott, M. P. (2012). *Molecular Cell Biology- 7th edition*. New York, NY: W. H. Freeman and Company.

Lowe, R. K. (1996). Background knowledge and the construction of a situational representation from a diagram. *European Journal of Psychology of Education, 11*, 377–397. doi:10.1007/bf03173279

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education, 14*, 225–244. doi:10.1007/bf03172967

Lowe, R., & Schnotz, W. (2008). *Learning with animation: Research implications for design.* Cambridge University Press.

Mason, L., Pluchino, P., & Tornatora, M. C. (2013). An eye-tracking study of learning from science text with concrete and abstract illustrations. *The Journal of Experimental Education, 81*, 356–384. doi:10.1080/00220973.2012.727885

Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology*, *41*, 172–187. doi:10.1016/j.cedpsych.2015.01.004

Mathai, S., & Ramadas, J. (2009). Visuals and visualisation of human body systems. *International Journal of Science Education, 31*, 439–458. doi:10.1080/09500690802595821

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

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*, 715–726. doi:10.1037//0022-0663.82.4.715

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*, 312–320. doi:10.1037/0022-0663.90.2.312

Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction, 12*, 107–119. doi:10.1016/s0959-4752(01)00018-4

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*, 43–52. doi:10.1207/s15326985ep3801_6

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86*, 389–401. doi:10.1037/0022-0663.86.3.389

Meijer, J., Veenman, M. V. J., & Van Hout-Wolters, B. H. A. M. (2006). Metacognitive activities in text-studying and problem-solving: Development of a taxonomy. *Educational Research and Evaluation, 12*, 209–237. doi:10.1080/13803610500479991

Paas, F., Tuovinen, J. E., Van Merriënboer, J. J. G., & Darabi, A. 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. doi:10.1007/bf02504795

Paivio, A. (1990). *Mental representations. A dual coding approach*. New York, NY: Oxford University Press.

Palincsar A. S., & Brown, A. L. (1994). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*, 117–175. doi:10.1207/s1532690xci0102_1

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33–40. doi:10.1037/0022-0663.82.1.33

Pohlmann, A., Fricke, W. F., Reinecke, F., Kusian, B., Liesegang, H., Cramm, R … Bowien, B. (2006). Genome sequence of the bioplastic-producing 'Knallgas' bacterium *Ralstonia eutropha* H16. *Nature Biotechnology, 24*, 1257–1262. doi:10.1038/nbt1244

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891. doi:10.3758/BRM.40.3.879

Pressley, M. (2000). Development of grounded theories of complex cognitive processing: Exhaustive within- and between-study analyses of think-aloud data. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 261–296). Lincoln, NE: Buros Institute of Mental Measurements.

Pressley, M., & Afflerbach, P. (1995). Verbal protocols of reading: The nature of constructively responsive reading. Hillsdale, NJ: Erlbaum.

Pribyl, J. R., & Bodner, G. M. (1987). Spatial ability and its role in organic chemistry: A study of four organic courses. *Journal of Research in Science Teaching, 24*, 229–240. doi:10.1002/tea.3660240304

Pressley, M., Borkwski, J. G., & Schneider, W. (1989). Good information processing: What it is and how education can promote it. *International Journal of Educational Research*, *13*, 857–867. doi:10.1016/0883-0355(89)90069-4

Pressley, M., El-Dinary, P. B., Gaskins, I., Schuder, T., Bergman, J. L., Almasi, J., & Brown, R. (1992). Beyond direct explanation: Transactional instruction of reading comprehension strategies. *The Elementary School Journal, 92*, 513–555. doi:10.1086/461705

Pressley, M., & Harris, K. (2006). Cognitive strategies instruction: From basic research to classroom instruction. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 265–286). Mahwah, NJ: Erlbaum. doi:10.4324/9780203874790.ch12

Quillin, K., & Thomas, S. (2015). Drawing-to-Learn: A framework for using drawings to promote model-based reasoning in Biology. *CBE-Life Sciences Education*, *14*, 1–16. doi:10.1187/cbe.14-08-0128

Raedts, M., Rijlaarsdam, G., Van Waes, L., & Daems, F. (2007). Observational learning through video-based models: Impact on student's writing accuracy of self-efficacy beliefs, task knowledge and writing performances (pp. 219–238). In P. Boscolo & S. Hidi (Eds), *Studies in Writing, 20. Writing and motivation*. Oxford, UK: Elsevier. doi:10.1108/s1572-6304(2006)0000019014

Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A. Minorsky, P. V., & Jackson, R. B. (2010). *Campbell Biology*. 9th ed. San Francisco, CA: Pearson Education.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, *22*, 27–57. doi:10.1093/applin/22.1.27

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64*, 479–530. doi:10.3102/00346543064004479

Roth, W.-M., & McGinn, M. K. (1998). Inscriptions: A social practice approach to representations. *Review of Educational Research, 68*, 35–59. doi:10.3102/00346543068001035

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, *13*, 141–156. doi:10.1016/s0959-4752(02)00017

Schönborn, K. J., Anderson, T. R., & Grayson, D. J. (2002). Student difficulties with the interpretation of a textbook diagram of Immunoglobulin G (IgG). *Biochemistry and Molecular Biology Education, 30*, 93–97. doi:10.1002/bmb.2002.494030020036

Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist, 32*, 195–208. doi:10.1207/s15326985ep3204_1

Schwonke, R., Berthold, K., & Renkl, A. (2009). How multiple external representations are used and how they can be made more useful. *Applied Cognitive Psychology, 23*, 1227–1243. doi:10.1002/acp.1526

She, H., & Chen, Y. (2009) The impact of multimedia effect on science learning: Evidence from eye movements. *Computers & Education, 53*, 1297–1307. doi:10.1016/j.compedu.2009.06.012

Sims, V. K., & Hegarty, M. (1997). Mental animation in the visuospatial sketchpad: Evidence from dual-task studies. *Memory and Cognition, 25*, 321–332. doi:10.3758/bf03211288

Snijders, T. A. B., & Bosker, R. J. (1994). Modeled variance in two-level models. *Sociological Methods & Research, 22*, 342–363. doi:10.1177/0049124194022003004

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology, 13*, 290–312. doi:10.2307/270723

Souvignier, E., & Mokhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction*, *16*, 57–71. doi:10.1016/j.learninstruc.2005.12.006

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257–285. doi:10.1207/s15516709cog1202_4

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295–312. doi:10.1016/0959-4752(94)90003-5

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–295. doi:10.1023/a:1022193728205

Szpunar, K. K., Jing, H. G., & Schacter, D. L. (2014). Overcoming overconfidence in learning from video-recorded lectures: Implications of interpolated testing for online education. *Journal of Applied Research in Memory and Cognition*, *3*, 161–164. doi:10.1016/j.jarmac.2014.02.001

Van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 43*, 16–26. doi:10.1080/00461520701756248

Van Gog, T., Paas, F., & Van Merriënboer, J. (2005). Uncovering expertise-related differences in trouble-shooting performance. Combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology, 19*, 205–221. doi:10.1002/acp.1112

Veenman, M. V. J. (2012). Metacognition in science education: Definitions, constituents, and their intricate relation with cognition. In A. Zohar & Y. J. Dori (Eds.), *Metacognition in Science Education: Trends in Current Research* (Vol. 40, pp. 21–36). Dordrecht, the Netherlands: Springer.

Verhoeff, R. P. (2003). *Towards systems thinking in cell biology education*. Utrecht, the Netherlands: CD-ß Press.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *Journal of Educational Research, 84*, 30–34. Retrieved from http://www.jstor.org/stable/40539680

Winn, W. D. (1982). The role of diagrammatic representation in learning sequences, identification, and classification as a function of verbal and spatial ability. *Journal of Research in Science Teaching, 19*, 79–89. doi:10.1002/tea.3660190110

Winn, W. D. (1988). Recall of the pattern, sequence, and names of concepts presented in instructional diagrams. *Journal of Research in Science Teaching, 25*, 375–386. doi:10.1002/tea.3660250505

Winn, W. D. (1991). Learning from maps and diagrams. *Educational Psychology Review, 3*, 211–247. doi:10.1007/bf01320077

Winn, W. D. (1993). An account of how readers search for information in diagram. *Contemporary Educational Psychology, 18*, 162–185. doi:10.1006/ceps.1993.1016

Winn, W. D., Li, T-Z., & Schill, D. E. (1991). Diagram as aids to problem solving: Their role in facilitating search and computation. *Educational Technology Research and Development, 39*, 17–29.

Winn, W. D., & Sutherland, S. W. (1989). Factors influencing the recall of elements in maps and diagrams and the strategies used to encode them. *Journal of Educational Psychology, 81*, 33–39. doi:10.1037/0022-0663.81.1.33

Winters, F. I., Greene, J. A., & Costich, C. M. (2008). Self-regulation of learning within computer-based learning environments: A critical analysis. *Educational Psychology Review, 20*, 429–444. doi:10.1007/s10648-008-9080-9

Wu, H., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education, 88*, 465–492. doi:10.1002/sce.10126

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*, 197–206. doi:10.1086/651257

Zito, J. R., Adkins, M., Gavins, M., Harris, K. R., & Graham, S. (2007). Self-regulated strategy development: Relationship to the social-cognitive perspective and the development of self-regulation. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 23*, 77–95. doi: 10.1080/10573560600837693

AUTHOR INDEX

Adkins, M., 67
Admiraal, W., 7, 13, 29, 31, 33, 43
Afflerbach, P., 10, 44, 45, 60, 61, 62, 86, 87
Aiken, L. S., 23, 26
Ainsworth, S., 13, 14, 86
Akyol, G., 44
Alexander, P. A, 65, 73, 82
Anderson, T. R., 29, 43, 65
Azevedo, R., 10, 44, 50, 51, 62, 66

Baddeley, A. D., 37
Bannert, M., 9, 30, 31, 37, 66, 86, 106, 111
Barron, K., 23
Berthold, K., 45
Bezemer, J., 7, 14, 27, 43
Bielaczyc, K., 28
Biggs, J., 66
Bloom, B. S., 14, 19, 87
Bodner, G. M., 31
Boekaerts, M., 10, 44, 66, 82, 87, 105, 106, 111
Borkwski, J. G., 67
Bosker R. J., 56, 58
Bowen, G. M., 7, 13, 14
Braaksma, M. A. H., 92
Brown, A. L., 28, 67
Brunner, L. L., 16, 17
Buckley, B. C., 33
Butcher, K. R., 13, 44, 50

Campbell, N. A., 7, 14, 33
Canham, M., 16, 17, 28, 31, 45, 46, 65, 66, 86
Carlson, R., 7, 15, 17, 27, 28, 46
Carpenter, P. A., 45, 61
Carter, G., 15, 31, 45, 66
Catley, K. M., 46
Chandler, P., 7, 15, 46
Chen, Q., 81, 89
Chen, Y., 45, 55, 61, 62, 76
Chi, M. T. H., 9, 30, 31, 44, 50, 51, 55, 61
Chittleborough, G., 7, 43, 65
Cohen, J., 19, 21, 60, 87
Cook, M. P., 15, 28, 29, 31, 45, 46, 60, 61, 66
Costich, C. M., 68
Cromley, J. G., 10, 13, 44, 50, 51, 62, 65, 66, 74
Crowe, A., 14, 19, 33, 87

Daems, F., 10, 28, 92
Darabi, A. A., 66
De Boer, H., 65
Delen, E., 68
Dignath Van Ewijk, C. C., 65
Dirks, C. P., 14, 87
Dole, J. A., 65, 67
Donais, E. L., 16, 17
Donker, A. S., 65
Dos Santos, V. J., 62, 69
Drits, D., 65
Duffy, G. G., 67

Ekstrom, R. B., 36, 37, 38, 85, 103, 104, 109
Ericsson, K., 43

Feltovich, P. J., 9, 30
Frazier, P., 23, 24
French, J. W., 19

Galembeck, E., 62, 69
Gallini, J. K., 10, 15, 16, 17, 28, 86
Gavins, M., 67
Gilbert, J. K., 14
Glaser, R., 9, 30
Graham, S., 11, 65, 67, 68, 73, 87, 105, 111
Grayson, D. J., 7, 14, 29, 43, 65
Greene, J. A., 68
Guthrie J. T., 9, 14, 19, 28, 29, 33

Haertel, G. D., 44
Hannus, M., 61
Hardy, M. A., 23
Harris, K. R., 11, 18, 65, 67, 68, 73, 87, 105, 111
Hattie, J., 66
Hayes, A. F., 76, 77, 81
Hegarty, M., 9, 15, 16, 17, 28, 29, 30, 31, 33, 36, 37, 38, 41, 43, 44, 45, 46, 51, 65, 66, 73, 86, 90, 104, 106, 109, 111
Heiser J., 16, 43
Holliday, W. G., 16, 17, 28
Holmqvist, K., 44
Hox, J., 56, 58
Hoyle, R. H., 18
Hyönä, J., 61

Jarodzka, H., 68, 74
Jing, H. G., 75
Judd, C. M., 18
Just, M. A., 15, 33, 45, 46, 51, 61

Kay, R. H., 68
Kimmerly N., 14, 29

Kindfield, A. C. H., 14, 30
King, A., 60
Knippels, M. C. P. J., 61
Körner, C., 46, 69
Kostons, D., 65
Kozhevnikov, M., 9, 31, 36, 37, 41
Kragten, M., 7, 13, 29, 31, 33, 43, 51, 74
Krathwohl, D. R., 9, 14, 87
Kress, G., 7, 14, 27, 43
Kriz, S., 43, 44, 45, 46, 65, 66, 73, 86

Larkin, J. H., 7, 9, 13, 29, 30, 65, 73
Li, T-Z., 15
Liew, J., 68
Lodish, H., 50
Loizou, A. Th., 13
Lowe, R. K., 15, 27, 28, 50, 51, 61, 66, 68
Luciw-Dubas, U. A., 10, 13, 44, 66
Lynch, J. G., 81, 89

Masia, B. B., 14, 87
Mason, L., 45, 46, 61, 62, 68, 74
Mathai, S., 14
Mayer, R. E., 9, 10, 15, 16, 17, 28, 30, 31, 36, 37, 41, 46, 71, 73, 86, 105, 106, 110, 111
McDermott, J., 30
McGinn, M. K., 43
McMillen, T. L. B., 31
Meijer, J., 10, 44, 50, 51, 63, 73, 86, 87
Meister, C., 67, 81
Mokhlesgerami, J., 65, 67, 81, 83
Moreno, R., 16, 30, 36, 71, 105, 110
Motes, A., 9, 31

Nokes, J. D., 65
Novick, L. R., 8, 46

Paas, F., 9, 44, 45, 48, 50, 60, 62, 66, 72, 86, 104, 109
Paivio, A., 16, 73
Palincsar A. S., 67
Pintrich, P. R., 66
Pirolli, P., 28
Pluchino, P., 45, 68
Pohlmann, A., 50, 76
Preacher, K. J., 76, 77, 81
Pressley, M., 10, 44, 60, 61, 62, 65, 67, 68, 73, 86, 87
Pribyl, J. R., 31
Purdie, N., 66

Quillin, K., 55, 76, 87

Raedts, M., 10, 28, 92
Ramadas, J., 14
Reece, J. B., 7, 14, 29, 43, 69
Renkl, A., 45
Rijlaarsdam, G., 7, 10, 13, 28, 29, 31, 33, 43, 92
Robinson, P., 69, 73, 83
Rosenshine, B., 67, 81
Roth, W.-M., 7, 13, 14, 43

Schacter, D. L., 75
Schill, D. E., 15
Schneider, W., 67
Schnotz, W., 9, 30, 31, 37, 66, 68, 86, 106, 111
Schönborn, K. J., 7, 14, 29, 43, 65
Schunk, D. H., 10, 66, 67, 68
Schwonke, R., 45, 46, 61, 62
Shade, C. K., 46
Shah, P., 31, 85
She, H., 45, 55, 61, 62, 76
Simon, D. P., 30, 43
Simon, H. A., 7, 13, 29, 30, 65, 73
Sims, V. K., 9, 29, 30, 31, 36, 37, 38, 41, 86, 104, 109
Snyder-Hogan, L. E., 10, 13, 44, 66
Sobel, M. E., 77, 81
Souvignier, E., 65, 67, 81, 83
Sungur, S., 44
Sutherland, S. W., 10, 15, 16, 25
Sweller, J., 7, 9, 15, 16, 27, 30, 33, 46, 82, 86, 104, 106, 109, 111
Szpunar, K. K., 75

Tekkaya, C., 44
Thomas, S., 55, 76, 87
Tix, A., 23
Tornatora, M. C., 45, 68
Treagust, D., 7, 14, 43, 65
Tuovinen, J. E., 66
Tversky, B., 16, 43

Van Der Werf, M. P. C., 65
Van Gog, T., 44, 45, 48, 50, 60, 62, 72
Van Hout-Wolters, B. H. A. M., 10, 44, 73, 86, 87
Van Merriënboer, J. J. G., 9, 45, 66, 86, 104, 109
Van Waes, L., 10, 28, 92
Veenman, M. V. J., 10, 44, 45, 63, 73, 86, 87, 106, 111

Walberg, H. J., 44
Wang, M. C., 44
Wenderoth, M. P., 14, 87
West, S. G., 23, 26
Wiebe, E. N., 15, 31, 45, 66
Willson, V., 68

Winn, W. D., 7, 9, 10, 13, 14, 15, 16, 17, 25, 27, 28, 29, 36, 37, 38, 40, 86, 104, 109
Wu, H., 31, 85
Zhao, X., 81, 83, 89
Zimmerman, B. J., 10, 66, 67, 68
Zito, J. R., 67

SUMMARY

Students in secondary Science education seem to have difficulties with comprehending diagrams. Process diagrams are an important type of representation in Biology for explaining processes like protein synthesis, compound cycles, etc. In this thesis, we aimed at getting deeper insight into students' difficulties comprehending process diagrams. These insights were used to inform the design of an intervention.

CHAPTER 1

In chapter 1, we define and describe process diagrams as a distinct type of diagram. Process diagrams are defined as representations of how systems function (e.g., photosynthesis) by abstractions in components and arrows and that contain information that is spatial, dynamic, and schematic in nature. Process diagram communicate information through the spatial organization of elements (e.g., components, arrows, labels), e.g., components in close proximity are more related. Arrows in a process diagram represent the dynamic functioning of a system, e.g., arrows might represent the amount of carbon flux per hour or biochemical reactions. Process diagrams are schematic in nature because they are simplified and symbolic representations of the real world.

Next, we present the aim and the research questions of this thesis. The aim of this thesis was to design an intervention that facilitates students' comprehending process diagrams. We designed three studies to examine students' difficulties with process diagrams. The general scheme was to study the interaction between three constituting elements: 1) the task, 2) the student, and 3) the diagram design. The first three studies provided insights that informed the design of the final intervention study. The research questions of this thesis are:

1) What is the relationship between features of the (1) task, (2) student, and (3) diagram, on the one hand, and difficulty of a diagram task, on the other hand?
2) What is the relationship between students' learning activities while studying various process diagrams, and their resulting comprehension of these diagrams?
3) What is the effect of a strategy training on learning from process diagrams.

The research questions are answered by four studies (chapter 2 to 5). The first research question is answered by two studies (chapter 2 and 3), the third study (chapter 4) answers the second research question and the fourth study (chapter 5) answers the third research question.

CHAPTER 2

In chapter 2, we focus on explanatory factors that predict students' difficulties with process diagrams. From 18 compulsory national Biology exams of secondary school

pre-university students all process diagram tasks (n = 64) were included in corpus. Features of the task, student, and diagram were related to the difficulty of that particular task, indicated by the cohort mean exam score. We defined two levels of cognitive task demand: tasks with a 'low' and tasks with a 'high' cognitive demand. Students' prior knowledge and familiarity with the components, arrows, and spatial arrangement of the process diagrams were coded by two experts. Diagram features, e.g., arrow and component type, were also coded.

A hierarchical regression analysis showed main effects for (1) the cognitive task demand, (2) the familiarity of the components, and (3) the number of components in a diagram. The cognitive task demand explained most of the variance (30%) in task difficulty. The familiarity of the components was also related to task difficulty. Diagrams that contain 'familiar' components showed higher mean exam scores and were therefore seen as less difficult. The number of components was negatively related to the task difficulty: The more components in a diagram the less difficult a task was. Information in a diagram might become more concrete if more components, i.e., bits of information, are added.

We also observed interactions. Within the category of tasks with a high cognitive demand, tasks about a diagram of which students have low prior content knowledge were more difficult than tasks about a diagram of which students have high prior content knowledge. Tasks with a high cognitive demand about a diagram with familiar arrows were, surprisingly, more difficult than tasks with a high cognitive demand about a diagram with unfamiliar arrows. This latter finding might be attributed to compensation for task difficulty by the large number of components in the diagrams involved.

The final model explained 46 percent of the variance in exam scores. The results of this study suggest that students have difficulties (1) with tasks that require a deeper understanding when the content is new, (2) with diagrams that use unfamiliar component conventions, and (3) with diagrams that have a small number of components and are therefore probably more abstract.

CHAPTER 3

In chapter 3, we focus on students' ability to solve process-diagram problems in Biology and the relationship with prior knowledge, spatial ability and working memory. For this purpose, we developed a test that represents process diagrams and adjacent tasks used in secondary education Biology. We included a total of 28 process diagrams on four biological topics, i.e., ecology, protein synthesis, dissimilation, and hormones. The process-diagram test consists of 97 tasks. Each topic of the process-diagram test contains tasks with a low cognitive demand and a high cognitive demand. Students' prior knowledge about the topics in the process-diagram test was measured by a test with 56 open and closed questions. Spatial ability and working memory tests were selected from the Ekstroms' kit of factor referenced cognitive tests (Ekstrom et al., 1976).

As hypothesized, scores on the prior knowledge test correlated positively with scores on tasks with a low and a high cognitive demand. We expected that prior knowledge would be related with tasks with a low cognitive demand because prior knowledge facilitates search (Winn, 1993) and with tasks with a high cognitive demand because knowledge schemata keep cognitive load low (Sweller, Van Merriënboer, & Paas, 1998).

Scores on a test that measured the ability to search for information in a complex spatial diagram were also, as hypothesized, positively correlated with scores on tasks with a low cognitive demand. Visualization and spatial orientation factors (Ekstrom et al., 1976) were, as hypothesized, not related (with exception of one of five tests for these factors) to tasks with a low or a high cognitive demand. This was expected because we assumed that process-diagram tasks do not require rotation or actually visualizing the movement of components.

Students' scores on a visual working memory test were, as hypothesized, positively related to tasks with a high cognitive demand. For tasks with a high cognitive demand, students need to build and explore an elaborate mental model that draws on the capacity of visuospatial working memory (Sims & Hegarty, 1997).

We conclude that the ability to solve process-diagrams problems involves the presence of prior knowledge, distinct spatial abilities, and visuospatial working memory capacity.

CHAPTER 4

In chapter 4, we examined students' learning activities while studying process diagrams, related to their resulting comprehension of these diagrams. Each student completed three learning tasks. Eye-tracking data was collected for all learning tasks: Verbal data was collected for the first two learning task by cued retrospective think-aloud protocol. Verbal data and eye-tracking data were collected as indications of students' learning activities. Students' comprehension was directly inferred from the verbal data from the first two learning tasks; comprehension of the third learning task was measured by a test.

For the verbal data, we applied a fine-grained coding scheme to optimally describe students' learning activities. The coding scheme distinguishes between an orientation phase and an elaboration phase. We defined three main categories of activities, i.e., cognitive, metacognitive and diagram learning activities. For the eye-tracking data, we used fixation time and transitions between areas of interest in the process diagrams as indices of learning activities.

Various learning activities were found that distinguished more and less successful students while studying process diagrams. Some distinct findings were that successful students were more likely to employ learning activities such as using the legend in the orientation phase; in the elaboration phase successful students more often give meaning to process arrows—80% between-student variance in comprehension score was predicted by the latter variable—and read the organizational levels. We also found that successful students were more likely to employ learning ac-

tivities such as activating prior knowledge and self-questioning, and that they spend more time in the main area of the process diagram. Spending more time in the main area of the process diagram predicted 65% between-student variance in comprehension score.

We also analyzed between-student and within-student between-diagram differences in behavior and comprehension score across learning tasks. Students employed successful learning activities consistently across learning tasks. Furthermore, compared to unsuccessful students, successful students used a more coherent approach of interrelated learning activities for comprehending a process diagrams.

CHAPTER 5

In chapter 5, we evaluated the effect of multiple-strategy training on learning from process diagrams. The training focused on a stepwise working-routine that included when and where to employ cognitive and metacognitive learning strategies and on affective strategies to invest effort in the implementation of this stepwise working-routine.

The study followed an experimental pretest-posttest design. Students (N = 180) were randomly assigned to the experimental or the control condition. A computer-based multimedia learning environment was designed to deliver the instructions for the pretest, intervention, and posttest phase. This multimedia environment was designed following the general guidelines for the design of multimedia instructional material (Mayer & Moreno, 2002). The pretest, intervention, and posttest were conducted in one single session in a classroom with computers in students' school.

The pretest phase consisted of a learning task followed by a test. The learning task was to "understand as much as you can" from a process diagram. When students finished studying the diagram, they had to rate their invested mental effort and perceived task difficulty before they could start with the pretest. The pretest contained eight multiple-choice items that tested students' comprehension of the process diagrams.

Students in the experimental condition were presented a multiple-strategy training based on cognitive strategy instruction (e.g., Harris & Graham, 1996). The training consisted of five phases: 1) emphasizing that diagrams are important and useful, 2) explaining the strategic approach, 3) providing a model of the strategic approach, 4) practicing with the strategic approach, and 5) providing feedback on practicing.

The posttest phase, like the pretest phase, consisted of a learning task followed by a test. Again, students first had to rate their invested mental effort and perceived task difficulty before they could start with the posttest.

Structured equation modeling was applied to examine the direct and indirect effects, through invested mental effort and perceived task difficulty, on learning from process diagrams. We observed an indirect positive significant effect of multiple-strategy training, through invested mental effort, on learning from process diagrams compared to the control group. The latter result might suggest that students in the experiment group were willing to invest more mental effort because of increased

strategy belief (Boekaerts, 1997). We observed no significant direct effect of multiple-strategy training on learning from process diagrams in the experiment group, compared to the control group.

CHAPTER 6

In the concluding chapter, we first present the main results of the four separate studies of this thesis and then some reflections upon the studies and directions for future research. Finally, implications for educational practice are presented.

We reflect on theories that supported this thesis. For this thesis, multimedia theories (Mayer, 2001; Schnotz & Bannert, 2003), theories that focus on learning from graphical representations (e.g., Hegarty, 2005), and cognitive load theory (Sweller, 1988; Sweller, 1994; Sweller et al., 1998) provided a framework on how students construct a mental model of an external representation, i.c., process diagram. Furthermore, research on self-regulated learning (Boekaerts, 1997) and (meta)cognition (Veenman, 2012) also supported the present thesis. We argue that the present thesis benefitted from the afore-mentioned theories; they served as a base to get deeper insight into students' difficulties with process diagrams. This thesis adds to previous research a detailed analysis of students' difficulties with a specific type of external representations, i.c., process diagrams.

Furthermore, we reflect on issues that concern how variables were defined and measured, ecological validity and generalization. We also discuss how the first three studies of this thesis have informed the final intervention study.

Finally some practical recommendations are presented. We think this thesis might help stakeholders of the Biology education community (e.g., students, teachers, teacher educators, instructional designers) to make informed decisions with regard to process diagrams when they teach, select or design learning materials, design exams, etc. We argue that students need more support when they have to comprehend process diagram. This thesis identified several conditions that teachers should consider when their students have to learn or solve problems with process diagrams. Teacher educators might inform prospective teachers about students' difficulties with process diagrams. Furthermore, this thesis might help instructional designers by making more informed decisions when selecting or designing process diagrams (and adjacent tasks and instructions) for educational materials.

SUMMARY (DUTCH)

Studenten hebben moeilijkheden met het begrijpen van diagrammen. Procesdiagrammen vertegenwoordigen een belangrijk type diagram in de biologie voor het beschrijven van processen zoals eiwitsynthese, kringlopen, etc. Het doel van dit proefschrift was om dieper inzicht te krijgen in de moeilijkheden die studenten ervaren met het begrijpen van procesdiagrammen. Deze inzichten zijn gebruikt voor het ontwerpen van een interventie.

HOOFDSTUK 1

In het eerste hoofdstuk worden procesdiagrammen als een specifiek type diagram gedefinieerd en beschreven. Procesdiagrammen zijn gedefinieerd als representaties van hoe een systeem werkt door middel van abstracties in componenten en pijlen. Tevens communiceren procesdiagrammen ruimtelijke, dynamische en schematische informatie. Procesdiagrammen geven ruimtelijke informatie weer door de organisatie van de elementen (e.g., componenten, pijlen, labels), e.g., elementen die dichtbij elkaar staan, zijn nauwer met elkaar verbonden. Pijlen representeren het dynamisch functioneren van een systeem, e.g., pijlen die de hoeveelheid koolstofstroom per uur of biochemische reacties weergeven. Procesdiagrammen zijn schematisch omdat ze een vereenvoudigde en symbolische weergave van de werkelijkheid zijn.

Vervolgens beschrijven we het doel en de onderzoeksvragen van dit proefschrift. Het doel van dit proefschrift was om een interventie te ontwerpen die studenten ondersteunt bij het begrijpen van procesdiagrammen. We hebben drie studies ontworpen om de problemen die studenten hebben met het begrijpen van procesdiagrammen in kaart te brengen. Deze studies richten zich op de interactie tussen 1) de taak, 2) de student, en 3) het ontwerp van het diagram. Deze studies gaven inzicht in de moelijkheden die studenten hebben en gaven input voor het ontwerp van de uiteindelijke interventiestudie. De onderzoeksvragen van dit proefschrift zijn:

1) Wat is de relatie tussen de eigenschappen van de (1) taak, (2) student, en (3) diagram, enerzijds, en de moeilijkheid van een diagramtaak, anderzijds?
2) Wat is de relatie tussen de leeractiviteiten van studenten bij het bestuderen van een procesdiagram en de hieraan gerelateerde begripsvorming van dit diagram?
3) Wat is het effect van een training in strategieën op het leren van procesdiagrammen?

De onderzoeksvragen worden in vier studies beantwoord (hoofdstuk 2 t/m 5). De eerste twee studies (hoofdstuk 2 en 3) beantwoorden de eerste onderzoeksvraag, de derde studie (hoofdstuk 4) beantwoordt de tweede onderzoeksvraag en de vierde studie (hoofdstuk 5) beantwoordt de derde onderzoeksvraag.

HOOFDSTUK 2

In hoofdstuk 2 onderzoeken we factoren die moeilijkheden van leerlingen met procesdiagrammen verklaren. Het onderzoeksmateriaal bestond uit achttien centrale examens Biologie vwo. Uit deze examens zijn alle taken meegenomen die een interpretatie van een procesdiagram (n = 64) vereisen. Eigenschappen van de taak, leerling en diagram werden gerelateerd aan de moeilijkheidsgraad van de diagramtaak (d.i. 1 - de gemiddelde cohortscore). We definieerden twee niveaus van cognitieve taakbelasting: taken met een 'hoge' en taken met een 'lage' cognitieve belasting. De veronderstelde voorkennis van studenten en hun bekendheid met componenten, pijlen en de ruimtelijke rangschikking van de procesdiagrammen zijn gecodeerd door experts. Eigenschappen van het diagram, e.g., type pijl en component, zijn ook gecodeerd.

Een hiërarchische regressieanalyse toonde aan dat (1) de cognitieve taakbelasting, (2) de bekendheid met de componenten en (3) het aantal componenten in een diagram verschillen verklaarden in leerlingscores. De cognitieve taakbelasting verklaarde de meeste variantie (30%) van de moeilijkheidsgraad van de diagramtaak. De bekendheid met de componenten droeg ook bij aan de verklaring van de moeilijkheidsgraad. Diagrammen met 'bekende' componenten werden als minder moeilijk ervaren. Ook het aantal componenten droeg bij aan de moeilijkheidsgraad van de diagramtaak: Hoe meer componenten hoe eenvoudiger de taak. Dit effect zou verklaard kunnen worden doordat de informatie in een diagram meer concreet wordt als er meer informatiedelen aan worden toegevoegd: Het diagram wordt minder ambigue en de kans op foutieve interpretatie wordt kleiner.

Verder werden er interactie-effecten gevonden van de cognitieve taakbelasting met eigenschappen van de leerling en het diagram. Post-hoc analyse van de interactie tussen cognitieve taakbelasting en voorkennis toonde een significant verschil aan in de moeilijkheidsgraad van de diagramtaken bij een 'hoge' cognitieve taakbelasting en de mate van aanwezige inhoudelijke voorkennis. Post-hoc analyse van de interactie tussen cognitieve taakbelasting en bekendheid met de pijlen liet een negatief effect zien op de moeilijkheidsgraad van de diagramtaak bij taken met een 'hoge' cognitieve belasting wanneer de pijlen 'onbekend' zijn. Dit effect lijkt op het eerste gezicht een beetje vreemd maar een nadere analyse laat zien dat twee van de drie diagrammen met een 'onbekende' pijlen veel meer componenten dan gemiddeld bevatten. Deze toename in de concreetheid door de grote hoeveelheid componenten zou mogelijk kunnen compenseren voor het effect van 'onbekende' pijlen.

Het definitieve regressiemodel verklaarde 46 procent van de variantie in de examenscores. De resultaten van deze studie suggereren dat studenten moeilijkheden hebben (1) met taken die een dieper inzicht vragen wanneer de inhoud nieuw is, (2) met diagrammen die onbekende componenten bevatten, en (3) met diagrammen met een klein aantal componenten waardoor ze waarschijnlijk meer abstract zijn.

HOOFDSTUK 3

In hoofdstuk 3 meten we de bekwaamheid van studenten ten aanzien van het oplossen van problemen met procesdiagrammen in het biologieonderwijs en de relatie met voorkennis, ruimtelijke vaardigheden en werkgeheugen. We hebben een test ontwikkeld met 28 procesdiagrammen over vier biologische onderwerpen, i.e., ecologie, eiwitsynthese, dissimilatie en hormonen. De test bestaat uit 97 taken. Elk onderwerp uit de test bevat taken met een 'lage' en 'hoge' cognitieve belasting. De voorkennis van de studenten is gemeten door middel van een test met 56 open en gesloten vragen. Testen voor ruimtelijke vaardigheden en werkgeheugen zijn afkomstig uit *Ekstrom's kit of referenced cognitive tests* (Ekstrom et al., 1976).

Scores op de voorkennistest correleren positief, zoals verwacht, met scores op taken met een 'lage' en 'hoge' cognitieve belasting. We hadden verwacht dat scores op de voorkennistest zouden correleren met scores op taken met een 'lage' cognitieve belasting omdat voorkennis de zoekopdracht vergemakkelijkt (Winn, 1993); voorkennis vergemakkelijkt taken met een 'hoge' cognitieve belasting omdat de aanwezigheid van kennisschema's de cognitieve belasting laag houden (Sweller, 1988; Sweller, 1994; Sweller, Van Merriënboer, & Paas, 1998). Scores op een test die de bekwaamheid meet om informatie te zoeken in een complex diagram (Ekstrom et al., 1976) waren, zoals verwacht, ook positief gecorreleerd met scores op taken met een 'lage' cognitieve belasting.

Scores op testen die visualisatie en ruimtelijke oriëntatie meten waren, zoals verwacht, niet gerelateerd (met uitzondering van één van de vijf testen) met taken met een 'lage' en 'hoge' cognitieve belasting. Dit was verwacht omdat we aannamen dat procesdiagramtaken geen rotatie en visualisatie vereisen. Scores op een test voor ruimtelijk werkgeheugen waren, zoals verwacht, positief gerelateerd met taken met een 'hoge' cognitieve belasting. Voor taken met een 'hoge' cognitieve belasting moeten studenten een uitgebreid mentaal model construeren en onderzoeken en dit vereist capaciteit van het visueel-ruimtelijk geheugen (Sims & Hegarty, 1997).

We concluderen dat de bekwaamheid van studenten ten aanzien van het oplossen van problemen met procesdiagrammen gerelateerd is aan de aanwezigheid van voorkennis, specifieke ruimtelijke vaardigheden, en capaciteit van het visueel-ruimtelijk geheugen.

HOOFDSTUK 4

In hoofdstuk 4 onderzoeken we de leeractiviteiten van studenten tijdens het bestuderen van procesdiagrammen en hun begripsvorming van dit diagram. Elke student maakte drie leertaken. We verzamelden eye-tracking data tijdens alle leertaken én verbale data tijdens de eerste twee leeractiviteiten door middel van het *cued retrospective* hardop-denk protocol. Eye-tracking data en verbale data werden verzameld als indicatoren voor de leeractiviteiten van de studenten. De begripsvorming van het procesdiagram werd direct afgeleid van de verbale data voor

de eerste twee leertaken; begripsvorming van het procesdiagram van de derde leertaak werd gemeten met een posttest.

De verbale data zijn geanalyseerd met een fijnmazig codeerschema om de leeractiviteiten van student te beschrijven. In het codeerschema onderscheiden we een oriëntatie- en een uitwerkingsfase. We definieerden drie hoofdcategorieën voor activiteiten, i.e., cognitieve, metacognitieve en diagram leeractiviteiten. De eye-tracking data zijn gebruikt voor het berekenen van de tijdsduur van de fixaties en transities tussen *areas of interest* als indicatoren van leeractiviteiten van studenten. Er zijn verschillende leeractiviteiten gevonden die succesvolle studenten vaker ontplooiden tijdens het bestuderen van procesdiagrammen. Enkele duidelijke verschillen waren dat meer succesvolle studenten vaker leeractiviteiten in de oriëntatiefase ontplooiden zoals de legenda gebruiken; in de uitwerkingsfase gaven meer succesvolle studenten vaker betekenis aan de pijlen van het procesdiagram—deze variabele verklaarde 80% van de variantie op de score van begripsvorming tussen studenten—en las deze groep studenten vaker de labels met de organisatieniveaus. We vonden ook dat meer succesvolle studenten vaker leeractiviteiten ontplooiden zoals het activeren van voorkennis en het zichzelf vragen stellen en dat ze meer tijd besteedden aan het kijken naar het gedeelte van het diagram waar het proces wordt weergegeven. Laatstgenoemde variabele verklaarde 65% van de variantie op de score van resulterend begrip tussen studenten.

We hebben ook de *between-student* en de *within-student between-diagram* verschillen in gedrag en resulterend begrip geanalyseerd. Studenten ontplooiden de leeractiviteiten consistent over leertaken. Verder gebruikten succesvolle studenten, vergeleken met minder succesvolle studenten, een meer coherente aanpak van samenhangende leeractiviteiten voor het begrijpen van procesdiagrammen.

HOOFDSTUK 5

In hoofdstuk 5 evalueren we het effect van een training met meerdere strategieën voor leren van procesdiagrammen. De training bevat een stapsgewijze werkwijze en bevat aanwijzingen waar en wanneer cognitieve en metacognitieve leerstrategieën ontplooid moeten worden. Tevens bevat de training affectieve strategieën om de studenten te stimuleren om inspanning te besteden aan het implementeren van deze stapsgewijze werkwijze.

De studie heeft een experimenteel pretest-posttest design. Studenten (N = 180) werden willekeurig toegewezen aan de experimentele- of de controleconditie. De instructies voor de pretest-, interventie-, en posttestfase werden geleverd door een digitaal multimedia leeromgeving. Deze multimedia leeromgeving was ontwikkeld volgens de richtlijnen voor het ontwerpen van multimedia-instructiemateriaal (Mayer & Moreno, 2002). De pretest, interventie en de posttest werden uitgevoerd in een enkele sessie in een computerlokaal op de school van de studenten.

De pretestfase bestond uit een leertaak gevolgd door een test. De leertaak was om ‘zoveel mogelijk te begrijpen’ van een procesdiagram. Wanneer studenten klaar waren met het bestuderen van het diagram dan moesten ze aangeven hoeveel

mentale inspanning ze hadden verricht en hoe moeilijk ze de taak vonden. Vervolgens konden de studenten beginnen met de pretest.

De posttest bestond uit acht meerkeuzevragen en testte hoeveel studenten van het procesdiagram begrepen hadden. De studenten in de experimentele conditie kregen een training aangeboden met meerdere strategieën gebaseerd op het model van cognitieve strategie instructie (e.g., Harris & Graham, 1996). De training bestaat uit vijf fasen: 1) benadrukken dat diagrammen belangrijk en nuttig zijn, 2) uitleg van de strategische aanpak, 3) strategische aanpak demonstreren met behulp van een model, 4) oefenen met de strategische aanpak en 5) terugkoppeling op het oefenen.

De posttestfase bestond, net als de pretestfase, uit een leertaak en een test. Wederom moesten de studenten, wanneer ze klaar waren met het bestuderen van het diagram, aangeven hoeveel mentale inspanning ze hadden verricht en hoe moeilijk ze de taak vonden. Vervolgens konden de studenten beginnen met de posttest.

Met behulp van een *structured equation model* zijn de directe en indirecte effecten, door mentale inspanning en ervaren moeilijkheid van de taak, op leren van een procesdiagram onderzocht. We zagen bij de experimentgroep, vergeleken met de controlegroep, een indirect positief significant effect, door mentale inspanning, op leren van een procesdiagram. Laatstgenoemde zou kunnen suggereren dat studenten in de experimentgroep bereid waren om meer mentale inspanning te leveren omdat ze overtuigd waren van de strategie (Boekaerts, 1997). We zagen geen direct effect van de training op het leren van procesdiagrammen bij de experimentgroep, vergeleken met de controlegroep.

HOOFDSTUK 6

In het afsluitende hoofdstuk presenteren we eerst de belangrijkste resultaten van de vier afzonderlijke studies in dit proefschrift, vervolgens enkele reflecties op de studies en richtingen voor toekomstig onderzoek. Ten slotte worden de gevolgen van dit proefschrift voor de onderwijspraktijk besproken.

We reflecteren op studies die dit proefschrift ondersteunen. Theorieën over multimedialeren (Mayer, 2001; Schnotz & Bannert, 2003), theorieën die zich richten op het leren met grafische representaties (e.g., Hegarty, 2005), en de *cognitive load* theorie (Sweller, 1988; Sweller, 1994; Sweller, et al., 1998) dienden als raamwerk voor hoe studenten een mentaal model construeren van een externe representatie, i.c., een procesdiagram. Verder wordt dit proefschrift ondersteund door onderzoek dat zich richt op zelfgereguleerd leren (Boekaerts, 1997) en (meta)cognitie (Veenman, 2012). We beweren dat dit proefschrift heeft geprofiteerd van de voorgenoemde theorieën; ze dienden als een basis om dieper inzicht te krijgen in de moeilijkheden die studenten hebben met procesdiagrammen. Dit proefschrift voegt aan de bestaande kennis toe, een gedetailleerde analyse van de problemen die studenten ervaren met een specifiek type representatie, i.c., het procesdiagram.

Verder reflecteren we op kwesties die betrekking hebben het definiëren en meten van de variabelen, ecologische validiteit en generalisatie. We bespreken ook hoe de

eerste drie studies van dit proefschrift hebben bijgedragen aan het ontwerp van de interventiestudie.

Ten slotte bespreken we enkele praktische aanbevelingen. We denken dat dit proefschrift de belanghebbenden van de biologie-onderwijsgemeenschap (e.g., studenten, leraren, onderwijsontwikkelaars) kan ondersteunen bij het maken van meer gefundeerde beslissingen wanneer ze lesgeven, leermateriaal selecteren of ontwikkelen, examens maken, etc. We beweren dat studenten meer ondersteuning nodig hebben bij het begrijpen van procesdiagrammen. Dit proefschrift heeft een aantal voorwaarden vastgesteld die leraren in overweging kunnen nemen wanneer hun studenten moeilijkheden ondervinden met het begrijpen van procesdiagrammen. Lerarenopleiders zouden toekomstige leraren op de hoogte moeten stellen van de problemen die studenten hebben met procesdiagrammen. Verder ondersteunt dit proefschrift onderwijsontwikkelaars om beter gefundeerde beslissingen te maken wanneer ze procesdiagrammen (en aanverwante taken en instructie) selecteren of ontwerpen voor onderwijsmateriaal.

ACKNOWLEDGEMENTS

First of all I would like to thank my supervisors—Gert Rijlaarsdam and Wilfried Admiraal. Trust is the first thing I think of when I reflect on the guidance you both gave me during this research project. You've allowed me to work on this thesis in a way that fitted my needs and interests. Our collaboration could be characterized by outbursts of working together interspersed with long silences. I was often surprised how you both found the time during these outbursts to give me your full attention and I've learned a lot from our constructive and high quality discussions during these times. The 'long silences' were also respected; you knew I like to sort things out on my own and trusted that work was getting done.

Next, I would like to thank all teachers, support staff, and school managers from O.S.G. West-Friesland. I am especially grateful to Jos Roebroek and the other members of the Biology section who helped me throughout this research project with advise and coding of data. I also would like to thank Ewald Weiss and Peter Snoek—my former executives at the O.S.G. West-Friesland—for their support in creating the conditions for this project.

I am grateful to René Alberts and his colleagues from the National Institute for Educational Measurement (CITO) for making the data available for the first study. I am also grateful to the students from the O.S.G. West-Friesland who volunteered to participate in the second and third study. My special thanks goes to the teachers and their students from the 7 schools who volunteered to participate in the final study of this thesis. I also would like to thank Douwe van der Kooi from the Amsterdam University of Applied Sciences—my current employer—for giving me the opportunity to work on the finalization of this research project.

Furthermore, I would like to thank my parents—Wim and Diny—and my parents-in-law—Henk and Elly—for helping out in several ways. Finally, my special thanks goes to my partner—Boëlla—for it would have been impossible for me to finish this research project without your support while raising a family with three children—Stijn, Isabel, and Joris.