In the Eye of an Expert

Conveying Perceptual Skills in Biological and Medical Domains

via Eye Movement Modeling Examples

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Dipl.-Psych. Halszka Maria Jarodzka

aus Wrocław, Polen

Tübingen

2010

Tag der mündlichen Qualifikation:	20.04.2011
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Peter Gerjets
2. Berichterstatter:	Prof. Dr. Stephan Schwan

Dedication

To my grandmother, Elżbieta Sapalska, who always believed in me.

I would like to thank...

...my supervisors and reviewers

Peter Gerjets, Katharina Scheiter, Stephan Schwan, and Barbara Kaup

...my former colleagues from the KMRC for an enriching time

Yvonne Kammerer, Birgit Imhof, Maria Opfermann, Gabrielle Cierniak, Tim Kühl, Anne Schüler, Alexander Eitel, Erlijn van Genuchten, Markus Armbruster, Vera Bauhoff, Juergen Buder, Kurt Langenbacher, Waltraud Lenz, Rosemarie Croizier, and many more

...my cooperation partners for their contribution to this thesis

Michael Dorr, Marcus Nyström, Thomas Balslev, and Berit Eika

Most of all I would like to thank Tamara van Gog and Kenneth Holmqvist,
who both taught and supported me so much that they became my mentors and friends.

...my new colleagues from CELSTEC for their support in the final stage of my thesis

Els Boshuizen, Jo Boon, Wendy Kicken, Mieke Haemers, Iwan Wopereis, Milou de Smet, Kees Pannekeet, Chantal Gorissen, Eniko Bezdan, Olga Firssova, Audrey Wigman, Ingrid Jonkman, Christa van Mierlo, Sandra Wetzels, Kim Dirkx, Renate de Groot, Saskia Brand-Gruwel, Wilfred Rubens, Martin van Dijk, Jeroen Berkhout, Ine Verstappen, Ludo van Meeuwen, Bettine Taminiau, and many more

Most of all I would like to thank Paul Kirschner for his unconditional support.

...my family and friends

Małgorzata Piegsa, Jan, Jakub and Julian Sowula, Elżbieta Sapalska, Waldemar, Karol, and Filip Sapalski, Marek, Piotr, Janina, Leon and all other Zenger family members, Margit and Karl-Ludwig Drumm, Kim Stalbovs, Melanie Wahl, and last but not least Paweł Jarodzki

Most of all I would like to thank Christian Drumm for his unconditional love.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Overview: Challenges in Teaching Perceptual Skills in Biological and Medical Domains

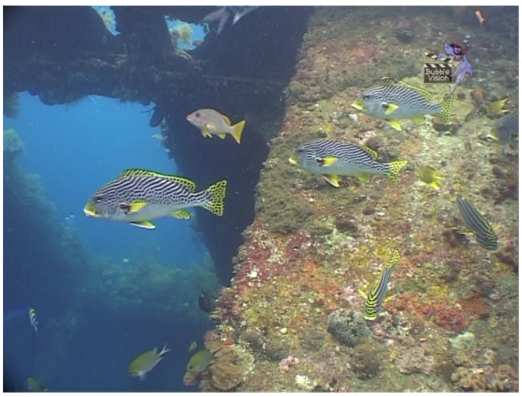
1.1 Introduction

Research on cognitive skill acquisition has provided numerous insights concerning the question of how people acquire and apply cognitive skills required for solving problems in particular in well-structured domains (e.g., mathematics, physics; cf. Van Lehn, 1996). In this research, to-be solved tasks are typically stated in a symbolic format, for instance, by presenting word problems or mathematical equations. However, this research so far has mostly ignored the fact that many real-world tasks require a problem solver to process and interpret complex, visual input (e.g., X-rays), thereby posing strong demands on perception (i.e., perceptual tasks, cf. Chase & Chi, 1981; Chi, 2006a). Hence, in these tasks people cannot reach high performance without extracting and interpreting relevant visual features. These features are not always of static, but may also be of dynamic nature. For instance, in biological or medical practice – that is, the domains investigated in the context of this thesis – often *motion patterns* have to be interpreted. This is the case in many areas of biological description of behavioral patterns (e.g., role of monkeys' facial patterns for breeding, Kingdon, 1980; or instinctive motion patterns of geese in rolling their eggs, Lorenz & Tinbergen, 1939) and in medical diagnosis (e.g., diagnosis of Parkinson – a neurological movement disorder, Jankovic, 2008; or Bell's palsy – paralysis due to the inflammation of a facial nerve, Ahmed, 2005). To conclude, many real-world tasks require not only cognitive, but also *perceptual skills* related to the processing of dynamic visual input.

In domains other than biology and medicine, in which motion patterns play a crucial role, like sports, it has been shown that experts compared to novices perceive motion patterns more efficiently and draw more correct conclusions based on them (e.g., Moreno, Reina, Luis, & Sabido, 2002). Hence, experts may possess more sophisticated perceptual skills than novices in these domains. This is also very likely to be the case in biological description of behavioral patterns and medical diagnosis that rely on recognizing and interpreting motion patterns. It is likely that these perceptual skills do not develop automatically, but rather need to be trained. Hence, conveying skills to extract and interpret motion patterns should be in the focus of instruction. The goal of this thesis is to analyze differences of experts and novices in terms of their underlying perceptual skills and based on these differences to develop a novel instructional method that is suited to convey these skills.

In this thesis, this approach will be illustrated for two tasks, namely the classification of fish locomotion patterns in marine zoology and the diagnosis of neurological diseases in pediatrics. In both domains specific motion patterns are important characteristics for classification and diagnosis. Hence, recognizing and interpreting these motion patterns is of central importance in these tasks. For instance, in marine zoology identification keys are used for species classification. In these classification keys references are made to how fish are moving as relevant features for species classification (Lieske & Myers, 1994). Similarly, diagnostic systems in pediatric neurology rely on motion patterns for the diagnosis of different types of epileptic seizures (cf. International League Against Epilepsy, 2010; Loddenkemper et al., 2005; Lüders et al., 1998). The following two paragraphs will describe in more depth that the ability to recognize motion patterns is important in the practice of these two tasks.

The first task deals with fish locomotion pattern classification in marine zoology. A marine zoologist is engaged in investigating and classifying fish species. This has to be done in both, in the laboratory as well as in the natural environment of the living fish. In the natural environment, motion patterns of fish are very prominent to the observer and are therefore often used for species classification. Imagine diving in the Indian Ocean (for a screenshot from a documentary see Figure 1.1(a)). You see a large number of colorful fish swimming in a colorful reef. Probably you would not be able to detect and remember specific fin shapes, spikes, color patches, or to count the number of gills of a single fish. Nevertheless, these features are all potentially relevant for species classification according to field guide books (e.g., Lieske & Myers, 1994). Detecting these features is especially difficult since fish quickly swim away from divers to a distance in which – under water – color and small details are difficult to identify. This is due to the fact that water acts as a strong filter for color (Emmerson & Ross, 1985; Kinney, Luria, & Weitzman, 1967). Motion, on the other hand, can be even recognized under conditions of otherwise low visibility. Moreover, most fish move constantly; therefore, their motion patterns constitute information that is always available to an observer – unlike color or the shape of a specific fin. Hence, it seems to be appropriate that classification keys not only address static, but also dynamic features, like these motion patterns, since, especially in realistic contexts the latter information may be more readily available.

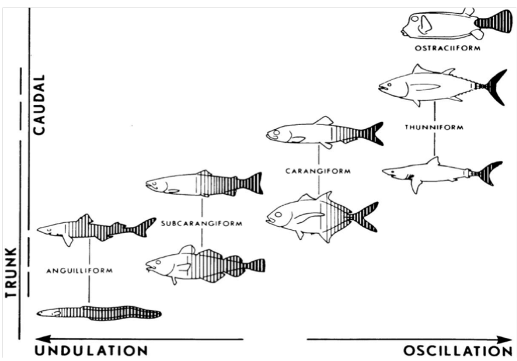


(a) Example of a real-life scene of reef fish. From *Diving in Bali,* by N. Hope, 2006, *Bubble Vision*.

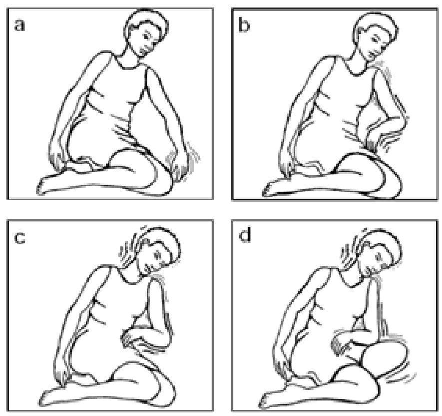
(b) Medical diagnosis based on motion patterns. Example from clinical practice (Image by T. Balslev).

Figure 1.1: Biological classification and medical diagnosis in practice.

The second task addressed in this thesis concerns the diagnosing of neurological diseases in pediatrics. In clinical practice motion patterns of patients are very prominent to the physician and are central to state a diagnosis of neurological diseases (Dreifuss & Nordli, 2000; Nordli, 2002; Nordli, Bazil, Scheuer, & Pedley, 1997). Imagine being in a pediatric ward of a hospital. A several months old girl lies in bed crying, clearly feeling uncomfortable (Figure 1.1(b)). She waves her arms and legs around and now and then she gets a muscle spasm and shortly looses consciousness. Those spastic and facial movements are clear symptoms of a neurological disease. Hence, it has been recommended to diagnose epileptic seizures in infants according to their ictal (i.e., seizure induced) motion patterns (Nordli, 2002).



(a) Example of an identification key for forward swimming modes in fish. From *Form, function, and locomotory habits in fish*, by C. C. Lindsey, 1978, *Fish Physiology, 7*, 1-88.



(b) Medical diagnosis based on static features. From www.epilepsyfoundation.org.

Figure 1.2: Biological classification and medical diagnosis in current education.

In contrast to how important, but difficult recognizing motion patterns in the above mentioned tasks may be (in biological classification see Law & Lynch, 1990; in medical diagnosis see Nordli et al., 1997), only little attention has been paid to the question of how to train underlying perceptual skills (for exceptions in biology didactics see Pfeiffer, Gemballa, Jarodzka, Scheiter, & Gerjets, 2009; for exceptions in medical education see Balslev, De Grave, Muijtjens, & Scherpbier, 2005). Rather, in educational textbooks for both tasks motion patterns are conveyed like conceptual knowledge. The motion patterns are described verbally in both, biological classification ("Filefish slowly swim by a wavelike motion of their second dorsal and anal fin", Lieske & Myers, 1994, p. 340) and medical diagnosis ("The patient opens his eyes with a frightened expression; the right arm is rotated inwards and stiff. He seems to be trying to ward off something menacing. He is screaming loudly throughout the episode, lasting 1 min"; D. Schmidt & Schachter, 2000, p. 10). Those written descriptions are often accompanied by static pictures. These pictures either provide a rather abstract representation of the underlying movement (e.g., by shading the part of the fish body that contributes to its forward propulsion, cf. Figure 1.2(a)) or they show only one particular state that occurs in the course of the movement (e.g., a typical body position of a child during an epileptic seizure; cf. Figure 1.2(b)). Verbal descriptions and static pictures may be suited to convey conceptual knowledge relevant to the task at hand, however, this knowledge does not guarantee that a person will be able to *recognize and interpret* these motion patterns in a living fish or infant.

Accordingly, written verbal descriptions and static illustrations appear to be not optimal to convey perceptual skills for accomplishing tasks that require extracting and interpreting visual information (in marine zoology: Gerjets et al., 2010; Imhof, Scheiter, & Gerjets, 2009; Pfeiffer, Scheiter, Kühl, & Gemballa, 2010; in medical education Balslev et al., 2005; Balslev, De Grave, Muijtjens, Eika, & Scherpbier, 2009). Thus, an instructional method that specifically aims at conveying such perceptual skills is likely to substantially improve current methods used in biological and medical education. Accordingly, the core aim of this thesis is to develop an effective instructional method to convey perceptual skills in both example tasks introduced, namely fish locomotion pattern classification and diagnosing neurological diseases in infants.

As methods to convey perceptual skills hardly exist, the developed instructional method is inspired by methods that have been used effectively to teach cognitive skills, namely, learning from classical worked examples (Sweller, Van Merriënboer, & Paas, 1998) and cognitive modeling (Collins, Brown, & Newman, 1989). Both approaches assume that it is more efficient to study the task performance shown by a skilled person (e.g., by reading a description of the solution procedure for a specific task or by observing a skilled person performing the task directly) than learning by unassisted problem-solving. The instructional method that is developed and evaluated in this thesis consists of two components, namely first, a domain expert's verbal explanations (i.e., modeling of cognitive processes) she or he makes while performing the task. Second, it consists in a video of a dynamic stimulus (e.g., a fish swimming or an infant moving) with the expert's eye movements superimposed onto this video (i.e., modeling of perceptual processes). These eye movements are supposed to provide information on how experts accomplish the task on a perceptual level and are assumed to guide learners' visual attention. Together, the expert's verbal explanation and her or his eye movement recordings superimposed onto a video are called *eye movement modeling examples* (EMME; Van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009).

1.2 Overview of this Thesis

To develop EMME, the tasks used in this thesis first have to be analyzed in depth to get a better understanding of how they are performed and where potential difficulties in performing them may occur. The following chapter (Chapter 2) presents different ways of analyzing a task and how these different types of task analysis may inform the design of instruction. Based on this overview a task analysis method is chosen and applied in this thesis, namely the so-called ”development of a task model” by Schaafstal and Schraagen (1992). This task analysis method consists of describing the stepwise procedure of the task, identifying the knowledge and skills required to perform the task, predicting difficulties that are likely to occur when performing the task, and drawing conclusions concerning instruction for this task. Chapter 3 presents a two-step analysis of the task to classify of fish locomotion patterns. It starts with a normative task analysis that reveals how the task *should* be performed (Chapter 3.1). This analysis is done with the help of domain experts and based on textbooks from the domain of marine zoology. In the second step an empirical task analysis is conducted (Chapter 3.2). Its goal is to verify and supplement the conclusions based on the normative analysis. This is done by comparing experts and novices while classifying fish motion patterns on a performance and process level. These processes are both of cognitive and perceptual nature. The main findings show that experts perform the task faster and more accurately compared to novices by attending earlier to relevant areas of the stimulus and using shortcuts, while novices get distracted by irrelevant, but salient features. In sum, these results confirm and supplement the findings of the normative task analysis.

In Chapter 4 the insights from the normative and the empirical task analyses are then used to develop EMME with the aim of conveying perceptual skills for the task of classifying fish locomotion patterns. EMME is derived from research on learning from classical worked examples (Sweller et al., 1998) and cognitive modeling (Collins et al., 1989). In Chapter 5, Study 2 of this thesis investigates two different versions of displaying the expert's eye movements in EMME in terms of how effective they guide students' attention and support different aspects of students' learning in comparison to a control group, where no eye movements are shown to learners. The main findings of the study suggest that students using EMME acquire perceptual skills that enable them to efficiently search and visually inspect relevant body parts of fish and to interpret their movement patterns. Interestingly, both variants of EMME foster different aspects of learner performance: Displaying the expert's eye movements as a spotlight fosters visual search, while displaying the expert's eye movements as solid dots enhances the interpretation of motion patterns. Since, none of these two versions appears to be able to support visual search and performance, one conclusion from Study 2 is that EMME need to be further optimized.

To test the generalizability of EMME to other domains it is applied to the second example task: diagnosing epileptic seizures in infants. To do so, in Chapter 6 a task analysis is performed that justifies the applicability of this method to the task. Based on this task analysis in Chapter 7 an adapted version of EMME is implemented in two improved versions to teach perceptual skills. The results of Study 3 largely confirm the findings from Study 2, in that EMME enable students to efficiently visually search the relevant body parts of the infants and interpret their motion patterns correctly. In particular, the spotlight version of EMME that is optimized compared to the version used in Study 2 enhances perceptual skills as evident at both, the process and the performance level compared to the alternative EMME version and a control group.

Finally, Chapter 8 summarizes and discusses the findings and draws conclusions for future research.

Part I

Task Analysis as a Theoretical Framework for Developing Instructional Support

Chapter 2

Task Analysis as a Prerequisite for Effective Instructional Design

As a first step in developing a novel instructional method I propose to conduct a detailed task analysis of the learning task to be acquired. An analysis of the task is necessary to define (1) all task steps that need to be executed to successfully perform the task, (2) knowledge and skills that are necessary to achieve each task step, (3) difficulties that are likely to occur when performing this task, and thus, (4) to define what should be in focus of an instructional training – a procedure that has been suggested by several researchers (e.g., Clark, Feldon, Van Merriënboer, Yates, & Early, 2008; Jonassen, Tessmer, & Hannum, 1999; Schraagen, 2006).

Task analysis is defined as specifying what a person is required to do, in terms of physical or cognitive processes, to achieve a goal, that is, in performing a task (Kirwan & Ainsworth, 1992; Smith & Ragan, 1999). A task analysis may be conducted for different reasons, like developing tests for the selection of personnel, deciding whether a task should be accomplished by humans or by machines, or supporting people's work by developing computer systems (Schraagen, 2006). One reason that emerged very fast when researchers and practitioners began to apply task analysis methods is to use its outcomes for trainings and the design of training equipment (Flanagan, 1954; Miller, 1962). And still nowadays –

according to Schraagen (2006) – task analysis focuses on understanding expert knowledge and performance to leverage that understanding into methods for training. Jonassen et al. (1999) explicitly stress the necessity of a task analysis for developing effective instructional material. Without a task analysis prior to instructional design, important components of the task or resources necessary to perform it might be missing in the training. With respect to instructional design "task analysis is a process of analyzing and articulating the kind of learning that you expect the learners to know how to perform" (Jonassen et al., 1999, p. 3).

2.1 Brief Overview of Task Analysis Methods

In basic psychological research usually the researcher creates a task that fits her/his research focus. In more applied research the task is predetermined by the job people perform (Schaafstal & Schraagen, 2000), like in the example of marine zoology or medical diagnosis. To investigate the given task a task analysis method has to be chosen. To choose the appropriate task analysis method one has to consider that there should be a close connection between the focus of the task analysis and the actions and processes required to perform the task. Most important, one has to consider whether these actions and processes are of overt (i.e., directly observable by the task analyst) or of covert nature (i.e., inert states of the task performer, like thoughts or the allocation of visual attention, that are not directly observable by the task analyst).

A number of well established methods exist that focus on the analysis of tasks requiring overt actions, like the manual control of machines (e.g., *time study* by Taylor, 1998/1911; *motion study* by Gilbreth (1920, as cited in Schraagen, 2006); *job analysis* by Drury, Paramore, Van Cott, Grey, & Corlett, 1987; Luczak, 1997; Meister, 1999).

Many methods are also available to analyze tasks that require mainly overt actions, but also some crucial covert processes like supervisory control of machines (e.g., *task analysis and task description* by Miller, 1953, 1962; the *critical incident technique* by Flanagan, 1954; or the *hierarchical task analysis* by Annett & Duncan, 1967). For a detailed overview on different task analysis methods see Schraagen (2006), Schraagen, Chipman, and Shalin (2000), and Jonassen et al. (1999).

The tasks of interest in this thesis require only covert processes. Hence, the following section presents an overview over task analysis methods for tasks that require mainly covert processes.

2.1.1 Analysis of Tasks Requiring Covert Processes

Tasks that are only or mainly composed of covert processes, like cognitive tasks, require a type of analysis that is called *cognitive task analysis* (Hoffman & Lintern, 2006; Hoffman & Woods, 2000; Schraagen et al., 2000). Cognitive task analysis refers to the process of identifying the knowledge and strategies that make up expertise in a particular domain and task (Glaser et al., 1985). The motivation to conduct a cognitive task analysis may be twofold: (1) acquiring or eliciting expert knowledge to understand knowledge structures underlying expertise and to understand expert decision making in naturalistic or field settings (Schraagen, 2006); (2) studying diverse domains of expertise for the purpose of developing better methods for instructional design and enhancing learning (Klahr, 1976). Cognitive task analysis can be applied to various cognitive tasks, like planning, maintaining situation awareness, decision making, etc. (Schraagen, 2006). Cognitive task analysis makes use of different techniques to elicit knowledge from experts, like observations, interviews, verbal reports, conceptual techniques that focus on concepts and their relations (for an overview see Cooke, 1994). In the following, I will present four examples of cognitive task analysis methods that are widely used according to Hoffman and Lintern (2006): the critical decision method, the work domain analysis, concept mapping, and the development of a task model.

The *critical decision method* (Crandall, Klein, & Hoffman, 2006; Hoffman, Crandall, & Shadbolt, 1998) requires a domain expert to recall critical decisions of her/his work by constructing a timeline. During this recall the expert is guided by the task analyst. This task analysis results in time-lined scenarios, which describe decisions, decision requirements, strategies, and informational cues. These outcomes may be used for training or system design. Hence, the focus of the critical decision method is on decision making and reasoning.

According to Hoffman and Lintern (2006) a *work domain analysis* is a functional analysis of an entire work domain. The task analyst constructs a so-called abstraction-decomposition matrix. The work domain is decomposed from organizational context, down to social collectives (teams), down to individual worker or individual component. Entities at each level constitute the means to achieve ends at the level above, the intent is to express means-ends relations between the entries of adjacent levels, with lower levels showing how higher-level functions are met, and higher levels showing why lower-level forms and functions are necessary. The worked domain analysis is used to describe, analyze, and design human-machine systems for process control and complex systems (Burns & Hajdukiewicz, 2004; Chow & Vicente, 2002; Lintern, Miller, & Baker, 2002; Naikar & Sanderson, 2001). The work domain analysis can fit knowledge and skills of the individual expert into the larger functional context of the organization and its purposes.

A third common method for cognitive task analysis is *concept mapping* (for a literature review see e.g., Crandall et al., 2006, for an application in task analysis see e.g., Hoffman & Lintern, 2006). The aim of this task analysis method is to generate models of knowledge. In that, one domain expert, and two researchers (one who acts as facilitator to create the concept map, and one to run the program) create concept maps. A set of concept maps hyperlinked together is a knowledge model of the expert. As a result, the practitioner knowledge of domain concepts such as relations, laws, and case types is represented. These knowledge models, in turn, can be used in to create knowledge bases or interfaces.

Schaafstal and Schraagen (1992) propose the *development of a task model* as a method for cognitive task analysis (see also Schaafstal, 1991, 1993; Schaafstal, Schraagen, & Van Berlo, 2000; Schraagen, 1993; Schraagen & Schaafstal, 1996). This method distinguishes three levels of analysis. On the first level, the global task structure – including the decomposition into subtasks – is described. On the second level of analysis, the procedure of each (sub)task is described. The third level of analysis consists of a description of the required knowledge and skills to perform each (sub)task. This task model is then used to predict and diagnose human cognitive behavior in terms of difficulties that are likely to occur and to provide design guidelines.

2.1.2 Summary and Conclusions

The question now is, which task analysis to choose to analyze the tasks described in the introduction (classifying fish locomotion patterns and diagnosing epileptic seizures in children). Three characteristics of these tasks need to be considered. First, both tasks do not require observable actions, but covert processes. Hence, as already stated above, task analyses that focus only or mainly on overt actions, like time study (Taylor, 1998/1911), motion study (Gilbreth, cf. Hoffman & Lintern, 2006), job analysis (Drury et al., 1987), task analysis and task description (Miller, 1962), the critical incident technique (Flanagan, 1954), or the hierarchical task analysis (Annett & Duncan, 1967) do not fit these cases. Instead, a task analysis that focuses on covert processes is required. Second, both tasks are already subtasks of broader tasks – the job of a neurological pediatrician or of a marine zoologist – and thus the according task analysis must be applicable to sub-tasks as well. For this reason, the work domain analysis (cf. Hoffman & Lintern, 2006) is not an appropriate method for this thesis. Instead, a task analysis is required that is applicable to a sub-task-level. Third, both tasks are strongly related to visually searching and interpreting perceptual input. Most task analyses, however, have a different focus: While the critical decision method (Crandall et al., 2006; Hoffman et al., 1998) focuses on decision making and reasoning, concept mapping (Crandall et al., 2006) focuses on knowledge on concepts and their relations. Thus, these two task analysis methods do not fit either.

To sum up, the task analyses mentioned in the latter paragraph are not entirely suitable for analyzing the tasks described in the introduction. Those tasks require a method allowing for the analysis of covert processes that accompany the inspection of a perceptual input on a sub-task-level. The method that most likely meets these requirements is the development of a task model suggested by Schraagen, Schaafstal and colleagues (Schaafstal, 1991, 1993; Schaafstal & Schraagen, 1992; Schaafstal et al., 2000; Schraagen, 1993; Schraagen &

Schaafstal, 1996). Although, this method addresses the global task structure as the methods described above, it decomposes and analyses the task on a sub-task-level as well and it allows for analyzing perceptual aspects of a task. Another advantage of their approach is that it addresses an unresolved issue of cognitive task analysis, that is, how to use its outcomes (Schraagen, 2006; Schraagen et al., 2000). Their approach proved to be successful to provide input for instructional design (Schraagen, 2006). Thus, the task analysis used in the current thesis will follow their approach as described in the following.

2.2 Task Analysis Method Used in this Thesis: "Development of a Task Model" According to Schaafstal and Schraagen (1992)

Schraagen, Schaafstal and colleagues have described their approach and its implementations several times (Schaafstal, 1991, 1993; Schaafstal & Schraagen, 1992; Schaafstal et al., 2000; Schraagen, 1993; Schraagen & Schaafstal, 1996), but the most detailed description of the procedure of their approach is from 1992 (Schaafstal & Schraagen, 1992). In that, the authors describe a method to analyze complex, cognitive tasks, like classification, diagnosis, and architectural design. This method requires the development of a task model. This **task model** is composed of three levels: (1) a task structure of top-level goals, (2) procedures to achieve the goals described in the task structure, and (3) a description of the domain knowledge and skills necessary to perform the task procedures. Afterwards, **conclusions** can be drawn based on this task model, for instance, in terms of difficulties that are likely to occur in performing the given task and in terms of instructional goals that need to be addressed for successful training. Two paths should be used to develop a task model: First, a **normative task analysis** is performed, which reveals how the task *should* be performed. Second, an **empirical task analysis** is conducted to investigate how the task actually *is* performed.

2.2.1 Development of a Task Model

2.2.1.1 First Level: Task Structure

The first level refers to complex tasks that are composed of many sub-tasks. Such tasks are called "whole tasks"[1] (cf. Van Merriënboer & Kirschner, 2007) or "generic tasks" (Chandrasekaran, 1983). The tasks under consideration in this thesis are subtasks of whole tasks on a higher and more generic level. Classifying fish locomotion patterns, for instance, is only one subtask of the higher-level task to classify fish species. It is not sufficient to classify a fish's locomotion pattern to know about its membership to a species. Instead, other characteristics have to be taken into account as well, like the behavior of the fish, its environment, its morphology, or its genetics (cf. Lieske & Myers, 1994). Similarly, diagnosing an neurological disease based on motion patterns is only one subtask in reaching a definitive diagnosis. An entire diagnosis of a neurological disease requires taking anamnesis, examinations of the blood, or using imaging techniques (cf. World Health Organisation, 2006). As the task of interest in this thesis are only subtasks, for the purpose of this thesis, I will skip the first level. The following sections describe the actual task analysis procedure applied in this thesis.

2.2.1.2 Second Level: Procedural Task Analysis

In this stage of task analysis (which is the **first** stage in the case of this thesis), Schaafstal and Schraagen (1992) analyze the procedure of performing a task. The authors define steps that need to be taken to achieve a goal, that is, finalizing the task. Other task analysis methods also use similar approaches to describe a task's procedure by either decomposing the tasks into subtasks (Klein, 1993; Taylor, 1998/1911), subgoals (Annett & Duncan, 1967; Catrambone, 1998), or into series of discrete actions or steps (Jonassen et al., 1999, Chapter 5, Chapter 9; Kieras, 2004). Schaafstal and Schraagen (1992) refer to the order and composition of these

[1] Whole tasks are tasks that are composed of all or almost all of the constituent skills, associated knowledge, and attitudes for real-life task performance (Van Merriënboer & Kirschner, 2007, p. 16).

steps as ”local strategies”[2]. The authors stress that in some cases different strategies – that is, a different order of the steps or omitting certain steps – may be used to perform a certain task (cf. also Annett & Duncan, 1967). Often expert performers display specific and powerful types of strategies that include omitting steps or using a different order of the steps, namely heuristics or shortcuts (e.g., Rasmussen, 1986).

2.2.1.3 Third Level: Analysis of the Required Knowledge and Skills

In the next stage of the task analysis (which is the **second** stage in the case of this thesis), Schaafstal and Schraagen (1992) analyze the knowledge and skills required to achieve each step of the task. It is important to note that the authors mainly address covert actions, like information processing steps, at this stage of task analysis. As the tasks of interest in this thesis require the inspection and interpretation of visual material, perceptual processes are likely to play a crucial role, too. Thus, I will include the analysis of potentially important perceptual skills[3] in this step (see Section 2.3.2). The results of the task analysis – the task model – can be conveniently displayed by means of tables. In such tables one axis represents the knowledge and skills, while the other axis represents the task steps. The cells of these tables point to the connection of task steps and the knowledge and skills they require.

2.2.2 Drawing Conclusions Based on the Task Model

2.2.2.1 Analyzing Potential Difficulties of the Task

After having analyzed the task procedure and the required knowledge and skills (i.e., after developing a task model), Schaafstal and Schraagen (1992) use these information to predict human cognitive behavior in terms of errors that are likely to be observed. According to the authors the main source of error results from an overload of working memory. They state that this would be especially true for novices. Their errors may occur on both levels: they may

[2] In contrast to ”global strategies” that can be defined on the first level (task structure).

[3] Adding different types of important knowledge or skills into the task analysis is also suggested by Schaafstal and Schraagen (1992): ”...and one may wonder whether it would not be beneficial to incorporate many different types of knowledge that play a role in the domain.” p. 27

lack the relevant local strategies (i.e., knowing which steps to perform and in which order) or the knowledge and skills to execute each task step. Other task analysis methods also place a high relevance on difficulties or (typical) failures that may occur in the task (Annett & Duncan, 1967; Flanagan, 1954; Jonassen et al., 1999; Militello & Hutton, 1998; Miller, 1962). As stated above, the tasks used in this thesis potentially include a high relevance of perceptual processes and are thus, prone to possible errors or difficulties related to the processing of perceptual input. Thus, I will include the analysis of potential difficulties from a perceptual perspective in this step.

2.2.2.2 Deriving Instructional Goals

Other conclusions that Schaafstal and Schraagen (1992) draw from a task model are related to the development of trainings (see also Clark et al., 2008; Jonassen et al., 1999). They recommend to draw conclusions for training from each level of the task analysis. Based on the task structure level – which is not present in this thesis – subgoal structures can be successfully trained (e.g., Soloway & Johnson, 1984). Based on an analysis of the procedural level, a training of the stepwise task procedure can be implemented (e.g., Pirolli & Anderson, 1985). Schaafstal and Schraagen (1992) explicitly state that training knowledge and skills can be successful only when a prior task analysis has shown, which knowledge and skills are crucial for the task of interest, otherwise instruction remains ineffective. When deriving instructional goals, I will also refer to the analysis of the difficulties that are likely to occur in a task.

2.2.3 Two Paths to a Task Model: Normative and Empirical

The last sections provided information on how to perform a task analysis and on what conclusions to draw from it according to Schaafstal and Schraagen (1992). The current section will describe how information on the task *content* can be collected. Schaafstal and Schraagen (1992) propose to conduct the analysis of a task in two manners: (1) by means of a normative analysis on how the task *should* be performed and (2) by means of an empirical analysis on how the task *is* actually performed by experts and novices. All task model levels (i.e., procedure, knowledge and skills) and all conclusions drawn from the task model (i.e., difficulties of the task, instructional goals) need to be addressed by both perspectives.

Following Schaafstal and Schraagen (1992) the **normative task analysis** should be conducted by analyzing handbooks, work procedures, educational materials used and so on, which may be quite useful in terms of identifying local strategies (for example: how to carry out a certain test) and the relevant domain knowledge. The authors warn, however, that this material may either not be entirely suited for the task at hand or not be written for educational purposes and thus, may include types of knowledge that are not used in real-life situations.

Additionally, Schaafstal and Schraagen (1992) recommend to compare the normative task model with the results of an **empirical task analysis**. Schaafstal and Schraagen (1992) advice to review the literature on existing studies that can be applied to the task of interest. If this review does not yield sufficient information to develop an empirical model for the given task, a study needs to be conducted that addresses this specific task. This study is intended to reveal whether people follow the strategies defined by the normative task analysis or whether they differ in some respect. The authors recommend an analysis of the cognitive processes conducted during task performance (for a detailed description of suitable methods for these analyses see Section 2.3.1). In particular, they recommend taking verbal protocols of experienced and less experienced performers (i.e., experts and novices, Schaafstal & Schraagen, 2000). The authors claim that this type of analysis is a feasible way of assessing strategies and the domain knowledge applied. The authors performed this task analysis method successfully by presenting experts and novices with realistic problems and asking them to think aloud while performing the task (Schaafstal, 1993; Schraagen & Schaafstal, 1996).

To perform a task analysis based on expert performance, it is important to find an appropriate definition of expertise. For the purpose of investigating expertise differences it is important to focus on people, who show consistently superior performance on a specified set of representative tasks for a domain (Ericsson & Lehmann, 1996; Ericsson & Smith, 1991). The major methodological challenge with regard to this approach is to identify standardized tasks that can be *reproduced* under laboratory conditions and that *capture* the essential aspects of a particular type of expert performance (Ericsson & Lehmann, 1996). For some types of expert performance, like typewriting or exceptional memory, the conditions under which these tasks are performed are already standardized so that they are easy to reproduce in the laboratory (Ericsson & Lehmann, 1996). For others, however, it is difficult to reproduce the exact same conditions under which a task is performed or even to design a collection of standardized tasks (e.g., in sports, Ericsson & Lehmann, 1996). For instance, for the tasks that are in the focus of this thesis the objects of interest are difficult to observe under controlled conditions either due to their rare occurrence (e.g., certain neurological diseases in infants) or due to their occurrence in places that are difficult to access (e.g., reef fish). Moreover, both tasks involve movements of living objects; therefore, it is unlikely that these objects (fish, infants) will move in the same way reliably with (ir)relevant features being observable in the same way across multiple observations. Hence, for tasks that require exposure to living beings and their behavior, which is rare to observe and difficult to standardize a good alternative is to present a set of video recordings of the relevant dynamic stimuli (i.e., swimming fish or movements of infants) to experts under laboratory conditions. This procedure mimics as closely as possible a real-life task situation while at the same time allowing for presenting tasks in a controllable and standardized way. This method has already been successfully applied in studies in sports (cf. Moreno et al., 2002; Raab & Johnson, 2007) and is hence also chosen for this thesis. All videos used were carefully chosen with the help of domain experts.

As already mentioned, the tasks used in this thesis are very likely to have a strong perceptual component. The task model approach does not comprise a way to investigate this component directly, because as a form of cognitive task analysis it focuses on knowledge and skills related to *cognitive* processes. Hence, this approach needs to be extended for the purpose of this thesis. Schraagen (2006) mentions an attempt to address the perceptual component in analyzing the task of shore-based pilotage (i.e., guiding a ship from the sea to the harbor entrance). Besides methods that focus on cognitive processes like thinking aloud, time on task and retrospective interviews, he also captured every five minutes photographs of the view ahead "in order to obtain a running record of the pilot's perceptual input" (Schraagen, 2006, p. 197). Unfortunately, those recordings were never analyzed, because they resulted in "too many data" (Schraagen, 2006, p. 198). It can be assumed that the author did not find an automated and standardized way to analyze this type of data. Thus, a better way of capturing perceptual processes is needed, for which clear guidelines for analysis exist. In line with typical observational techniques used for task analysis, the perceptual processes might also be observed directly by means of eye tracking (for a detailed description of this method see Section 2.3.2). Eye tracking has already been successfully used for task analytical approaches in, for instance, pilot interactions with flight management systems (Diez et al., 2001). Since researchers state that different contexts demand different task analysis methods (Jonassen et al., 1999, p. 5) and since technological developments usually resulted in changes in task-analysis methods (Schraagen et al., 2000), adding the analysis of perceptual processes (visual attention) of experienced and less experienced performers by means of the eye tracking might be an important improvement for the empirical stage of task analysis according to Schaafstal and Schraagen (1992).

2.3 Capturing Covert Processes

As stated above, a task analysis should reveal a collection of information necessary to develop a model of the task (i.e., which steps and in which order were taken to accomplish the task and which knowledge and skills were applied to achieve each step). Schraagen and colleagues postulated that an important part of a task analysis is to study domain experts while performing the task of interest to gain this information (i.e., empirical task analysis; Schaafstal & Schraagen, 2000, 1992; Schaafstal et al., 2000; Schraagen, 2006, 1993; Schraagen et al., 2000; Schraagen & Schaafstal, 1996). For tasks that require covert processes (cognitive or perceptual) a methodological challenge is how to gain access to these information needed for the task model. Researchers have designed a variety of techniques to get insight into these covert processes (Cooke, 1994; Ericsson & Lehmann, 1996; Hoffman, 1987, 1995; for methods specifically used for task analysis see Jonassen et al., 1999, Chapters 25-31). Usually, the cognitive processes mediating task performance are investigated by means of verbal reports (Ericsson, 2006; Ericsson & Simon, 1993).

However, the tasks of interest in this thesis have a strong perceptual component beyond the cognitive processes involved in that pursuit. Hence, for a thorough analysis of these tasks, both the cognitive *and* perceptual processes accompanying task performance are important. Perceptual processes, are usually captured by means of eye tracking (Duchowski, 2003; Holmqvist et al., 2011). To evaluate the existing literature on verbal reports and eye tracking it is important to understand their potentials and drawbacks. Hence, in the remainder of Section 2.3, both methods will be described in greater detail and a combination of both approaches that allows for methodological triangulation will be recommended for the tasks investigated in this thesis.

2.3.1 Investigating Cognitive Processes by Using Verbal Reports

This section first describes what verbal reports are. Next, it presents how verbal reports can be used as a measure of cognitive processes. In particular, it will be discussed how verbal reports have been used for task analysis purposes. Finally, shortcomings of the verbal reporting technique will be addressed, which explain why an additional process tracing technique is required to obtain a thorough picture of the perceptual tasks of interest in this thesis: eye tracking. This technique will be described in the subsequent section.

2.3.1.1 The Origin and Idea of Verbal Reports

Initially, the easiest and most common way to gather insight into cognitive processes accompanying expert performance was to interview experts (Ericsson, 2006). Expert interviews, however, have been shown to suffer from severe problems. For instance, it is questionable whether experts are able to describe their thoughts, behaviors, and strategies so that they are understandable to less skilled people[4] (Ericsson, 2006). Moreover, it has been found a long

[4] A less skilled person would be the task analyst in this case, who tries to get an insight into the task.

time ago already that multiple experts provide inconsistent reports (Binet, 1893/1966). Even worse, discrepancies have been found between reported and observed behavior (Watson, 1913). For this reason Watson (1920) and Duncker (1945) introduced a new method of thought analysis: *verbal reporting*. This type of verbalization has been shown to retain the underlying structure of thought, that is, the cognitive processes unchanged (Chi, 2006a; Ericsson & Simon, 1980, 1984, 1993).

The central assumption underlying the use of verbal reports is ”that it is possible to instruct subjects to verbalize their thoughts in a manner that does not alter the sequence and content of thoughts mediating the completion of a task and therefore should reflect immediately available information during thinking.” (Ericsson, 2006, p. 227). In verbal reporting participants are expected to verbalize their ”inner speech”. These verbalizations are supposed to deliver information on the knowledge that is currently activated and on changes with regard to this knowledge. According to Ericsson and Simon (1993) the following two assumptions are important for the analysis of verbal protocols: (1) The verbalizations can be described as states that correspond to the contents of working memory (i.e., to the information that is in the focus of attention); (2) the information verbalized is a verbal encoding of the information in working memory. That is, only content can be found in the data that is ”on the participant’s mind”, respectively in the participant’s focus of attention.

Verbal reporting can be applied to study the structure of expert performance and to access cognitive processes that mediate expert performance (Ericsson, 2006; Ericsson & Simon, 1984, 1993). This technique has been successfully used in a variety of domains to study expertise differences, like designing surveys (Sudman, Bradbrun, & Schwarz, 1996), learning a second language (Green, 1998), text comprehension (Ericsson, 1988; Pressley & Afflerbach, 1995), or developing computer software (Henderson, Smith, Podd, & Varela-Alvarez, 1995; Hughes & Parkes, 2003). The technique of verbal reporting has also been used directly for the purpose of task analysis (for an overview see Schraagen, 2006). Schaafstal (1993), for instance, compared verbal protocols of novice and expert performers in troubleshooting while operating a paper mill. Novices showed an unsystematic and very simple approach to the task as compared to experts, which caused novices to not being able to detect and solve errors correctly. Schaafstal et al. (2000) reported similar findings based on the same technique in the domain of weapon engineering. Detailed analyses of verbal protocols, followed-up by an observational study on the training (Schraagen & Schaafstal, 1996) lead to a revision of the training. The revised training was more efficient in terms of shorter training duration and better learning results compared to the common training beforehand. Thus, the verbal reporting technique used in expertise studies for task analysis purposes have already been demonstrated to be of help for instructional-design purposes.

Since verbal reporting during task performance is suspected to slow down task performance (e.g., Karpf, 1973) verbal reports may also be recorded *after* a person has performed the task to reduce potential interferences between verbal reporting and task performance. To avoid forgetting and fabulating of passed thoughts, verbal reports can be cued by recordings of own overt actions (like mouse clicks on a webpage) or covert processes (eye movements) during task performance. This method is called **cued retrospective reporting** (Van Gog, Paas, Van Merriënboer, & Witte, 2005). In this case participants do not provide verbal reports during task performance, but *afterwards* while seeing a recording of their own eye movements from task performance (for a detailed description of this method see Section 3.2.2.2). First studies emphasize the value of using a recording of own eye movements to stimulate verbal reporting (i.e., cued retrospective reporting) for task analysis purposes to elicit knowledge from experts (Seagull & Yan, 2001).

A crucial part in the use of verbal reports is its coding (Chi, 2006a). The data should be coded in the context of the task. Hence, a normative task analysis needs to be done beforehand to specify functional problem states according to which single utterances can be categorized (see Section 3.1). This procedure is necessary to figure out which states and in which sequence each participant went through to get an insight into the process of performing the task.

2.3.1.2 Shortcomings of Verbal Reports

Besides the opportunities to gain insight into cognitive processes while performing a task and the fact that verbal reports are generally accepted to provide valid information concerning a person's cognitive processes (H. A. Simon & Kaplan, 1989), we have to accept several shortcomings in using this method for task analysis. Much of experts' knowledge is supposed to be automated and thus, not explicit (Chi, 2006a; Speelman, 1998). Hence, experts often have difficulties to verbalize it (Chi, 2006b). And even if experts do verbalize their knowledge, they use a lot of technical terms in doing so (Reimann & Rapp, 2008). Furthermore, their verbalizations often do not contain descriptions in enough detail. This is in particular true for perceptual processes (Ericsson & Simon, 1993). Thus, even though this method can provide information on perceptual processes, for instance, when experts note more relevant features of pictures than novices do (Wineburg, 1991) it does so in a quite indirect way and at a coarse level of granularity. Verbalizing perceptual processes requires what Ericsson and Simon (1984, 1993) call ”level 2” verbalizations. Level 2 verbalizations refer to verbalizing processes that are not inherently given in a verbal form (i.e., encoding of information that is not isomorphic with language; Ericsson & Simon, 1984, 1993). Before being able to verbalize those processes, the participant has to ”translate” them into a verbal form. This translation, however, is difficult and thus, often not comprehensively executed (Ericsson & Simon, 1980). Moreover, it is more likely to result in distortions compared to ”level 1” verbalizations, where processes are verbalized that are already available in working memory in a verbal form (e.g., in mental arithmetic). Thus, verbal protocols of experts in the empirical task analysis of the tasks used in this thesis are very likely to be incomplete or circumscribed with technical terms. This may be particularly true for the perceptual skills involved that will be described in Section 3.1.2. These skills, however, are assumed to be of central importance for performing these tasks. Thus, to address the shortcomings of the verbal reporting technique, an additional method to capture perceptual processes is needed, namely eye tracking. This method will be discussed in the following section.

2.3.2 Investigating Perceptual Processes by the Use of Eye Tracking

Verbal reporting is a widely used method to elicit knowledge during task performance for task analysis purposes (Schaafstal, 1993; Schaafstal et al., 2000; Schraagen, 2006; Schraagen & Schaafstal, 1996; Seagull & Yan, 2001). Eye tracking, on the other hand, has so far hardly been used for task-analytical purposes (for an exception, see e.g., Diez et al., 2001) and is thus, not an established method for task analysis as verbal reports are. As perceptual processes are very likely to play a crucial role in performing tasks that heavily rely on perceptual input, this method will be used for task analysis purposes in this thesis. Since eye tracking is still a sparsely used technique in educational psychology, at least as compared to verbal reporting, this section starts by introducing the methodology in greater detail. Next, it will be described how eye tracking was used to investigate perceptual and even cognitive processes.

2.3.2.1 Eye Tracking as a Methodology to Analyze Distribution of Visual Attention

Eye tracking is a method of tracking the movement of the eyeball(s) in order to infer where a person looks at. In earlier times, eye tracking was difficult and effortful for both, the researcher and the participant (for an overview of eye tracking history see Wade & Tatler, 2005). Modern eye tracking systems allow for an unintrusive tracking for the participant; moreover, constant development of hard- and software enables researchers to achieve robust and detailed data for various types of stimuli. The eye tracking method used in this thesis is the pupil-and-corneal reflection method (for the very first implementation, see Dodge & Cline, 1901). When measuring eye movements with this method, an infrared light source is directed towards the eye. This causes a reflection on the cornea (corneal reflection). A video camera captures images of the eye with the corneal reflection. An image processing software recognizes the brightest (corneal reflection) and the darkest spot (pupil) on this video (cf. Figure 2.1). When the eye moves, the distance between both spots changes. In combination with parameters of the environment (i.e., coordinates of the screen and distance from the eyes to the stimulus monitor), it can be inferred at which position of the environment the eye was directed to at which point in time.

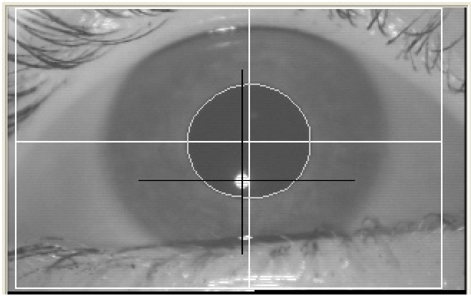


Figure 2.1: Image of the eye as the eye tracking system ”sees” it (www.smivision.com).

The resulting eye tracking data are usually aggregated into different types of eye movement parameters, so-called fixations and saccades (for more details see Duchowski, 2003; Holmqvist et al., 2011). Fixations stabilize the focus of the eye on a certain position of a stimulus. Fixations include all minimal eye movements within a small spatial range (e.g., according to Duchowski, 2003: less then 5 degrees of visual angle) and within a limited amount of time (e.g., according to Duchowski, 2003: 150 - 600 ms). Saccades, on the other hand, are rapid eye movements that move the eye’s focus from one position to another position of a stimulus and are movements over a far larger spatial range then fixations (cf. Duchowski, 2003). When a moving object is presented within a stimulus and the eye follows it (as it is the case for the tasks used in this thesis), another type of eye movements occurs, namely *smooth pursuit* (Dodge, 1903). In that, the eyes match the speed of the moving object, which results in far slower motion then saccades. It is generally considered that when measuring a fixation or smooth pursuit, the attention to that position is captured, that is, that information from the stimulus is perceived by the observer (cf. eye mind-assumption; Just & Carpenter, 1976). In contrast, the eyes are ”blind” during saccadic movements, that is, information is not taken in from the stimulus by the observer (Holmqvist et al., 2011). Hence, fixations and smooth pursuit movements are the eye movement parameters of interest when investigating attention allocation. However, while fixations and saccadic movements of the eye can be easily detected by contemporary eye tracking software, there is no satisfactory way to detect smooth pursuit, yet (Holmqvist et al., 2011). When applying this contemporary eye tracking software to eye tracking data that is likely to contain smooth pursuit, this data would be biased. The smooth pursuit eye movements would probably be incorrectly classified as saccades, because saccades also include movements over larger spatial ranges then fixations. Hence, the researcher would incorrectly assume that no information intake took place, although during smooth pursuit eye movements information is extracted from the stimulus, . Hence, when dealing with eye tracking data that are very likely to include smooth pursuit movements, it is recommended to use temporally raw data instead (i.e., data that has not been aggregated by eye tracking software into eye movement parameters).

2.3.2.2 Eye Tracking as a Measure of Perceptual and Cognitive Processes

The most direct way to investigate perceptual (here: visual) processes is via eye tracking. Eye tracking data allow for drawing conclusions about *which part* of the stimulus the participant attended to, for *how long*, and in *what order*. Thus, these kind of data give insight into visual search and the allocation of (overt) visual attention.

Although eye tracking has hardly been used for task analysis purposes, it has been used to improve instructional design in other ways. For instance, in educational psychology, eye tracking is used to study how different designs of multimedia (i.e., combinations of text and pictures, Mayer, 2005b) affect perceptual processes during learning (Schwonke, Berthold, & Renkl, 2009). Moreover, eye tracking is suited to study how (developing) expertise affects processing of visual stimuli (like static pictures, animations, or videos, cf. Canham & Hegarty, 2010). Both approaches (indirectly) contribute to better design of instructional materials or procedures. Although eye tracking recently has become more popular in educational psychology, direct investigations of cognitive and perceptual processes for instructional design purposes are still rare (for changes in that see two current special issues: Van Gog & Scheiter, 2010; Scheiter & Van Gog, 2009). The only slowly emerging popularity of eye tracking in educational psychology might go back – besides technical issues – to an important shortcoming of this method, which will be described in the following paragraph. In the next section, we will see how to compensate for this shortcoming.

Eye tracking data tells us where on the stimulus and possibly for how long a person allocated her or his visual attention. But how can we know, *which* cognitive process is involved? For instance, if a person in a fish locomotion classification task looks for a long time at a fin, does this long gaze indicate s/he is interpreting the fin's motion, or does it indicate uncertainty whether it is the relevant fin that produces propulsion? In other words, eye movement data alone is often ambiguous. One approach to compensate for this shortcoming is to add another data source to resolve this ambiguity. This combination of approaches is called *methodological triangulation*.

2.3.3 The Benefits of Using Multiple Methods: Methodological Triangulation

Methodological triangulation refers to the use of more than one methodological approach in investigating a research question in order to enhance confidence in the ensuing findings (Denzin, 1970). If research was founded on the use of one single research method it might suffer from limitations associated with that method or from the specific application of it. Thus, methodological triangulation offers the prospect of enhanced confidence, credibility, and persuasiveness of a research account through verifying the validity of the findings by cross-checking them with another method (Bryman, 1984). ”Once a proposition has been confirmed by two or more independent measurement processes, the uncertainty of its interpretation is greatly reduced. The most persuasive evidence comes through a triangulation of measurement processes” (E. J. Webb, Campbell, Schwartz, & Sechrest, 1966, p. 3). Since recording verbal protocols of participants is a method that allows researchers to gain insight into participants’ cognitive processes while inspecting a visual stimulus or performing a task, they are a good supplementation to eye tracking data. Thus, for analyzing tasks that heavily rely on perceptual input, like those used in the current thesis, both methods – eye tracking and verbal reports – should be used in conjunction.

In sum, the task model approach recommends both, a normative task analysis and an empirical analysis of expert-novice differences in terms of differences in strategies applied and procedures used, knowledge and skills used, and misconceptions and faulty procedures that novices may have (Schaafstal & Schraagen, 1992). This information, in turn, should be used as input for instructional design (Schraagen, 2006). I will follow this approach in the next two chapters for the first experimental domain in this thesis, namely the classification of fish locomotion patterns.

Chapter 3

Analyzing the Task of Classifying Fish Locomotion Patterns

The following chapter will provide an analysis of the task of classifying fish locomotion patterns as described in Section 2.2. The first part of the chapter (Section 3.1) provides a normative task analysis and the second part of the chapter (Section 3.2) provides an empirical task analysis.

3.1 Normative Task Analysis Based on Theoretical Considerations

According to Schaafstal and Schraagen (1992), a normative task analysis is composed of three analyses, which will be addressed in this chapter. First, the procedure of the task has to be described as discrete steps that need to be performed to solve the task. Second, the knowledge and skills that are required to execute each step have to be identified. Both analyses result in a task model that is depicted in a tabular format. Third, based on this task model difficulties that may occur in this task are predicted. In the end of this chapter, this task analysis will be used to derive instructional goals when teaching the task.

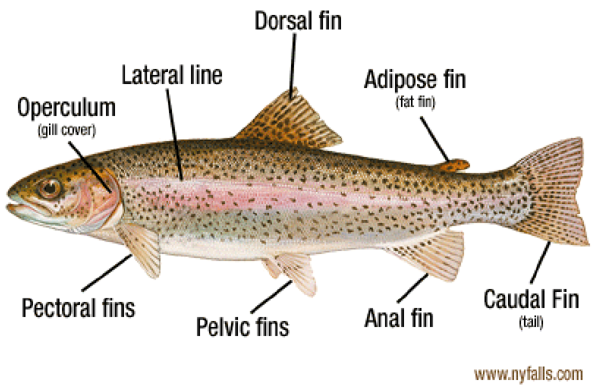


Figure 3.1: The naming of a fish's body parts. From www.nyfalls.com

3.1.1 Analyzing the Task Procedure

As domain of interest investigated in Study 1 and Study 2 of this thesis the different locomotion patterns of reef fish are chosen. Locomotion pattern used by reef fish for swimming forward can be classified in different groups according to which body parts contribute to the propulsion required to swim forward and in what manner these body parts move (Lindsey, 1978; for an overview of fish body parts see Figure 3.1). For the purpose of this thesis, four different locomotion patterns were chosen with the help of a domain expert and teacher in the field of marine zoology (S. Gemballa, professor of marine zoology, personal communication, May, 2007), namely balistiform, (sub)carangiform, labriform, and tetraodontiform. A fish swimming balistiform uses its dorsal and anal fin in an undulating (i.e., wave-like) way to generate thrust, a fish swimming (sub)carangiform uses its caudal fin in an undulating way, a fish swimming labriform uses its pectoral fins in an oscillating (i.e., paddle-like) way, and a fish swimming tetraodontiform uses its dorsal and anal fin in an oscillating way (cf. Figure 3.2 for illustration). For a detailed description of fish locomotion patterns see Blake (2004), Lindsey (1978), or P. W. Webb (1984).

To classify a fish locomotion pattern, the following steps should be applied (Lindsey, 1978): First, it has to be decided which parts of the body is used to produce propulsion. This can either be the body itself or the fins. Second, it has to be decided how these parts moves. This can either be in an undulating (i.e., wavelike) or an oscillating (i.e., paddlelike) way. Third, each combination can be assigned to a certain class: balistiform, (sub)carangiform, labriform, or tetraodontiform. This latter decision is purely based on conceptual knowledge about how different fish locomotion patterns are defined. For an overview of the steps required to perform this task and how they interrelate see Table 3.1.

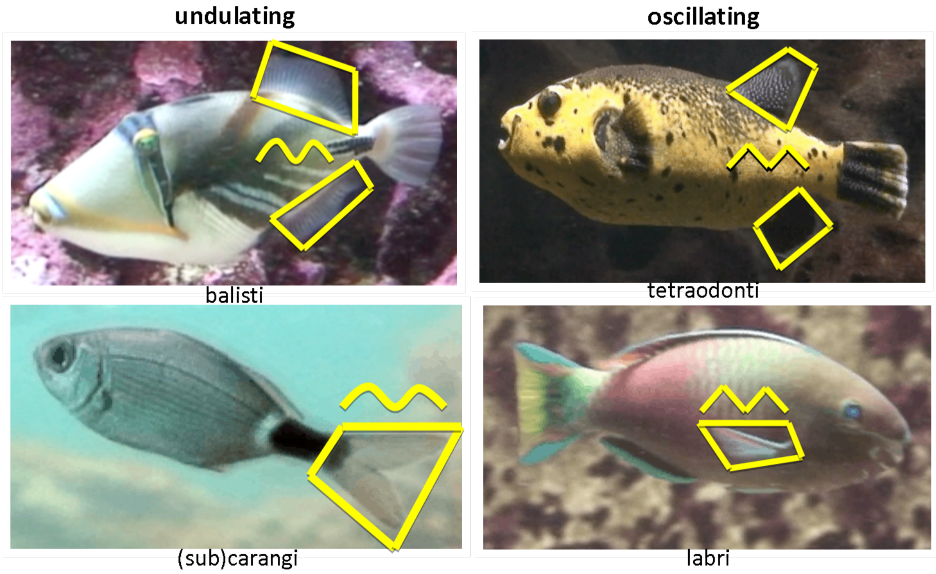


Figure 3.2: The four locomotion patterns investigated in Study 1 and Study 2 of the current thesis: (sub)carangiform, balistiform, labriform, tetraodontiform.

Table 3.1:
Procedure of the Task to Classify Fish Locomotion Patterns

Step 1: Specifying body parts involved in generating thrust				
Anal fin	✓			✓
Caudal fin		✓		
Dorsal fin	✓			✓
Pectoral fin			✓	
Step 2: Specifying the motion pattern of these body parts				
Oscillating			✓	✓
Undulating	✓	✓		
Step 3: Classification of the locomotion pattern				
Balistiform	✓			
(Sub)carangiform		✓		
Labriform			✓	
Tetraodontiform				✓

3.1.2 Identifying the Knowledge and Skills Required to Perform the Task

The task to classify fish locomotion patterns is a **knowledge-rich task**, because not all information needed to perform the task is given in the task itself (Van Lehn, 1989). Hence, prior-knowledge or training is required to solve this task. The question is what type of knowledge or skill is required for this task and hence, should be in focus of training. Since this task requires the inspection and examination of a perceptual input – namely videos of swimming fish in Study 1 and 2 – without which the task cannot be performed, this task can be described as being a **perceptual task** (Chase & Chi, 1981; Chi, 2006a). Hence, the skill to inspect and examine swimming fish is supposedly important to teach. In the following, we will see how to describe the knowledge and skills required to perform this perceptual and knoweldge-rich task.

For **perceptual tasks**, it may be appropriate to distinguish between the knowledge that refers to concepts or facts in the domain and the skills required to deal with the perceptual input. This distinction was proposed in an adjacent perceptual domain: olfaction. Parr, Heatherbell, and White (2002) refer to recognizing an odor – which requires to actually smell it – as perceptual skills, whereas they refer to naming an odor – which requires elaborating about the smell and categorizing it – as semantic knowledge. Based on their study of expertise differences in wine judging they state that the perceptual skills are the crucial factor of expertise (and not semantic knowledge). Hence, distinguishing the skills and knowledge required for the (visual) task of classifying fish locomotion patterns into perceptual skills and conceptual knowledge appears to be useful. This distinction can be applied to the current task by distinguishing conceptual knowledge about the definition of different locomotion patterns and perceptual skills to inspect and examine a fish swimming according to a certain locomotion pattern.

These perceptual skills, however, do not refer to perceptual processes only, but also to cognitive processes that require a concrete perceptual input. Chi (2006a) elaborates on methods to uncover the various representational differences in experts and novices (e.g., recall, perceiving, categorizing, verbal reporting). For differences in terms of *perceiving* she proposes to determine ”...what experts and novices see (literal stimulus features) and perceive (meanings of the features or patterns of features).” (p. 172). Her description of representational differences on a perception level may be applied to the current task: The term ”seeing” can be related to skills required to perform the first task step (specifying body parts involved in generating thrust) , while the term ”perceiving” can be related to skills required to perform the second task step (specifying how the relevant body parts move). A similar distinction has been made by Manning, Gale, and Krupinski (2005) in their review of research on ”perception” in medical imaging (e.g., X-rays or mammograms). They state that research in this area is mostly concerned with errors in diagnosis based on medical images ”due to visual search, detection of radiological features, and their interpretation” (p. 684). Again, a connection can be drawn to the perceptual skills required to perform the current task: Step 1 requires the visual search and detection of relevant areas and Step 2 their interpretation. Thus, a respective differentiation of perceptual skills will be included in the description for the current task as well.

Table 3.2 provides an overview of the task procedure and the required knowledge and skills in the form of a task model.

Table 3.2:
Task Model for the Task to Classify Fish Locomotion Patterns

Task steps	**Required knowledge and skills**
Perceptual input mandatory (steps based on **perceptual skills**):	
1. Specifying *body parts* involved in generating thrust	Visual search and identification of relevant body parts
2. Specifying the *motion pattern* of these body parts	Visual inspection of relevant body parts and interpretation of their motion
Perceptual input NOT mandatory (steps based on **conceptual knowledge** only):	
3. *Classification* of the locomotion pattern	Assignment of observations to the correct class

As this task is a **knowledge-rich task**, it requires knowledge structures in long-term memory that can be described as schemas (cf. Cooper & Sweller, 1987). In the task of classifying fish locomotion the top-level schema might comprise the following information: a fish locomotion pattern is defined by the body parts used to swim and the type of their movement. Lower-level schemas might represent the characteristic features of the different fish locomotion patterns (but also characteristics of different body parts and of different motion patterns). When a schema is activated, it needs to be instantiated with information specific to the current example. Schema instantiation happens during the first two steps defined in the task procedure above. The first step requires to efficiently visually search for the body parts that produce propulsion to swim forward and to identify them as such (i.e., denoting them as being relevant) to fill the first slot of the schema. The second step requires to visually inspect and interpret the motion of the relevant body parts to fill the second slot of the schema. This means that when inspecting a swimming fish the relevant body parts need to be focused on, while all other irrelevant elements have to be ignored. When the slots of the top-level schema have been instantiated, a schema-specific solution procedure can be executed as the third step of task accomplishment (Van Lehn, 1989). In the current task, this might result in the activation of a lower level schema specific for the fish locomotion pattern at hand that contains further information on this fish locomotion pattern (for instance, its technical term). To execute this solution procedure, that is, the last step of this task, the according conceptual knowledge is required. It thus results from the task analysis that conceptual knowledge is necessary but not sufficient for successful task performance. Before the third task step can be performed, the first two steps need to be executed. These first two steps essentially require – in contrast to the third step – a perceptual input, for instance, in form of a video of a swimming fish.

In sum, the knowledge-rich and perceptual task to classify fish locomotion patterns is composed of conceptual knowledge and perceptual skills. Executing these ”perceptual” skills – i.e., skills that cannot be executed without perceptual input – requires both cognitive and perceptual processes. These perceptual skills are composed of two parts (cf. Chi, 2006a; Manning et al., 2005):

1. Detecting **which** body parts are used to swim: *Visually searching* the relevant body parts (i.e., quickly looking at them) and *identifying* the relevant body parts (i.e., denoting them as being relevant).

2. Detecting **how** these body pats are used: *Visually inspecting* those relevant body parts (i.e., looking at them for a sufficiently long[1] time) and correctly *interpreting* the motion of those relevant body parts (i.e., describing their motion correctly).

Visually searching (part of task step 1) and inspecting (part of task step 2) are perceptual processes that can be tested by capturing the visual attention of the task performer while inspecting a video of a swimming fish (e.g., by means of eye tracking). Identifying these body parts as relevant (part of task step 1) and interpreting its type of motion (part of task step 2) are cognitive processes that can be tested, for instance, by means of questionnaires administered after having inspected a video. Since to perform these skills a perceptual input is mandatory, they will be referred to as **perceptual skills** for the purpose of this thesis.

3.1.3 Predicting Difficulties

The normative analysis resulted in a task model for the task to classify fish locomotion patterns. Based on this task model this section derives possible difficulties that may occur when performing this task. Potential difficulties may be caused by the procedure of the task itself as well as by characteristics of the perceptual input used in Study 1 and 2 of this thesis, which consists of video recordings of swimming fish.

3.1.3.1 Difficulties Caused by the Task Procedure

As described above *three* steps need to be executed to perform the task of classifying fish locomotion. In the first step – detecting which body part(s) are crucial for thrust propulsion – unexperienced people (i.e., novices) may either detect the wrong body parts or may not

[1]Instead of being distracted by salient, but irrelevant areas, that is, in relation to the total viewing time of the perceptual input.

detect all body parts. In the second step, the motion may be interpreted incorrectly (although the second step only refers to a dichotomous decision; i.e., either undulating or oscilating motion). In the final step, the results of the first two steps may be assigned to a wrong class.

Accordingly, novices may fail at all three steps of this task. The first two steps, however, are a necessary prerequisite for the third step. Without pursuing the first two steps correctly one must fail to complete the last one. Hence, the first two steps are of a particular importance in this task. As already mentioned, these two steps cannot be executed without a perceptual input. Thus, in the next section difficulties that might be caused by characteristics of the perceptual input are described in greater detail.

3.1.3.2 Difficulties Caused by the Perceptual Input

Under real-life task performance conditions, the perceptual input in the fish-locomotion-classification task would probably be a reef fish swimming forward in its natural habitat. To standardize the experimental conditions, video recordings of reef fish were used. However, it is important to note that the core characteristics of both types of perceptual input (i.e., living fish vs. video recordings) are similar and so are the difficulties that they may cause for the observer. This is the case, because the video camera that is controlled by a diver in the fish's natural habitat substitutes for the observer him- or herself. Four exemplary screenshots of the videos used in Study 1 and Study 2 are presented in Figure 3.2. In the following, to analyze the characteristics of these stimuli, research on learning from dynamic visualizations is consulted as this research area has focused on challenges that may occur when inspecting highly information-dense and dynamic input. Accordingly, findings from this research area can be well applied to real-life scenarios, where the perceptual input is characterized by the same features.

Information-dense Perceptual Input Hampers Visual Search and Identification of Relevant Body Parts (Step 1). The stimuli in this task comprise a large amount of irrelevant information. As can be seen from the exemplary screenshots in Figure 3.2, many elements like the colorful patterns on the fish body or the reef in the background are present, which are irrelevant to the task of classifying fish locomotion patterns. Indeed, research on learning from highly realistic representations has shown that irrelevant features are challenging for learners (Scheiter, Gerjets, Huk, Imhof, & Kammerer, 2009), since they may attract attention and thus guide it away from more relevant elements (Dwyer, 1976). Moreover, relevant elements may be blurred or otherwise difficult to detect. Thus, it is likely for people unexperienced in this task to have difficulties to allocate their visual attention to the relevant body parts (e.g., certain fins) and interpret their motion without being distracted by irrelevant information (e.g., colorful patches on the fish's body). Hence, these features of fish swimming in the reef are likely to add to the difficulty for novices to perform the first step of this task.

A related difficulty of the perceptual input used in this thesis is the relation of the visual saliency of visual elements and their thematic relevance. As described by Schnotz and Lowe (2008) ideally both the visual saliency and the thematic relevance should correspond to each other, however, when using real-life scenes – as reef fish swimming forward – this desired relation is not necessarily given. Dynamism, for instance, in itself captures people's attention (Hillstrom & Chai, 2006). This is problematic, however, if the most dynamic features are not the thematically relevant ones (Lowe, 2003, 1999). For the fish locomotion task, thematic relevance implies visual saliency as the relevant body parts are moving and, thus, are visually salient. However, this is not the case vice versa; that is, not all visually salient body parts or other elements of the representation are relevant (e.g., something in the background may move or the fish's fins may move passively driven by the flowing water). Hence, it is very likely that unexperienced people get distracted by irrelevant movements that a fish might execute (i.e., movements not relevant for thrust propulsion to swim forward but for instance for stabilization). Again, this issue adds to the difficulty for unexperienced people to execute the first step of this task.

Dynamism Hampers Visual Inspection and Interpretation (Step 2). Another core feature of the reef fish swimming forward is that they are dynamic, that is, they induce changes of the perceptual input over time. According to Lowe (2003, 2004) temporal change may occur in three different ways: (1) Transformations (appearance changes), which are changes in the size, shape, color, and texture of single elements. This type of dynamism also occurs in the task to classify fish locomotion patterns. For instance, when a fish swims labriform by flapping its pectoral fin, this will appear as if the fin's shape and size are changing. (2) Translations (position changes) refer to the movement of single elements across the scene that the observer views. This type of movement also occurs in the perceptual input of interest in the thesis, either due to the movement of the entire fish (e.g., a fish escaping from the observer's view) or due to the movement of the observer him- or herself. (3) Transitions (inclusion changes) occur when single elements appear or disappear. This type of dynamism may also occur in the perceptual input used in Study 1 and Study 2, when a fish is occluded in part by another object. In addition to distinguishing these different types of dynamism, Lowe (2003, 2004) points out that all types of dynamism may occur for several elements at a time. However, this is not the case for the fish locomotion patterns addressed in Study 1 and Study 2. In the videos used for these studies, either only one relevant body part moves (e.g., caudal fin) or, if two body parts are moving (e.g., dorsal and anal fin), then they move in exactly the same way. Hence, it is sufficient to observe the motion of only one of both parts in detail.

The processing of dynamic information may cause difficulties. For instance, research on learning with dynamic representations yielded that they cannot be accurately perceived if they are too complex or too fast (Tversky, Morrison, & Betrancourt, 2002). The videos of swimming fish used in this thesis include fast motion of single fins. Rapid motion, however, causes high cognitive demands that may complicate information processing (Ayres & Paas, 2007; Lowe, 1999, 2003; K. Meyer, Rasch, & Schnotz, 2009; Tversky et al., 2002). Moreover, dynamic information is transient (i.e., only temporary available) which means that the observer cannot access passed or changed information to compare it to the current status (Hegarty, 1992). Hence, the observer has to keep passed information activated in working memory (e.g., the last position of a fin) to relate it to the current information (e.g., the current position of a fin). To do so, the observer needs already a schema of the observed object to create an internal representation of the passed information so it can be compared to the current information. This, however, may be very difficult, in particular, for novices. In the task of classifying fish locomotion patterns this may imply that inspecting and interpreting the motion of a fast moving body part that often change their shape or position (i.e., the second step of this task) is very difficult, in particular for people with little experience in this task (i.e., novices).

3.1.4 Instructional Goals

To perform the task of classifying fish locomotion patterns the skill to detect and interpret motion of swimming fish is supposedly important to teach. This assumption is supported by three findings from the normative task analysis.

First, the analysis of the task procedure revealed that the task to classify fish locomotion patterns is composed of three steps: detecting which body parts produce propulsion, interpreting how these body parts move, and assigning this information to a locomotion pattern. Two out of these three steps rely on inspecting the perceptual input (i.e., the swimming reef fish). Only the last step can be executed without looking at the fish. This last step, however, cannot be executed correctly without the correct execution of the first two steps. Thus, inspecting the perceptual input and drawing the correct conclusions based on it, is central to this task. Consequently, based on the procedural task analysis, teaching the execution of these steps that are inherently bound to the perceptual input should be a major instructional goal, since they built the basis for successful task performance of the task to classify locomotion patterns of fish.

Second, as the two most important steps in the task to classify fish locomotion patterns require the inspection of a perceptual input, skills that enable those steps should be in focus of an instructional method. Such skills are to visually search and identify the relevant body parts as well as to visually inspect them to interpret their motion. Since these skills require perceptual input, I refer to them as ”perceptual skills” for the purpose of this thesis. The last step can be carried out without the perceptual input and relies on conceptual knowledge only. As pointed out in the introduction to this thesis, there are already well established methods to convey conceptual knowledge, however, hardly any to convey perceptual skills. Hence, teaching perceptual skills should be a main instructional goal in the task to classify fish locomotion patterns.

Third, analyses of the difficulties that are likely to occur in the task of classifying fish locomotion patterns revealed that the stimuli used in Study 1 and Study 2 as perceptual input are very likely to cause difficulties with regard to the above described perceptual skills. As they include many irrelevant information which the observer has to ignore, and as they are dynamic, relevant information may be missed or difficult to detect (for instance, because some motions may appear too fast). Both aspects require sophisticated perceptual skills when relevant body parts are difficult to detect and their motion is difficult to interpret. In sum, difficulties may occur at all three steps of the task. Since two out of three steps of the task to classify fish locomotion patterns require perceptual skills, again perceptual skills are particularly important.

Overall, the normative analysis of the task to classify fish locomotion patterns revealed that the major instructional goal in this task should be to convey perceptual skills. The following chapter reports an empirical task analysis that investigated whether the findings of the normative task analysis are in line with the actual performance of this task.

3.2 Empirical Task Analysis Based on Research Findings – Study 1

The last section provided a normative analysis of the task to classify fish locomotion patterns. According to Schaafstal and Schraagen (1992) conducting a task analysis requires to test and supplement the findings from the normative task analysis with an empirical task analysis, which is in the focus of the current section. The purpose of this empirical task analysis is to investigate the cognitive and perceptual processes of experts and novices when performing this task. In a first step, research conducted on tasks that are similar to the task of classifying fish locomotion patterns – in that all require covert cognitive and perceptual processes – will be reviewed. These findings will be used to generate hypotheses on task performance in the classification of fish locomotion patterns. As argued in Section 2.3.2 perceptual processes are best investigated by means of eye tracking. This method, however, has hardly been used directly for task analysis purposes. Thus, this section refers to literature from expertise difference studies, which use eye tracking or address perceptual processes by other means, in order to shed light on three issues: First, the task *procedure* actually used by experts in perceptual tasks, that is, strategies that experts are likely to use, will be addressed. Second, the *knowledge and skills* required to perform perceptual tasks, that is, attending to relevant information in the stimulus, will be addressed. Third, it will be addressed which *difficulties* are likely to occur in perceptual tasks for novices, like being distracted by salient but irrelevant information. Finally, it will be stressed what the limitations of the available findings are, namely that most of the studies investigated tasks that require the examination of *static* stimuli (i.e., information is not transient as it is the case for the stimuli used in this thesis) and that these studies did not address *perceptual strategies* directly. Based on limitations of available findings, a study will be carried out on the task of classifying fish locomotion patterns to complete the empirical task analysis as the first study of this thesis. For an overview of studies investigating expertise differences in perceptual tasks and on the stimuli they used and methods they applied see Table 3.3.

3.2.1 Empirical Findings on Covert Processes Involved in Perceptual Tasks

Expertise research has found several well established effects: Experts solve problems better and faster (De Groot, 1946/1978; Reingold & Charness, 2005). A reason for those effects is that experts are characterized by a large amount and good organization of knowledge (Chase & Simon, 1973; Gobet & Simon, 1996, 2000). From these findings it can be assumed for the empirical task analysis in the given task that experts will classify the fish locomotion patterns more quickly and more often correctly than novices, because their knowledge is organized in an efficient way. The mere analysis of the task outcome, however, does not suffice for a thorough analysis of the task. This is particularly the case, since the last section has shown (see Section 3.1) that the crucial steps in the task performance are of covert nature (i.e., cognitive and perceptual). Thus, in the following conclusions that can be drawn from investigating perceptual and cognitive processes will be discussed.

3.2.1.1 Task Procedure: Strategies that Experts Are Likely to Use

The normative analysis of the task procedure described one way of how the task should be executed (cf. Section 3.1.1). However, when actually performing a task, people with different levels of expertise may use different strategies (i.e., use different task steps or orders of task steps) to perform a task, as mentioned in Section 2.2.1.2. There are at least two open questions according to this issue as will be stressed in the following.

First, one may ask whether experts' strategies are in line with those postulated in the normative task analysis, that is, whether experts use supposedly optimal strategies that also a novice should use or whether experts act within their domains in a very different way than novices should do. For tasks that require cognitive processes only it has been shown that experts' highly integrated knowledge structures enable them to sometimes use **shortcuts** during task performance (e.g., in medical diagnosis based on written descriptions of patient cases: Boshuizen & Schmidt, 1992). These shortcuts are sometimes referred to as "knowledge encapsulation", that is, several steps of a task are represented and thus used as one step only (H. G. Schmidt & Boshuizen, 1992). The use of shortcuts by experts results in very different strategies for different expertise levels. This finding could also be shown for the use of cognitive processes in a perceptual task, namely navigation of river pilots (Schraagen, 1993). In contrast, in a medical perceptual task it has been found that experts process and report visual case information more elaborately than novices (e.g., for diagnosing X-ray pictures see Lesgold, 1984), which rather seems to favor the idea that experts optimize strategies that are used imperfectly by novices.

With regard to the task of classifying fish locomotion patterns it might be assumed that experts will be able to use knowledge-based shortcuts: The normative, procedural task analysis in Chapter 3.1.1 described an approach to the task what can be characterized as a *locomotion description strategy* and that is taught to university students of biology when they receive training in marine zoology. An application of this strategy would lead the observer to attend only to those parts of the fish body that are crucial for the fish's locomotion (i.e., the fins and the body) and ignore the parts that are irrelevant to the locomotion (i.e., eyes or colorful patches). Both novices and experts have to rely on the locomotion description strategy when classifying the locomotion of unfamiliar fish. However, with familiar fish, experts may be able to automatically retrieve knowledge on the specific locomotion pattern associated with this particular fish species from memory. In such a case a *species classification strategy* might be used, where other features of the fish become potentially relevant. Specifically, some species can be easily recognized due to salient static features (e.g., a specific colorful pattern on the fish's body characteristic of a particular fish species). Experts can use these features to classify a fish, and upon activating that particular schema, the knowledge on its locomotion pattern might become automatically activated. For novices, this type of knowledge-based shortcut is not available, which is why they need to rely on the locomotion description strategy.

Second, it is unclear whether experts' perceptual strategies are the same for all experts or whether they differ among experts (cf. Medin et al., 2006). For example, Medin, Lynch, Coley, and Atran (1997) showed that while some experts were very similar in their approaches to the task of categorizing trees (namely, parks maintenance personnel vs. scientific taxonomists), other experts differed in their approach to the same task (namely, landscapers vs. scientific taxonomists). Moreover, Medin et al. (2006) showed in another study that even in well-structured domains, like the categorization of freshwater fish, expertise does not always lead to shared conceptualizations (while some experts categorized freshwater fish into the same categories, others did not). The authors concluded that even if the outcome of a categorization process is similar across experts, the underlying processes might not necessarily be so. This conclusion, however, might be due to the characteristics of biological taxonomies in which one has to deal with a diversity of features of the various species. Taxonomies are often invented, that is, they are conceptual schemas not inherent to a domain. In particular in biology there are multiple categories for the same species depending on the focus of the taxonomy (e.g., morphology, genetics, etc.). Furthermore, categorizing species is based on multiple features, which do not need to be considered in a hierarchical order. Experts may differ in the features that they emphasize in order to achieve a categorization depending on their prior experiences, that is, on their learning history. For instance, an expert with lots of outdoor experience will potentially focus on other features (e.g., those easily observable in a natural setting) than an expert who mostly deals with formalin preparations and textbooks (e.g., features requiring a close inspection). Hence, these differences in learning history, which will be reflected in the organization and accessibility of knowledge, may affect how experts attend to features (so called top-down or exogenous guidance of visual attention, Corbetta & Shulman, 2002; Posner, 1980) yielding diverse perceptual strategies for experts. On the other hand, novices' attention may be mainly guided by features of the stimulus, for instance, by salience (so called bottom-up or exogenous guidance of visual attention, Lowe, 1999). Consequently, their perceptual strategies might be rather homogeneous compared to those of experts.

3.2.1.2 Knowledge and Skills Required to Perform this Task: Attending to Relevance

The normative analysis of the knowledge and skills required to perform the task of classifying fish locomotion patterns stressed the importance of perceptual skills (cf. Section 3.1.2).

Hence, findings on perceptual aspects of expertise will be reviewed in this section. Many studies that have addressed the issue of how experts perceive and interpret visual representations used the eye tracking methodology for this purpose (e.g., Antes & Kristjanson, 1991; Charness, Reingold, Pomplun, & Stampe, 2001; Vogt & Magnussen, 2007). In particular, these studies provide information on attention allocation through eye movement analyses.

As mentioned above, allocation of visual attention is guided by exogenous and endogenous cues (Stelmach, Campsall, & Herdman, 1997). Individuals with a higher level of expertise allocate their attention based more on endogenous cues (i.e., they know what features are important) and therefore process complex visual material more efficiently, for example, by focusing faster or longer on the relevant features while ignoring potentially salient, but irrelevant information (e.g., Canham & Hegarty, 2010; Charness et al., 2001; Haider & Frensch, 1999; Lesgold et al., 1988; Lowe, 1999; Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003; Van Gog, Paas, & Van Merriënboer, 2005). Haider and Frensch (1999) stated in their information-reduction hypothesis that with increasing expertise people learn to distinguish between relevant and irrelevant information and therefore concentrate on processing mostly relevant information. Using a letter string task in which the location of relevant information was varied, Haider and Frensch (1999) corroborated this hypothesis with eye movement data. This was confirmed in several further studies. In the domain of art, for instance, Antes and Kristjanson (1991) found that experts (i.e., artists) compared to novices (i.e., non-artists) had higher fixation densities on important aspects of the paintings. Charness et al. (2001) also found in their studies of expertise effects in chess performance that experts had a greater proportion of fixations on relevant rather than on irrelevant areas.[2] Hence, it is very likely that experts in the domain of marine zoology also possess sophisticated perceptual skills enabling them to direct attention quickly to those areas in the fish videos that are relevant to perform the task of classifying fish locomotion patterns, while ignoring other areas.

[2] In both studies (i.e., Antes & Kristjanson, 1991; Charness et al., 2001) relevant areas were determined a priori by an independent expert in the field of study. The same procedure has been applied for all three studies within this thesis.

3.2.1.3 Difficulties in this Task: Distraction by Saliency

The normative analysis of difficulties that may occur in the task of classifying fish locomotion patterns showed that novices are very likely to experience difficulties in visually searching for the relevant body parts and in interpreting their motion due to the features of the perceptual input. From an empirical perspective, it can be assumed that to the same extent to which experts possess the skill to attend quickly and for a sufficient time to relevant information within a visual stimulus, novices might fail in those skills (see references above). Often, novices do not only miss relevant information, but also get distracted by information that is salient but irrelevant for task performance (Lowe, 1999). That is, novices' attention allocation is mostly guided by exogenous cues (Stelmach et al., 1997). In the task of classifying fish locomotion patterns it is also very likely that novices will get distracted by salient, but irrelevant features, like colorful patches on the fish's body.

Table 3.3:
Overview of Studies Showing the Superiority of Experts in Attending to Relevant Areas

Eye tracking	
Domain	Reference
Static stimuli	
Letter-string task	Haider and Frensch (1999)
Chess	Charness et al. (2001); De Groot and Gobet (1996); Gobet and Charness (2006); Jongman (1967); Reingold and Charness (2005)
Electrical circuit	Van Gog, Paas, and Van Merriënboer (2005)
Lesion detection in mammographic images	Krupinski (1996); Nodine, Kundel, Lauver, and Toto (1996)
Virtual slide microscopy	Krupinski et al. (2006)
Weather maps	Canham and Hegarty (2010)
Art	Antes and Kristjanson (1991); Vogt and Magnussen (2007)
Dynamic stimuli[1]	
Car driving	Underwood et al. (2003)
Sports	Goulet, Bard, and Fleury (1989); Helsen and Starkes (1999); Moreno et al. (2002); Ripoll (1991)

[1] In all presented studies either biased parameters and imprecise mobile eye trackers were used, or only single basic eye tracking parameters were reported.

Verbal reports	
Domain	Reference
Static stimuli	
Chess	Charness (1981, 1991); De Groot (1946/1978)
Physics problem-solving	D. P. Simon and Simon (1978)
Sport situation descriptions	French et al. (1996); Ward, Hodges, Williams, and Starkes (2004)
Pulley systems	Chi, Feltovich, and Glaser (1981)
Electrical circuit	Van Gog, Paas, and Van Merriënboer (2005)
X-rays and skin disorders	Lesgold et al. (1988); Myles-Worsley, Johnston, and Simons (1988); Norman, Coblentz, Brooks, and Babcook (1992); Wolf et al. (1994)
Diagnosis of microscopic pathology	Crowley, Naus, Stewart, and Friedman (2003)
Distinguish types of rocks	Burton, Shadbolt, Hedgecock, and Rugg (1988)
Evaluating helicopters' routes	Geiwitz, Klatsky, and McCloskey (1988)
Incorporating historical documents	Wineburg (1991)
Dynamic stimuli	
Weather maps	Lowe (1999)[2]
Perceptual diagnosis of ECG	Simpson and Gilhooly (1997)
Films of sport events	Pinheiro and Simon (1992); Ste-Marie and Lee (1991)
Video-tape of a classroom lesson	Sabers, Cushing, and Berliner (1991)

[2] This study did not investigate verbal, but written reports. Still, the findings are in line with the remaining studies.

3.2.1.4 Limitations of Available Findings

Many of the reported findings on perceptual tasks can be used to draw conclusions for the current empirical task analysis. However, as these studies investigated tasks that require the examination of static stimuli, their empirical findings do not entirely fit to the task investigated in this thesis. Thus, these conclusions have to be regarded as mere assumptions and need to be verified by means of investigating dynamic stimuli. Moreover, strategies used in the task procedure, like knowledge-based shortcuts, have not been investigated on a perceptual level yet. Hence, hardly any assumptions can be made on how an actual task procedure of experts and novices would be in a perceptual task. This issue needs to be investigated directly by means of eye tracking. Both issues are addressed in the following in more detail.

First, the above reported studies used *static* representations. Dynamic representations – like the dynamic information used in this thesis – may, however, evoke different processes in the observer. Research on learning from visualizations, for instance, often found that learning from static pictures differs from learning with videos or animations (for a recent meta-analysis see Höffler & Leutner, 2007). Only a few eye-tracking studies have been conducted, though, that investigated expertise differences on dynamic stimuli and these studies only used single basic parameters. For instance, Moreno et al. (2002), investigated novice and expert gymnastic coaches inspecting videos of gymnastic techniques and indicating errors in performance. They found that experts had longer and fewer fixations than novices, and attributed this to the fact that experts attended more to informative (i.e., relevant) areas and ignored uninformative (i.e., irrelevant) ones. However, this assumption was not tested in the study, because eye movements were not analyzed in terms of which positions on the stimulus participants looked at (i.e., no distinction between relevant and irrelevant information was made in the analysis of eye movement parameters). Rather the authors interpreted only single basic parameters, like fixation times and *assumed* that experts to have looked at relevant areas.

Other studies that made the attempt to investigate not only single basic parameters, but moreover, to assign eye tracking parameters to positions on the stimulus, that is, to investigate at which position a person looked at, calculated *fixation* parameters on dynamic stimuli (e.g., Underwood et al., 2003). However, this is a very problematic analysis of eye tracking data from dynamic stimuli as was discussed in Section 2.3.2.1. For instance, it is very likely that eye movements following a moving object (so-called smooth pursuit) are wrongly assumed to be eye movements that relocate the eyes' focus of attention during which the eye does not see (so-called saccades). Hence, findings from studies reporting fixation parameters while using dynamic stimuli are probably not valid to a certain extent. Moreover, these studies were conducted with mobile, head-mounted eye trackers that do not allow for a detailed analysis of eye movement data in relation to specific spatial and temporal events (see Section 3.2.2.4). Instead, this methodology only allows for a coarse, manual coding of eye movement data that are rather imprecisely overlaid on a video of the viewed stimulus. Thus, an eye tracking analysis of perceptual processes involved in examining a dynamic stimulus is missing so far.

Second, studies investigating cognitive processes have shown that experts use specific strategies, like shortcuts. The answer on the question, whether experts use similar or different strategies seems to differ across tasks and types of expertise. However, these detailed characteristics of experts' strategies have not been investigated for perceptual processes, yet. Studies on perceptual processes of experts and novices investigated so far only single basic eye tracking indicators (e.g., number or duration of fixations) when inspecting dynamic stimuli. Research on identifying expertise effects on perceptual *strategies* that is, complex patterns of eye movements when inspecting dynamic stimuli, however, is still lacking. Hence, it is not clear whether experts would use shortcuts while performing the task to classify fish locomotion patterns also on a perceptual level and whether they would be rather similar or diverse when doing so.

In sum, although research on expertise differences in attention allocation on visual representations has provided interesting findings, none of these studies matches exactly the needs for an empirical task analysis for the task of classifying fish locomotion patterns. That is, not enough is known (1) on perceptual strategies that experts use in the procedure of dynamic, perceptual task, (2) on perceptual skills, like attending to the relevant information, that are required in dynamic stimuli, (3) and on difficulties, like being distracted by salient but irrelevant information, that may occur in inspecting dynamic stimuli. Thus, the first study of this thesis investigates eye movement data of experts and novices in terms of a detailed assignment of spatial events ("dynamic" AOIs, see 3.2.2.4), the appropriate temporal events (raw data in terms of the temporal dimension), and an analysis of the perceptual strategies used, by means of a monitor-fixed eye tracker with a higher spatial and temporal resolution, which allows for detailed analyses of the eye movement data.

3.2.1.5 Research Questions - Hypotheses

The aim of the study is to verify and supplement the assumptions derived from the literature overview on expert-novice differences in conducting perceptual tasks. The following hypotheses were tested for this purpose.

It was hypothesized (Hypothesis 1) that experts would perform more accurately and faster than novices on a classification of fish locomotion patterns (reflected in higher correctness rates and shorter mean viewing times of the stimulus)[3]. The respective test, however, mostly served as a manipulation check concerning the assignment of individuals to different levels of expertise.

More important, the task **procedure** was investigated. This was addressed by the following three hypotheses:

Hypothesis 2 addressed the question, whether hints on the use of shortcuts in the task procedure could be found in experts. In particular, it was expected that experts would verbalize less information than novices due to knowledge encapsulation and thus use fewer words in their description of how they accomplished that task[4]. This would hint toward the fact the experts do not need to execute as many steps as novices, but rather chunk several steps into one. In line with this reasoning, their verbalizations were expected to contain more encapsulating technical terms[5].

Next, it was investigated, whether experts would use a specific shortcut strategy in their task procedure: It was explored whether experts attend to features that are relevant for either locomotion description or use shortcuts and thus, attend to features that are relevant for species classification by means of analyzing the gaze duration on these features and amount

[3] e.g., De Groot (1946/1978); Reingold and Charness (2005)
[4] Ericsson and Simon (1980)
[5] Boshuizen and Schmidt (1992)

of utterances related to these features. It was predicted that experts would attend longer to and mention more often features relevant to the species than novices (Hypothesis 3).

Moreover, it was investigated whether expertise yields either more diversity or more homogeneity in terms of the perceptual strategies by analyzing experts' and novices' gaze sequences. It was assumed that novices' perceptual strategies would be guided by the salience of single features leading to a more homogeneous gaze pattern. On the other hand, experts' perceptual strategies were assumed to be controlled more endogenously. Hence, experts were expected to be characterized by rather heterogeneous gaze patterns depending on their individual learning history that would lead to different processing strategies (Hypothesis 4).

Finally, one hypothesis referred to both, the **skills and knowledge** required (as shown in the process data of experts) as well as to **difficulties** (as shown in the process data of novices) that occur when performing this task:

Based on prior findings with static, visual representations[6] it was hypothesized (Hypothesis 5) that also when using dynamic videos the process data of experts would show that they attend more to relevant information than novices, who would rather attend to perceptually salient, but potentially conceptually irrelevant information[7].

3.2.2 Method

3.2.2.1 Participants and Design

Participants were 21 individuals (10 women, 11 men, $M_{age} = 26.57$ years, $SD = 5.98$) with two different levels of expertise. All participants had normal or corrected-to-normal vision. Of them seven were experts, that is, professors, PhDs, or advanced PhD students, with a mean age of 31.43 years ($SD = 8.54$). The novices were 14 biology students of the University of Tuebingen, Germany, with a mean age of 24.14 years ($SD = 1.51$). The novices had basic knowledge of fish anatomy and terminology, but had very little, if any experience with

[6] Antes and Kristjanson (1991); Charness et al. (2001); Haider and Frensch (1999)
[7] cf. Lowe (1999)

classifying different locomotion patterns. It was determined via a questionnaire that experts were more interested in fish and had more relevant practical experience with fish locomotion (i.e., they engaged more frequently in snorkeling and diving than novices), because not only years of experience in a domain, but also the nature of this experience is important (Ericsson & Lehmann, 1996). Participants were compensated €10 for their participation.

3.2.2.2 Apparatus and Materials

Eye Tracking Equipment. Participants' eye movements were recorded with a Tobii 1750 remote eye tracking system with temporal resolution of 50 Hz on a 17" monitor using ClearView 2.7.0 software (www.tobii.com). Participants' verbal data were recorded by Camtasia 3.0 software using a standard microphone attached to the stimulus PC. The eye tracking data were analyzed with ClearView 2.7.0 software and self-programmed ruby tools (cf. Appendix Study 1).

Stimulus Materials. The materials consisted of four digital videos (for screen shots see Figure 3.3) in an audio video interleave format (.avi), sized 720 × 576 pixels (corresponding to 9.6 × 7.68 inches). Each video depicted a single fish swimming, whereby each fish deployed a different locomotion pattern (i.e., tetraodonti-, subcarangi-, labri-, and balistiform) that had to be described by the participants. The videos were rather short (8.79 s on average), but looped automatically until participants stopped them. This was done to avoid an artificial situation for both groups. On the one hand, novices might not be able to describe the locomotion pattern at all after a too short presentation. On the other hand, if experts were forced to look at a stimulus that they had interpreted already, they might start to look at completely irrelevant features just out of boredom.

Verbal Reports. Cued retrospective reporting is a verbal reporting procedure in which participants verbalize their thought processes during task performance after completing

Figure 3.3: Screenshots from the four videos used in the study.

the task, based on a cue of their performance. The cue used here consisted of the videos with the recordings of participants' own eye movements superimposed onto the video (Van Gog, Paas, Van Merriënboer, & Witte, 2005). This so called "gaze replay" showed fixations which included all eye movements within a range of 50 pixels and a duration of at least 200 ms. Moreover, a so-called gaze trail of 500 ms was included. This gaze trail showed the path the eyes have passed within the last 500 ms. For an example screenshot of the gaze replay see Figure 3.4. The gaze replay was in real time and although participants were not allowed to pause it, they could watch it again.

3.2.2.3 Procedure

The experiment was run in individual sessions of approximately 20 min. At the beginning, the eye tracking system was adjusted to the individual features of the participant based on a nine-point calibration. Before watching the videos, participants received the following instruction: "While watching the video, please take a look at the way the fish swims. Sub-

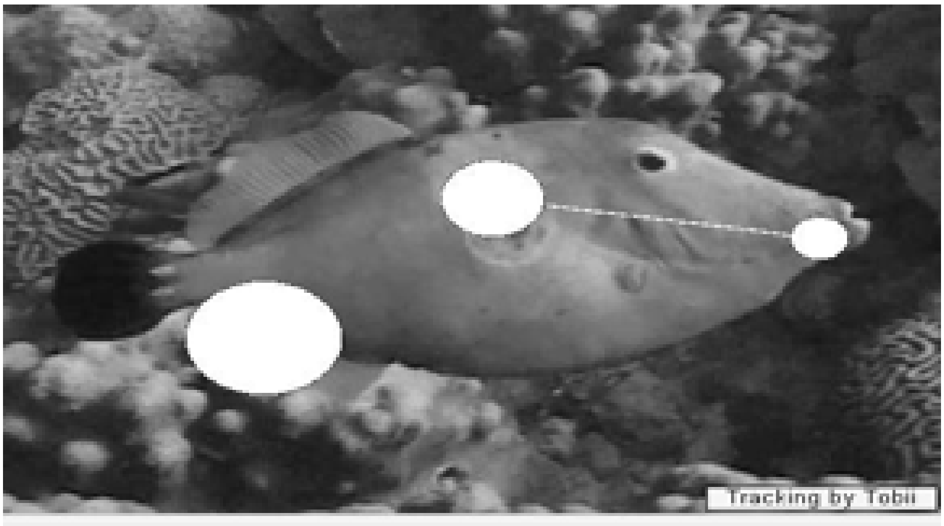


Figure 3.4: Screenshot of an exemplary gaze replay (gaze plot on 100 ms segment). In the replay, only one dot indicating the fixation was visible at a time and moved across the screen (suggested here by the different dots).

sequent to watching the video, you will be asked to describe the fish's locomotion pattern. You will be allowed to watch the video as often as you like."[8] Then, participants watched the looped video while their eye movements were recorded until they stopped it themselves. After having watched the video, participants were asked to describe the locomotion pattern of the depicted fish verbally. Subsequently, they were asked to provide the cued retrospective reports (Van Gog, Paas, Van Merriënboer, & Witte, 2005), for which participants received the following instruction: "In the following you will see a red dot moving over the screen. This is the recording of your eye movements. The lines that are drawn by the dot represent the path of your eye movements. The size of the dot corresponds to your gaze duration, that is, the bigger the dot is the longer you have looked at this point. Please watch the replay and tell me what you were thinking during your first viewing". This procedure was repeated for all four videos.

[8]This wording is a bit less specific than the request to classify the fish's locomotion. Nevertheless, experts understood this wording as a request to classify the fish's locomotion, while novices were still able to approach the task. If the wording of the instruction would include a specific request for a classification, novices would have not been able to perform this task at all as they have no knowledge on this type of classification.

3.2.2.4 Data Analysis

Duration and Correctness of Description. The first dependent variable was the mean viewing duration per video. The second dependent variable was the correctness of the description of the locomotion pattern after watching each video. Based upon Lindsey's (1978) description of fish locomotion patterns the following coding system for the correctness of description was applied. Naming the correct technical term for describing the locomotion pattern yielded 1 point, whereas naming a wrong technical term or none yielded 0 points. Two other dependent variables were obtained from the descriptions of the locomotion pattern. First, it was coded whether participants described correctly *which* part of the fish body had been moving. Second, it was coded whether participants described correctly *how* the part of the fish body had been moving. With regard to the latter two aspects, the technically correct and complete answers yielded 1 point each, a correct but colloquial or only partially complete description yielded 0.5 point, and everything else yielded 0 points. Thus, participants could receive a maximum of 1 point for each of the three dependent variables per video. The coding system is summarized in Table 3.4. Two raters conducted the coding of the verbal data independently, and showed complete agreement.

Eye Movement Data.

Preparation of Eye Tracking Data from Dynamic Stimuli. The eye tracking recordings for this thesis resulted in very information dense data, namely an x- and y-coordinate every 20 ms or 5 ms, respectively[9]. For further calculations, these data had to be assigned to different locations on the dynamic stimulus and then to be aggregated into different types of eye movement parameters.

Eye movements were aggregated according to **spatial events**, that is, eye movement parameters were assigned to locations on the videos. In general, stimuli can be divided into semantic areas of interest (AOIs). AOIs are precisely specified areas of an object (e.g., a fin,

[9]The density in time depends on the temporal resolution of the eye camera, which differed for Studies 1 and 2 (20 ms) versus Study 3 (5ms).

Table 3.4:
Coding System for the Locomotion Pattern Description

	Technical term?	Which part moves?		How does it move?	
	1 point	1 point	0.5 point	1 point	0.5 point
Video 1	Tetraodonti	Anal and dorsal fin	e.g., upper and lower fin	Oscillating	e.g., moving like a paddle
Video 2	Subcarangi	Caudal fin	e.g., back part	Undulating	e.g., wavelike
Video 3	Labri	Pectoral fin	e.g., frontal fin	Oscillating	e.g., moving like a paddle
Video 4	Balisti	Anal and dorsal fin	e.g., upper and lower fin	Undulating	e.g., wavelike

a patch of colorful stripes, the mouth) for which it can be determined whether and for which amount of time a participant is looking at this area. Each part of the fish body was an AOI (all fins, the body, and salient features, like eyes, colorful stripes, or taints). For this thesis, the AOIs were selected with the help of domain experts.

When using video as stimuli, where there are many moving objects, static AOIs are often of little use. Thus, ”dynamic” AOIs have to be created, which follow the form, size and position of the objects as they move. As all calculations for this thesis had to be performed on temporally raw data and on dynamic AOIs, manufacturer-built in calculations could not be used. Instead, data preparation and aggregation was done manually and with self-programmed tools as described in the following. Each video was divided into segments of 100 ms each. For each segment AOIs were defined manually. The length of the segments was determined based on the time for which AOIs did not change within the stimulus used in this study. In a first step, the data for each AOI were aggregated per video (i.e., across all 100 ms segments). In a second step, the data for the AOIs were aggregated across all four videos according to whether the body part represented by an AOI was: (a) relevant for the locomotion description strategy only; (b) relevant for the species classification strategy only; (c) relevant for both strategies at the same time; or (d) irrelevant for both classification strategies (see Table 3.5). The assignment of AOIs to these four categories was determined a priori by a domain expert (cf. Antes & Kristjanson, 1991; Charness et al., 2001). We refer to these aggregated AOIs as AAOIs in the remainder of the present chapter. The AAOIs were not all equal in size; however, this is not problematic as comparisons were made only between groups, instead of between AAOIs within each group.

For the videos used in the studies on fish locomotion – but also in Study 3 on infant motion – smooth pursuit eye movements are very likely to occur (i.e., movements when the eyes follow a motion, Dodge, 1903). As already stated above (Section 2.3.2.1), these types of eye movements cannot be detected by contemporary eye tracking software (Holmqvist et al., 2011). As a consequence, for the studies conducted in the course of this thesis it was not possible to revert to well established algorithms that allow to aggregate the large amount of eye tracking data into eye movement parameters, like fixations ans saccades. Hence, all calculations had to be performed on raw data with respect to the temporal distribution.

This procedure resulted in a very time-consuming manual data preparation for each video example. Thus, the duration and the number of videos used in the studies of this thesis had to be reduced to a minimum.

Resulting Dependent Variables. The first dependent variable was the distribution of gazes across the four different AAOIs for the first 4 seconds. This is the time for which each participant had at least watched each video (i.e., minimum viewing time was 4.15 s). This limitation was induced to render gaze durations comparable across participants, even though participants watched the videos for different amounts of time. To calculate the second dependent variable, sequence analyses for the different AAOIs were conducted to compare series of gazes to each other based on the so-called Levenshtein distance (Levenshtein, 1966). The Levenshtein distance is a measure used in computer science to assess the edit distance between two character strings. This edit distance is the minimal number of edit operations needed to transfer one string into another. Edit operations are insertions, deletions, or substitutions of single characters. The edit distance can be converted into a percentage value. In the present study the Levenshtein distance was used to compare the sequences of the gaze locations across participants with regard to their similarity (cf. Feusner & Lukoff, 2008).

To determine the Levenshtein distance, a string of AAOIs that a participant had looked at in a given order (e.g., AAOI 1, AAOI 3, AAOI 1, AAOI 4, ...) was the input data. The number of edit operations (i.e., insertions or deletions of AAOIs) needed to transform the AAOI sequence of this participant into that of another participant describes the similarity between the two sequences of AAOIs. The Levenshtein distance was determined for the gaze sequences of experts as well as for novices to analyze the similarity of the strategies used within groups. In particular, Levenshtein distances were computed for pairs of AAOI strings, whereby each person provided one AAOI string. This procedure resulted in a similarity score for each possible pair of experts and for each possible pair of novices. The calculation of the distribution of gazes and sequence analysis required self-programmed tools, which are described in Appendix Study 1.

Verbal Reports. The verbal reports obtained during watching the gaze replay were analyzed with regard to different aspects. First, the contents of the participants' initial utterances (i.e., the first term mentioned) of each gaze replay were analyzed to determine whether they referred to either relevant or irrelevant features for both classification strategies (see Table 3.5) in order to investigate participants' initial response to the task. As a second variable, it was analyzed for the complete gaze replay whether the verbal reporting referred to features that were either relevant according to one or both classification strategies, or irrelevant according to both strategies (see Table 3.5). Third, the total number of words used during retrospection was counted. Fourth, the number of different technical terms that were used during retrospection was determined.

Table 3.5:
Overview of Relevant and Irrelevant Parts of the Presented Fish According to the Locomotion Description Strategy and the Species Classification Strategy for All Four Videos

	Relevant	Irrelevant	Technical term
		Locomotion description strategy	
Video 1	Dorsal and anal fin	Pectoral, caudal fin, and eye	Tetraodonti
Video 2	Caudal fin	Dorsal, anal, pectoral fin, and stripes	Subcarangi
Video 3	Pectoral fin	Dorsal, anal, caudal fin, and eye	Labri
Video 4	Dorsal and anal fin	Caudal, pectoral fin, eye, and taint	Balisti
		Species description strategy	
Video 1	Dorsal, anal, pectoral fin, and eye	Caudal fin	Arothron hispidus
Video 2	Stripes	Caudal, dorsal, anal, and pectoral fin	Anisotremus virginicus
Video 3	Pectoral fin	Dorsal, anal, caudal fin, and eye	Thalassoma lunare (female)
Video 4	Taint	Dorsal, anal, pectoral, caudal fin, and eye	Cantherines macrocercus

3.2.3 Results

For all statistical tests reported here a significance level of .05 is used.

3.2.3.1 Performance: Duration and Correctness of Description

Novices ($M = 24.69s, SD = 12.15$) had significantly longer mean viewing times for the four videos than experts ($M = 10.93s, SD = 4.06$), $F(1, 19) = 8.32, p < .01, \eta_p^2 = .31$.

Results of a 2 (expertise) × 3 (performance) MANOVA, in which the three performance measures were naming the locomotion pattern correctly, describing which body part had been moving, and describing how each body part had been moving, showed a main effect of expertise, $Pillai's\ trace = .58, F(1, 19) = 7.32, p < .01$, indicating that expertise positively affected the overall correctness of the description of the locomotion pattern. The univariate analyses showed that experts mentioned some technical terms for locomotion pattern ($M = 0.39, SD = 0.32$), whereas none of the novices did $(M = 0.00, SD = 0.00), F(1, 19) = 20.82, p < .01, \eta_p^2 = .54$. All experts were able to identify the body part relevant to the locomotion pattern ($M = 1.00, SD = 0.00$), and although the novices' performance was also quite good in this respect ($M = 0.77, SD = 0.31$), it did not reach that of experts, $F(1, 19) = 3.89, p = .06, \eta_p^2 = .18$. Experts and novices did not differ in the way they described how the relevant parts of the fish body moved (for experts $M = 0.29, SD = 0.26$; for novices $M = 0.23, SD = 0.22), F < 1, ns$.

3.2.3.2 Eye Movement Data

An ANOVA with expertise as independent variable and gaze duration for the first four seconds of each video for each of the four AAOI types as dependent variable was conducted. The ANOVA revealed that experts looked significantly longer than novices on the AAOI relevant for the species classification strategy, $F(1, 19) = 7.69, p = .01, \eta_p^2 = .29$, and marginally longer on the AAOI relevant for both strategies, $F(1, 19) = 3.10, p = .095, \eta_p^2 = .14$. Gaze durations on the AAOI relevant

Table 3.6:
Means (and SD) in Milliseconds for the Four AAOIs as a Function of Expertise Level Per Video

AAOI	Experts	Novices
Distribution of gazes for the first 4 s of each video		
Relevant for species classification	375.00 (234.33)	160.36 (124.60)
Relevant for locomotion pattern description	85.71 (93.02)	122.50 (143.56)
Relevant for both strategies	268.57 (47.85)	204.27 (89.72)
Irrelevant for both strategies	2054.30 (280.53)	2336.80 (216.85)

for locomotion pattern description did not differ between groups, $F < 1, ns$. Finally, novices looked longer on the AAOI irrelevant for each strategy compared to experts, $F(1, 19) = 6.53, p = .02, \eta_p^2 = .26$. Means and standard deviations are displayed in Table 3.6.

Sequence Analysis. An ANOVA was conducted with the similarity scores as dependent variable and the type of comparison (comparisons among experts vs. comparisons among novices) as the independent variable. A significant main effect of type of comparison was found on Levenshtein distance, $F(1, 107) = 9.28, p < .01, \eta_p^2 = .08$. The group of experts ($M = 67.41\%; SD = 5.62\%$) was less similar than the novices group ($M = 72.09\%; SD = 6.49\%$), that is, the gaze behavior of experts was more heterogeneous than that of novices.

3.2.3.3 Verbal Reports

First, the initial utterances during watching the gaze replay were analyzed by conducting ANOVAs with expertise as independent variable. Experts mentioned contents relevant for classifying the fish species significantly more often ($M = 1.57, SD = 1.27$) than novices ($M = 0.21, SD = .43$), $F(1, 19) = 13.53, p < .01, \eta_p^2 = .42$. However, the experts did not differ from novices in the case of mentioning content relevant for classifying the locomotion pattern, $F < 1, ns$, of contents relevant for both classification strategies, $F(1, 19) = 3.07, p = .10$, or of irrelevant content, $F < 1, ns$. Means and standard deviations are displayed in Table 3.7.

Second, ANOVAs, with expertise as independent variable, were conducted in order to investigate utterances of the participants that occurred during the whole task performance. Experts mentioned features that were relevant to the locomotion description strategy significantly more often than novices, $F(1, 19) = 23.84, p < .01, \eta_p^2 = .56$, as well as features that were relevant according to the species classification strategy, $F(1, 19) = 9.11, p = .01, \eta_p^2 = .32$. However, there were no significant differences for features relevant to both strategies, $F(1, 19) = 1.49, p = .24$, and irrelevant statements, $F < 1, ns$. Means and standard deviations are displayed in Table 4.

Third, an ANOVA for the overall number of words used during retrospection, with expertise as independent variable, revealed no differences between experts ($M = 242.14, SD = 110.54$) and novices ($M = 247.36, SD = 151.44$) $F < 1, ns$. Finally, a similar ANOVA for the complete gaze replay showed that experts tended to use more technical terms during retrospection ($M = 16.29, SD = 4.75$) than novices ($M = 11.14, SD = 5.78$), $F(1, 19) = 4.12, p = .06, \eta_p^2 = .18$.

Table 3.7:
Means (and SD) in Percent for the Four Content Types as a Function of Expertise Per Video in the Two Time Slots

Content type	Experts	Novices
Initial utterance[a]		
Relevant for species classification	39.00 (31.75)	5.25 (10.75)
Relevant for locomotion pattern description	21.50 (22.50)	28.50 (23.75)
Relevant for both strategies	14.25 (19.75)	28.50 (16.50)
Irrelevant for both strategies	7.25 (12.25)	12.50 (16.25)
Utterances for entire task performance[b]		
Relevant for species classification	57.25 (23.75)	26.75 (20.75)
Relevant for locomotion pattern description	78.50 (17.25)	50.00 (9.75)
Relevant for both strategies	50.00 (0.00)	42.75 (15.25)
Irrelevant for both strategies	60.75 (43.00)	75.00 (26.00)

[a]*Note.* These values represent the percentage of participants stating an utterance in one of theses categories. Values do not add up to 100%, because off-topic utterances were not taken into account.
[b]*Note.* These values represent the percentage of utterances stated per category by each participant. A maximum of 100% for each category could have been achieved, if the categories were exhaustively described for each video.

3.2.4 Discussion

The present study aimed at verifying and supplementing the assumptions derived from the iterative overview on expert-novice differences in conducting perceptual tasks. For this purpose, expertise differences in classifying fish locomotion patterns were investigated, both at a performance and a process level. It was hypothesized that experts would perform more accurately and faster than novices (Hypothesis 1) and that experts would verbalize less information than novices and thus use fewer words, because they use more encapsulating technical terms during retrospection (Hypothesis 2). Furthermore, Hypothesis 3 addressed the issue of which relevant features the experts would use (i.e., for locomotion description strategy or species classification strategy, or both). Hypothesis 4 addressed the issue of whether expertise would yield either more diversity or more homogeneity in terms of the perceptual strategies used. Finally, Hypothesis 5 predicted that the process data of experts would show that they attended more to relevant information than novices, who would attend more to irrelevant information.

In line with Hypothesis 1, performance differences in favor of experts were verified. Experts were able to perform the task faster, as indicated by shorter mean viewing times of the stimuli, and more accurately, as indicated by their better description of locomotion patterns and their higher use of correct technical terms. As expected, novices were not able to do so. Not only did they not use technical terms, but they were also less able to identify the parts of the fish body relevant for the locomotion pattern displayed. While both experts and novices refrained from appropriately describing how the crucial body parts moved, experts named the correct technical term for the locomotion pattern, which by definition encapsulates this information.

The main focus of this study, however, was to investigate process differences between experts and novices both at a cognitive level, by means of verbal reporting, and at a perceptual level, by means of eye tracking. The results from verbal reporting revealed that although experts used indeed more technical terms, which encapsulate their knowledge (cf. Boshuizen & Schmidt, 1992), they did not use less words than novices, contrary to the assumptions of Ericsson and Simon (1980) and Hypothesis 2. This finding may be due to the fact that novices also verbalize little because they lack knowledge. However, it might also be due to the use of cues for retrospective verbal reports. Cued retrospective reporting seems to be a suitable method to enable experts to inspect their cognitive processes during task performance (Van Gog, Paas, Van Merriënboer, & Witte, 2005).

Moreover, the results showed that experts did not primarily focus on features that were crucial for identifying a locomotion pattern. Instead they mainly concentrated on features that allowed identifying the fish species; these features were not related to a locomotion pattern at all. In line with Boshuizen and Schmidt (1992) this finding indicates that experts use knowledge-based shortcuts, like using the features associated with a fish species, activating the appropriate schema, which also contains the knowledge on the locomotion pattern. Evidence that experts used this strategy can be found at both the conceptual and the perceptual level. For instance, experts often began their verbal reports by mentioning the fish species and they had longer gaze durations on AAOI relevant for species classification compared to novices. These findings confirmed Hypothesis 3.

Furthermore, the similarity of sequences of gaze allocations was investigated at a perceptual level. The results showed that experts had a more diverse gaze pattern compared to novices. This finding is in line with Hypothesis 4 and with the results of Medin et al. (2006) on experts classifying fish species. There are several possible explanations for this finding: It might either be that experts act upon their individual case-based knowledge rather than upon a more generic knowledge base that many experts share. Or it might be that experts use diverse strategies, while novices use no strategies at all and thus their gazes converge around a more neutral pattern that might be, for instance, influenced by salience. Indeed, it has to be noted that the group of experts was rather diverse (i.e., advanced PhD students as well as professors), which also might have led to a diversity in strategies resulting in a high variability of the perceptual patterns.

Finally, the cognitive and perceptual processes were analyzed with regard to whether experts would attend more to relevant rather than irrelevant features compared to novices during the task as indicated by Hypothesis 5. The comparison of initial utterances of the verbal reports between experts and novices showed that experts considered more of the relevant information in the beginning of their task processing. Furthermore, analyzing the distribution of gaze durations on AAOIs revealed that experts also attended more to relevant areas compared to novices. Finally, the verbal reports for the entire duration of task accomplishment showed that the experts' attention remained focused on relevant areas. Thus, the findings of the present study show, in line with results from studies on static stimuli (Antes & Kristjanson, 1991; Charness et al., 2001; Haider & Frensch, 1999), that experts attend to more relevant features of a complex dynamic stimulus than novices. Hence, Hypothesis 5 was confirmed.

3.3 Conclusions of the Normative and the Empirical Task Analysis with Regard to Instructional Goals

This chapter provided an analysis of the task to classify fish locomotion patterns. This task analysis consisted of two parts, namely a normative analysis that investigated how the task should be performed (Section 3.1) and an empirical task analysis that investigated how the task is actually performed (Section 3.2). Both task analyses resulted in similar conclusions that can be drawn for instructional design based on the three important aspects of a task analysis (according to Schaafstal & Schraagen, 1992): the procedure of the task, the knowledge and skills required to perform the task, and the difficulties that occur during performing the task. For an overview of the findings from both task analyses see Table 3.8.

First, the normative analysis of the *task procedure* showed the importance of perceptual skills for the task procedure. The empirical analysis revealed more details on the task procedure. When performing this task, experts are fast, use shortcuts based on encapsulated knowledge. Thus, instruction for novices cannot be designed based directly on experts' task performance. Moreover, this study revealed that experts have diverse approaches to performing the task of classifying fish locomotion patterns. Thus, there is probably not one optimal approach to this task. Rather, for instructional purposes it would be more helpful to carefully choose the task approach of one expert to design instruction based upon it.

Second, the normative analysis of the *knowledge and skills* required to perform this task revealed that the crucial skills in this task are to detect the relevant body parts and to interpret their motion. The empirical task analysis furthermore revealed that experts outperform novices in this task by being faster and more precise in performing the task. Experts have shown to outperform novices by better attending to relevant body parts. Thus, perceptual skills play a crucial role in this task and should be a focus of the instructional design for this task. Moreover, experts have no difficulties in verbalizing how they proceed. Thus, instruction can be designed based on experts' cognitive knowledge and skills.

Finally, the normative analysis of *difficulties* that are likely to occur in this task showed that visually detecting the relevant body parts in the context of many irrelevant elements might be difficult for novices. Moreover, interpreting the motion of these body parts while their motion is complex and fast might be even more difficult for novices. The empirical task analysis revealed additional findings on the difficulties that occur in performing this task. Novices perform slow and inaccurate. They have severe deficiencies in perceptual skills in that they become easily distracted by irrelevant areas, while attending less to relevant areas compared to experts. This was found in both their perceptual and cognitive processes. These findings support the assumptions derived from the normative task analysis that novices need instruction that would foster their perceptual skills.

The next part of this thesis will present how these findings can be used to develop an appropriate instructional method for the fish locomotion classification task.

Table 3.8:
Summary of Conclusions Drawn from Both Normative and Empirical Task Analysis on Instructional Goals to Teach Classifying Fish Locomotion Patterns

Conclusions from...	Consequence for instruction
Normative task analysis	
Procedure: The task is composed of three hierarchical steps: the first two are based on perceptual input, the third is solely based on conceptual knowledge	1. Correctly processing the perceptual input is a prerequisite to successfully perform the task
Knowledge and skills required: The task requires perceptual skills (i.e., detecting the relevant body parts and interpreting their movement) and conceptual knowledge (assigning the observations to a category)	2. Since conceptual knowledge can be trained with traditional methods, a method to convey perceptual skills is required
Difficulties assumed to occur: The perceptual input includes irrelevant and dynamic (i.e., transient) information	3. Novices' visual attention should be guided
Empirical task analysis	
Procedure: Experts use shortcuts in performing a task	1. Experts' strategies in performing the task cannot be directly taught to novices
Knowledge and skills required: Experts outperform novices by looking faster and longer to relevant information	2. Experts possess the perceptual skills for performing this task, which should be taught to novices
Difficulties assumed to occur: Novices get easily distracted by salient, but irrelevant information	3. Novices need training to ignore irrelevant information and attend to relevant information

Part II

Teaching Perceptual Skills

Chapter 4

Development of a Novel Method to Teach Perceptual Skills for the Classification of Fish Locomotion

Part I of this thesis presented an analysis of the task to classify fish locomotion patterns. First, theoretical considerations provided a normative task analysis. Second, an empirical task analysis was provided by comparing experts and novices while they were actually performing the task. The conclusion from both task analyses is that an important instructional goal should be to teach novices perceptual skills. As described in Chapter 3.1.2, the perceptual skills – i.e., skills that cannot be executed without perceptual input – required to classify fish locomotion are necessary for two steps in the task procedure (cf. Chi, 2006a; Manning et al., 2005):

1. Detecting **which** body parts are used to swim: *Visually searching* the relevant body parts (i.e., quickly looking at them) and *identifying* the relevant body parts (i.e., denoting them as being relevant).

2. Detecting **how** these body parts are used: *Visually inspecting* those relevant body parts (i.e., looking at them for a sufficiently long[1] time) and correctly *interpreting* the motion of those relevant body parts (i.e., describing their motion correctly).

[1]Instead of being distracted by salient, but irrelevant areas in relation to the total viewing time of the perceptual input.

Biology education for marine zoologists, however, hardly focuses on conveying these perceptual skills. Current education on classification of fish's locomotion patterns rather addresses the physical principles underlying fish locomotion (Nachtigall, 2002) and conceptual knowledge (Bone & Moore, 2008; Hoar, Randall, & Conte, 1997). In those cases, in which perceptual skills are addressed, instruction often tries to convey these skills in the same manner as conceptual knowledge. For instance, instructional material makes use of either verbal description of the motion ("Filefish slowly swim by a wavelike motion of their second dorsal and anal fin." Lieske & Myers, 1994, p. 340), or complex, static pictures (e.g., Figure 4.1, Nachtigall, 2006).

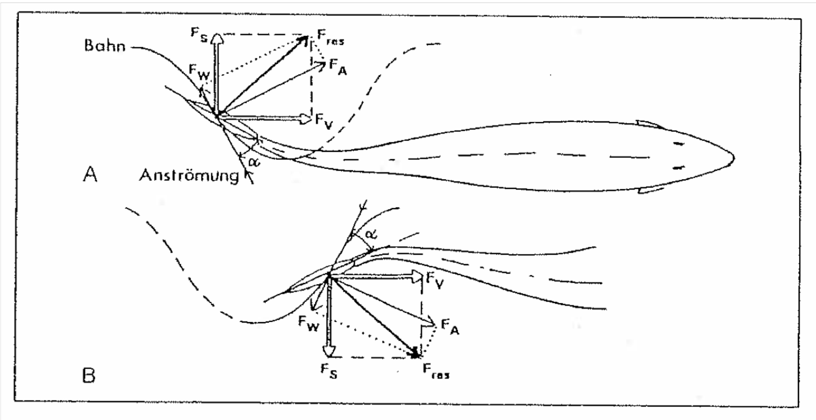


Figure 4.1: Static depiction of an anguiliform locomotion pattern. Adapted from "Ökophysik: Plaudereien über das Leben auf dem Land, im Wasser und in der Luft", by W. Nachtigall, 2006, p. 125.

Consequently, no established instructional method for conveying these skills exists in this area. This fact also holds true for other domains that rely on perceptual input. Although many researchers stress the importance of perceptual skills for different tasks (e.g., in sports: Williams & Ward, 2003, for medical diagnosis: Nodine et al., 1996, in Arts: Antes & Kristjanson, 1991), hardly any elaborated ideas for teaching methods exist. For instance, some researchers recommend to train perceptual skills by simply extensively repeating subtasks that require these skills (in sports e.g., Farrow, Chivers, Hardingham, & Sacuse, 1998). These training, however, aim at improving the *speed* of perception and to automate the connection between perception and physical reaction to it (e.g., reaction to an opponents service in tennis: Farrow & Abernethy, 2002) and not at improving the detailed analysis of a perceptual input.

On the other hand, a lot of research has been conducted on the initial acquisition of *cognitive skills*, like writing argumentative texts or solving mathematical equations, yielding successful instructional methods. The major difference between cognitive and perceptual skills (as defined above) is whether merely cognitive actions play the central role or whether the examination of a perceptual input is in focus.

In the field of acquiring cognitive skills example-based learning has proven to work particularly well. Example-based learning has been studied from two different perspectives: Cognitivistic approaches to learning investigated examples by providing a written, worked-out problem solutions to study (i.e., classical worked examples; see Atkinson, Derry, Renkl, & Wortham, 2000; Sweller et al., 1998). Socio-constructivistic research investigated examples by allowing students to observe an expert performing a task live or on video (i.e., cognitive modeling; see Bandura, 1977; Collins et al., 1989). Both approaches are discussed in the next sections based on the overview provided by Van Gog and Rummel (2010) that contains a detailed discussion on differences and commonalities of both instructional methods (see also Renkl, 2010). With the difference between cognitive and perceptual skills in mind the following section will review instructional methods that teach cognitive skills and discuss their possible application to teaching perceptual skills.

4.1 Instructional Methods to Teach Cognitive Skills

This section discusses instructional methods for the initial acquisition of cognitive skills. In a second step, these methods will be used to inform the instructional design of the novel instructional method for the acquisition of *perceptual skills* (EMME). I will first provide an overview on learning from classical worked examples. Subsequently, as an extension of this approach I will discuss learning from process-oriented worked examples. Finally, I will provide an overview on learning by cognitive modeling.

4.1.1 Learning from Classical Worked Examples

4.1.1.1 Classical Worked Examples and their Areas of Application

Classical worked examples are composed of a problem formulation, its final solution, and a procedure for solving the problem, that is, the solution steps (Van Merriënboer, 1997; Renkl, 2002). All parts are presented as written text, sometimes including graphs (Van Gog & Rummel, 2010). The solution procedure is usually presented in a step-by-step manner (Atkinson et al., 2000). For an example see Table 4.1

All applications of the worked-example method have been confined to conveying *cognitive* skills, like mathematics (e.g., Paas, 1992), physics (e.g., Reisslein, Atkinson, Seeling, & Reisslein, 2006; Van Gog, Paas, & Van Merriënboer, 2006, 2008), applying instructional design models (Hoogveld, Paas, & Jochems, 2005), or reasoning about legal cases (Nievelstein, Van Gog, Van Dijck, & Boshuizen, 2010). For an detailed overview over tasks that classical worked examples have been applied to see Renkl (2010) as well as Van Gog and Rummel (2010). It has not been shown, yet, whether classical worked examples can also be used to teach perceptual skills.

4.1.1.2 The Worked Example Effect and its Explanation

The ”worked example effect” (Sweller et al., 1998) refers to the finding that learning with classical worked examples results in higher learning outcomes, while less time and effort needs to be invested, compared to learning by unassisted problem solving (for reviews, see Atkinson et al., 2000; Renkl, 2010; Sweller et al., 1998; Van Gog & Rummel, 2010). It is important to note that classical worked examples have proven to be particularly suited for the initial phase of skill acquisition (Atkinson et al., 2000). If implemented in later stages of skill acquisition, classical worked examples even hamper learning compared to letting learners solve problems by themselves. This effect is known as the ”expertise reversal” effect (Kalyuga, Ayres, Chandler, & Sweller, 2003; Kalyuga, Chandler, Touvien, & Sweller, 2001).

Table 4.1:
Classical worked example adapted from "From example study to problem solving: Smooth transitions help learning", by A. Renkl, R. K. Atkinson, and U. H. Maier, 2000, as cited by Atkinson, Derry, Renkl, & Wortham (2000).

PROBLEM TEXT: From a ballot box containing 3 red balls and 2 white balls, two balls are randomly drawn. The chosen balls are not put back into the ballot box. What is the probability that a red ball is drawn first and a white ball second?

SOLUTION:

STEP 1:

Total number of balls:	5
Number of red balls:	3
Probability of red ball on first draw:	$3/5$

STEP 2:

Total number of balls after first draw:	4
Number of white balls:	2
Probability of white ball on second draw:	$2/4$

STEP 3:

Probability that a red ball is drawn first and a white ball is second:	$3/5 * 2/4 = 6/20 = 3/10$

ANSWER: *The probability that a red ball is drawn first and a white ball is drawn second is 3/10.*

The worked examples effect is most often explained against the backdrop of the Cognitive Load Theory (Sweller, 1988; Sweller et al., 1998; Van Merriënboer & Sweller, 2005). This theory is based on the assumption that learning takes place in working memory, which has a limited capacity, while the learned information is stored in the unlimited long-term memory. Moreover, this theory is based on the assumption that schema-based problem solving is most efficient and a characteristic of performance on an expertise level. Schemas are knowledge structures stored in long-term memory that can be activated in working memory to efficiently perform a task (Sweller & Cooper, 1985). A schema provides a ”framework” for solving problems of a specific category that needs to be filled with information specific to a given problem. Afterwards, the solution procedure, which is attached to the schema, is activated and the task is performed (Van Lehn, 1989). The Cognitive Load Theory assumes that problem solving and worked example study evoke different cognitive processes. When novices learn by solving conventional problems they use general problem-solving strategies such as means-ends analysis. When applying such strategies learners continuously search for operators – actions that lead from one problem state to the following one – to reduce the difference between the current problem state and the goal (Sweller, 1988). The search processes impose a high load on working memory, but they are not effective for learning. Sweller and Cooper (1985, p. 61) state that ”Problem-solving search and schema acquisition may be unrelated, even mutually incompatible activities. Attention to one may, at least in part, preclude attention to the other”. That means that when a novice performs a task, s/he focuses on searching for a solution by means-end strategies, and thus, has no free cognitive capacity to learn. Even though such strategies may allow learners to succeed in solving the problem eventually, they contribute very little to learning. Classical worked examples, in turn, reduce the use of suboptimal problem-solving strategies, allowing the novice learner instead to devote all the available cognitive capacity to study the worked-out solution procedure so that in the end, the novices are able to perform comparable tasks on their own (Sweller, 1988; Sweller et al., 1998).

Classical worked examples deliver information on solution steps and seldom a rational for why certain steps are chosen. This information, however, may be important to enhance understanding. In line with this reasoning, Van Gog, Paas, and Van Merriënboer (2004) have introduced an extension of classical worked examples by adding process-oriented information. This was meant to further enhance learning (Van Gog et al., 2004). Classical worked examples may be described as a product approach (Paas & Van Merriënboer, 1994). Product approaches focus on the results of each step of an effective task performance. What classical worked examples miss is the information underlying the selection and application of each solution step (Van Gog et al., 2004). Thus, they might not be sufficient to promote understanding, which is a prerequisite to achieve a certain level of expertise (Van Gog et al., 2004). Understanding a procedure involves both knowledge of its domain (knowledge about that domain's objects and events) and of its teleology (knowledge of purpose of the solution steps, Ohlsson & Rees, 1991). In line with that reasoning, instructional design of examples that aims at increasing understanding has shown to be effective (Gerjets, Scheiter, & Catrambone, 2004). Process approaches are meant to increase learners' understanding as well by making novices mimic experts' problem-solving behavior during training (Paas & Van Merriënboer, 1994).

4.1.1.3 Process-Oriented Worked Examples and Tasks They Have Been Applied to

Process-oriented worked examples are worked examples that additionally include process-oriented information. Process-oriented information is knowledge on how and why solution steps are taken and why they are performed in a particular order. Thus, process-oriented worked examples do not only contain the problem formulation, solution steps, and its final solution (all elements of product-oriented worked examples), but also principled information ("why") and experts' strategic information ("how") (Van Gog et al., 2004).

Van Gog (2006) explains the rationale behind process-oriented worked examples as follows. Classical worked examples are meant to prevent unnecessary search processes and thus to free learners' cognitive capacities (see above). The cognitive capacity that is than available may be used for cognitive activities that contribute to learning. Learners', however, seldom use this cognitive capacity spontaneously for cognitive activities that promote learning. Thus, it is important to stimulate learners to do so by effective instructional design. One possibility to stimulate learners is to add process-oriented information to classical worked examples.

Understanding the ”why”and ”how” of solution steps is considered important to recognize and apply the relevant parts of a previously learnt procedure (Catrambone, 1996, 1998; Gott, Parker-Hall, Pokorny, Dibble, & Glaser, 1993). Some learners may be able to self-explain this information (Chi, Bassok, Lewis, Reimann, & Glaser, 1989), which in turn contributes to learning (Atkinson, Renkl, & Merrill, 2003). Learners in the initial stage of skill acquisition, however, may lack the domain knowledge necessary to do so (Chi et al., 1989; Renkl, 1997). By providing process-oriented information to worked examples that includes the rationale behind the steps, learners may use the cognitive capacity that is freed-up by receiving a worked-out solution to study this process-oriented information, which would impose mental effort that contributes to learning (Van Gog et al., 2004).

Since process-oriented worked examples are a rather new instructional method, they have not been implemented often, yet. So far, they have been applied to complex cognitive tasks, like electric circuit troubleshooting (Van Gog et al., 2006), troubleshooting in a chemical processing plant (Darabi, Nelson, & Palanki, 2007), training older adults to use a ticket vending machine (Struve, 2008), and to training higher education teachers to apply instructional design (Hoogveld et al., 2005).

4.1.1.4 Findings on the Use of Process-Oriented Worked Examples

The empirical evidence for the effectiveness of process-oriented worked examples is mixed. Although some of the presented implementations find an advantage of providing process-oriented information to worked examples (Hoogveld et al., 2005; Struve, 2008; Van Gog et al., 2008) others failed to show an advantage of process-oriented worked examples over product-oriented worked-examples (Darabi et al., 2007; Van Gog et al., 2006). Nevertheless, all argue that process-oriented worked examples have the potential to increase learning if implemented in an appropriate manner. Given these studies four issues are of importance in designing process-oriented worked examples, which are of particular importance for the development of the novel instructional method (EMME) in this thesis:

First, a *split-attention effect* – that is mutually referring information sources that cannot be understood in isolation – should be avoided by presenting information in an integrated format (Chandler & Sweller, 1992). Van Gog et al. (2006) assume that process-oriented worked examples are only effective for learning if the form of the process information does not induce mental effort that hinders learning. The authors argue that since they presented part of the process information in the worked examples as written text, this may have caused the split attention effect. The authors assume that if all process information was presented in an auditory format, the effects would have been far better. Similar, the process and product informations in the study by Darabi et al. (2007) were presented in a split-attention format (Chandler & Sweller, 1992). Considering this fact and arguing in line with Van Gog et al. (2006) it is not surprising that the worked examples in this study were not helpful. Struve (2008) in contrast, who presented process-oriented worked examples as video (including audio), and thus without splitting the attention across the representations, could show that adding process-oriented worked examples increases learning to some extent in comparison to classical worked examples only.

Second, presenting the process-oriented information so that it *increases task complexity* should be avoided or otherwise learning might be impaired (cf. overload by so-called intrinsic cognitive load: Sweller, 1988). Hoogveld et al. (2005) discuss several reasons why their implementation of process-oriented worked examples was inferior to their implementation of product-oriented worked examples. Amongst others, they argue that the combination of the complex learning task and the ”why” and ”how” information might have overloaded participants, resulting in a too high task complexity and thus leaving no cognitive capacity for learning to occur. In line with this reasoning, Van Gog et al. (2006) assume that process-oriented worked examples are only effective for learning if the content of the process information does not induce too much mental effort that hinders learning. The authors assume that the content of process information may change the task itself and thus increase the mental effort caused by the task. Thus, the authors conclude that it needs to be addressed in further research whether and how additional support in form of process information added to worked examples can be effective to novice learners.

Third, too much *irrelevant information*, that is ”noise” may hinder participants from learning. Hoogveld et al. (2005) argue that the process-oriented information, which was presented as expert’s verbalization, seemed to include too much ”noise” (i.e., the process information may be embedded in information not directly relevant to the task at hand), which in turn might have overwhelmed learners. In line with that explanation, research on learning by observation on videos has shown that recognizing the relevant aspects of expert behavior requires a minimum level of experience and that inexperienced people run the change of getting lost in irrelevant details (Jentsch, Bowers, & Salas, 2001).

Fourth, an *expertise reversal effect* should be avoided. This effect states that if too skilled people receive instructional assistance, their performance is lower in comparison to unassisted problem solving (Kalyuga et al., 2003, 2001). In the case of process-oriented worked examples this means that persons with a certain expertise level should already be familiar with the process information, that is, presenting this process information to them makes this information *redundant*. This assumption is supported by a study by Van Gog et al. (2008). In that study the authors could show that process-oriented worked examples are helpful only in the initial stage of skill acquisition. Neglecting the expertise reversal effect when conveying knowledge by means of process-oriented worked examples can even hinder learning. For instance, the negative finding of Darabi et al. (2007) may be due to the fact that their participants were not novices, but more and less experienced engineering students. From this point of view their finding is in line with the well-established expertise reversal effect.

4.1.1.5 Applicability for Teaching Perceptual Skills in Classifying Fish Locomotion

As described in Section 3.1.2, the task to classify fish locomotion patterns requires the acquisition of several schemas at different levels. For instance, learning to detect the body parts that are relevant for different locomotion patterns or learning to interpret motion patterns. As classical worked examples have shown to facilitate schema acquisition, they may provide an appropriate instructional approach to this task as well. In addition, process-oriented worked examples present cognitive processes information required to solve the task, to the learner to understand the rationale behind each task step. The schema required for the current task, however, should additionally be composed of perceptual aspects. Furthermore, one aspect of the underlying task is dynamism, as described in Section 3.1.3.2. That means that part of the presented information is transient and thus, is time critical. That is, the task performer has to attend to a certain part of the example at a certain point in time (e.g., a fin while flapping) or s/he might miss important information to solve the task. In general, worked examples, however, do not involve any time critical events. Instead all relevant elements are available to the learner at any point in time. A method that has been successfully applied to teach time critical events (like motor skills), and is thus able to display time critical processes is cognitive modeling. This well-established instructional method will be described in the following section.

4.1.2 Learning from Cognitive Modeling

4.1.2.1 Cognitive Modeling and its Explanation

Bandura introduced learning from modeling in 1977. He could show that observing a model performing a task makes the observer learn to carry out the task her-/himself. Since this time learning from modeling has been implemented broadly: The model may display behavior in a natural manner or behave rather didactically (cf. Van Gog & Rummel, 2010). The learner can observe the model's behavior live, as a video recording, as a screen recording of the model's computer screen, or as an animation (see Bandura, 1977; Collins et al., 1989; Wouters, Paas, & Van Merriënboer, 2008). Thus, in contrast to worked examples, which are written and static representations (apart from single applications of process-oriented worked examples), modeling *always* includes dynamic visual information.

Although modeling has proven to be more effective than learning by unassisted problem solving (e.g., Couzijn, 1999; Kitsantas, Zimmerman, & Cleary, 2000) – as for worked examples – this effect is usually not explained in comparison to other types of learning, but in terms of the general cognitive processes occurring during observational learning (for a detailed description see Van Gog & Rummel, 2010). Bandura (1977, 1986) assumes that learners acquire a cognitive representation of the model's behavior that enables them to exhibit this behavior at later situations. In that, the attentional processes of the learner to the relevant aspects are of great importance.

Modeling aims at teaching overt (i.e., observable) *processes* that experts use to handle tasks (Collins et al., 1989). Collins et al. (1989) have expanded this method to conveying *cognitive* processes, which they called cognitive modeling[2]. It teaches the use of domain knowledge in solving problems and carrying out tasks. In the case of cognitive modeling, however, the crucial actions of the model are hidden cognitive processes and thus not readily available to learners for observation. To learn from these processes the learner must be able to observe them anyhow. Thus, the model has to externalize her/his cognitive processes that are usually carried out internally. This is done via methods from cognitive research: thinking aloud (cf. thinking aloud in Section 2.3.1).

Collins et al. (1989) explain the effect of cognitive modeling by the fact that it helps to integrate the knowledge and the skills to develop expertise. In that, observation plays a key role since it aids learners in developing a conceptual model of the target task or process prior to attempting to execute it. It provides learners with an advanced organizer that allows them to concentrate more of their attention on execution than would otherwise be possible (cf. prevention of unnecessary search in worked examples). Knowledge is thus learned in terms of its use within different contexts, encouraging both a deeper understanding of the meaning of the concepts and facts themselves and a rich web of memorable associations between them and problem-solving contexts (so-called situated learning: Resnick, 1987, cf. process-oriented worked examples).

[2]Cognitive modeling is the first part of the cognitive apprenticeship approach. It is followed by ”coaching”, ”scaffolding” (including ”fading”), ”articulation”, ”reflection”, and ”exploration”. For the scope of this thesis, I am only interested on the first modeling part. In further research, the following parts might also be addressed.

4.1.2.2 Findings on Cognitive Modeling

Like using worked examples, cognitive modeling aims at the acquisition of skills, where it has shown to be an efficient way of fostering learning (Collins et al., 1989). This is particularly true for less structured domains, like writing (e.g., Braaksma, Rijlaarsdam, & Van den Bergh, 2002; Couzijn, 1999), communication/collaboration (e.g., Rummel & Spada, 2005), and motor skills (e.g., Blandin, Lhuisset, & Proteau, 1999). For an overview of tasks to which cognitive modeling has been applied to see Renkl (2010) and Van Gog and Rummel (2010). Note that both, modeling communication/collaboration skills as well as modeling motor skills include time critical events. When observing a model displaying how to collaborate with others efficiently (e.g., Rummel, Spada, & Hauser, 2009) or executing a meaning-free experimental motor skill (e.g., Blandin et al., 1999) it is important for the learner to attend to relevant aspects of the modeled behavior at a certain point in time.

4.1.2.3 Comparison of Worked Examples and Cognitive Modeling

All three approaches – classical worked examples, process-oriented worked examples, and cognitive modeling – have many points in common: In contrast to other approaches that postulate that learning should result from a problem-solving process conducted by the learners themselves (e.g., discovery learning: Bruner (1967), or problem-based learning: Hmelo-Silver, Duncan, and Chinn (2007)), all three here presented approaches are a way to learn from problem solutions provided by others. In all cases, for cognitive tasks the model has to externalize cognitive processes – i.e., demonstrating them by verbalizing or writing them down –, so they are observable for the learner. Furthermore, researchers from both areas assume that it is impossible or even dangerous to let the learner discover all required knowledge by own experience. It is supposed to be more efficient to ”borrow” knowledge from others (Bandura, 1986, 1977; Sweller, 2004; Sweller & Sweller, 2006).

It has to be noted, however, that cognitive modeling differs crucially from worked examples in two aspects: the nature of the modeled behavior and the role of attentional processes of the learner (for more elaborations on that topic see Van Gog & Rummel, 2010).

For a comparison of classical worked examples, process-oriented worked examples, and cognitive modeling in so far it is relevant in the context of this thesis see Table 4.2.

Table 4.2:
Comparison of Classical Worked Examples, Process-oriented Examples, and Cognitive Modeling

	Classical worked examples	**Process-oriented worked examples**	**Cognitive modeling**
Presentation content	Worked-out solution to a problem	Worked-out solution with process information	An expert performing a task directly
Presentation format	Written	Either written or auditive explanations of why s/he is doing certain steps	Live or video recording of a person demonstrating and explaining the performance of a task
Possibility to convey time-critical events	not given	usually not	given and successfully implemented
Nature of the modeled behavior	Didactically optimal	Natural or didactical	Natural or didactical
Nature of attentional processes	All presented steps are relevant for the learner to attend to, little distracting information	May include distracting information	May include a great deal of distracting and transient information

Nature of the Modeled Behavior. A model may demonstrate behavior with the primary intention to efficiently perform a task (i.e., a natural behavior of experts) or with the primary intention to teach the performance of this task to novices (i.e., a didactic version of expert performance). The solution procedure presented in worked examples can be seen as a kind of expert's problem solving model (Atkinson et al., 2000). This model, however, does not present actual natural expert behavior, but rather a didactical version of it (Van Gog & Rummel, 2010). Thus, worked examples usually present a didactical and optimal solution procedure.

Cognitive modeling research has studied the question which expertise level a model should have (e.g., Blandin et al., 1999). The model usually is a teacher or an expert who is behaving didactically (e.g., Schunk, 1981; S. J. Simon & Werner, 1996). However, the model could also be a peer student with only an equal or slightly higher level of performance than the learner, in which case the demonstrated procedure may even contain errors (e.g., Braaksma et al., 2002; Schunk & Hanson, 1985). How to decide on which level of expertise a model should have? On the one hand, experts provide explanations that are abstract and thus, may be difficult for novices to understand. Hence, instructions that are provided by people with an intermediate expertise level may lead to better learning (Hinds, Patterson, & Pfeffer, 2001). On the other hand, for solving novel tasks, abstract explanations provided by experts may be more beneficial with regard to transfer (Hinds et al., 2001). Thus, Schunk (1981) suggests that expert models might lead to better learning than models with less expertise. This was supported by a study by Boekhout, Van Gog, Van De Wiel, Gerards-Last, and Geraets (in press), who showed that physiotherapy students learned better from expert worked examples than from advanced student worked examples. According to these findings an expert competence level seems to be appropriate to serve as a model.

Still, Van Gog and Rummel (2010) argue that experts, who behave in a natural (i.e., non-didactically) manner would not be appropriate as models as their task performance might be highly automatized. This means that they perform fast and skip task steps (Blessing & Anderson, 1996; Kalyuga & Sweller, 2004). This fact in turn might impose difficulties onto learners, in terms of perceiving and following the model's behavior (Van Gog & Rummel, 2010). This was also shown for the task to classify fish locomotion patterns in Study 1 of this thesis in that experts used shortcuts to classify a fish's locomotion based on classifying its species. Furthermore, as automated behavior does not require controlled processing, the expert model may have difficulties in verbalizing her or his behavior (Chi, 2006a; Speelman, 1998). Consequently, an expert model that is instructed to behave more didactical seems to be the appropriate choice. In Section 4.2.3 it will be discussed in greater detail how to instruct an expert model to behave didactical.

Role of Learners' Attentional Processes. Another important difference between worked examples and cognitive modeling stressed by Van Gog and Rummel (2010) are the demands they evoke for attentional processes in terms of the amount of irrelevant information they present and the transience of relevant information.

The solution procedure in worked examples is presented in a didactically manner, i.e., in a way in which students should learn to solve a problem and not necessarily how another person would actually solve it (e.g., experts might tend to skip steps, while people with low expertise might make errors: Blessing & Anderson, 1996; Kalyuga & Sweller, 2004). Hence, all presented steps are relevant and thus, the learner has to attend to all of them to incorporate them in the cognitive schema. Moreover, the solution procedure in classical worked examples is usually presented in a written manner hence, little distracting information is present. In contrast, cognitive modeling displays actual behavior of a model (e.g., Blandin et al., 1999; Rummel et al., 2009), which may result in a large amount of irrelevant and distracting information. Regardless of the presentation format, the learner may easily attend to irrelevant information as the model's tone of voice, salient but irrelevant objects present on the screen, etc. Thus, novices learning by cognitive modeling should not get distracted by irrelevant information.

Moreover, both instruction types differ in the transience of information (i.e., time-critical events) they present. Worked examples present a complete overview of the solution procedure to the learners. Cognitive modeling, on the other hand, either builds up that overview step-by-step (e.g., McLaren, Lim, & Koedinger, 2008), or only one step is present at a time and after it is executed and not present any more the next step occurs (e.g., Wouters, Paas, & Van Merriënboer, 2009). This is, for instance, the case if the model verbalizes her/his thoughts or if the problem statement the model is working on includes dynamic changes (as it is the case for the fish locomotion). In this case, the transience of the information requires learners to maintain each presented step in working memory while attending to the step that is currently being executed and processing them in relation to each other. This can be very cognitively demanding, especially for novice learners, and might hamper learning (e.g., Ayres & Paas, 2007).

Van Gog and Rummel (2010) thus conclude that for learning from cognitive modeling, it is crucial that the learner attends to the relevant aspects of the modeled behavior, and ”selective attention is, therefore, one of the crucial subfunctions in observational learning” (Bandura, 1986, p. 51). The learner’s attention, however, may be easily distracted by irrelevant information or is difficult to maintain across dynamic changes. Thus, cognitive modeling is demanding from an attentional perspective, which in turn may overwhelm learners.

4.1.2.4 Applicability for Teaching Perceptual Skills in Classifying Fish Locomotion

Cognitive modeling presents an expert model that demonstrates and explains the performance on a task, usually in a didactic manner. This method has proven to be effective in numerous studies, hence, it is an established method. Moreover, this method has shown to convey tasks that include time-critical events (e.g., motor skills). The normative task analysis has shown that this is an important issue for the task of interest in this thesis (Section 3.1). Thus, this instructional method might provide interesting hints on how to teach classifying fish locomotion patterns as well. There are at least two issues that need to be considered, however. First, similar to learning from worked examples, cognitive modeling has not been applied to teaching perceptual skills, yet. Second, for the success of this instructional method it is important to support learners in attending to the relevant aspects of the modeled behavior.

Based on the insights gained from this review, implications for designing an instructional method to teach perceptual skills will be derived in the next section.

4.2 Developing an Instructional Method to Teach Perceptual Skills

This section begins with conclusions on how to teach perceptual skills that can be drawn from instructional methods for teaching cognitive skills (Section 4.1) and from the task analysis conducted in Part I of this thesis. Next, I discuss how to present process information within the to-be-developed instructional examples. This questions is of particular importance for perceptual skills. Furthermore, it will be discussed how to instruct experts to behave in a didactical manner in order to be an appropriate model to novices. These considerations will be integrated into the development of a novel method to teach perceptual skills – eye movement modeling examples (EMME). This section ends with a first empirical study using EMME and consequences that have to be drawn from that study.

4.2.1 Requirements for Teaching Perceptual Skills

This section provides a summary of conclusions that can be drawn on designing an instructional method for teaching perceptual skills from classical worked examples (cf. Section 4.1.1), process-oriented worked examples (cf. Section 4.1.1.3), cognitive modeling (cf. Section 4.1.2), and the task analysis conducted in Part 1 of this thesis. Table 4.3 provides an overview of the conclusions drawn in the last section. Conclusions drawn from the task analysis are summarized in Table 3.8.

To teach the task of classifying fish locomotion patterns the instructional method should convey *perceptual skills* that enable to visually detect the relevant body parts and interpret their movement (→ task analysis). Three instructional methods that have proven to be effective for early skill acquisition, that is, to convey skills to novices, were presented in the last section: classical worked examples, process-oriented worked examples, and cognitive modeling. Furthermore, these methods foster the acquisition of schemas, they provide process information, and at least cognitive modeling can provide time-critical information.

Table 4.3:
Consequences Drawn from Teaching Cognitive Skills for the Design of an Instructional Method to Teach Perceptual Skills

Cognitivistic	**Classical Worked examples**
	1. presentation of examples of swimming fish 2. presenting fish, naming its locomotion, and naming which body part moved how 3. appropriate for participants naive towards the task (i.e., novices)
	Process-oriented worked examples
	1. verbal explanation of why and how solution steps are chosen 2. presenting process information verbally and integrated within the example 3. using presentation format that reduces amount of information 4. guiding learners' (visual) attention 5. using presentation form of process information that is meant to guide learners' attention; capturing how well learners' attention was guided
Socio-contructivistic	**Cognitive modeling**
	1. videotaping the model while s/he is classifying the fish's locomotion and replaying this recording to the learner 2. recording model's processes while s/he is classifying the fish's locomotion and replaying this recording to the learner 3. expert model, who is instructed to behave in a didactical manner 4. guiding learners' (visual) attention

Accordingly, a method to convey perceptual skills should use *examples* for instructional purposes (→ worked examples). These may be implemented by presenting videos of swimming fish and naming their respective locomotion pattern. In this case, the video of the fish represents the problem statement and the technical term of the locomotion pattern is the final solution. As worked examples also present *solution steps* on the way to the final solution the instructional method should present all hierarchical three steps (What is moving? How is it moving? What is the technical term?) within the example.

Both, process-oriented worked examples and cognitive modeling recommend to additionally present the *processes* underlying these solution steps. As described in both task analyses the crucial processes in this task are two-fold, namely cognitive as well as perceptual. Thus, both types of processes should be included in the example. *Cognitive processes* might be presented as a verbal overlay on the video examples (→ process-oriented worked examples). It is unclear, however, how to present the *perceptual processes*. The only hints provided are to use a kind of recording of the expert model's processes (→ cognitive modeling) that does not overwhelm the cognitive capacities of the learners (→ process-oriented worked examples). The issue on how to present the perceptual processes will be the topic of Section 4.2.2. It is clear, however, that the function that needs to be met is to guide learners' (visual) attention (→ classical worked examples, process-oriented worked examples, theoretical task analysis, empirical task analysis) towards the relevant body parts of the fish.

Another important issue is whose processes to record for modeling purposes. In line with the empirical task analysis and cognitive modeling research an expert should be chosen as the *model*. According to the empirical task analysis processes of only one single expert should be presented, because experts use heterogenous task approaches and an average approach is very likely to be meaningless. The empirical task analysis, process-oriented worked example research and cognitive modeling research allow for the conclusion that the expert model has to be instructed to behave in a didactical manner. How to design such an instruction, however, is not clear yet and will thus be the topic of Section 4.2.3.

4.2.2 How to Represent Cognitive and Perceptual Processes?

As already described cognitive and perceptual processes need to be included into examples used for instruction of perceptual skills.

Usually in worked examples all information on the solution procedure is presented in a written manner. This type of presentation is very likely to interfere with the problem formulation in this task – the video of the swimming fish – in terms of splitting the learner's attention between both representations (for negative effects of split attention on learning see e.g., Chandler & Sweller, 1992). The same may be true for the cognitive process information. Some researchers implemented process-oriented information by making an expert in the respective domain, who is providing the problem's solution, verbalize her/his cognitive processes and explain why s/he is choosing certain steps while performing the task (e.g., Wouters et al., 2008). Other researchers implemented this information as a written text (e.g., Darabi et al., 2007). According to research on learning with multimedia it is preferable to present text in an audio format when it accompanies videos (or animations) instead of presenting it in an written format (so-called ”modality-effect”, cf. Low & Sweller, 2005; Mayer, 2005a; Sweller et al., 1998). Hence, cognitive processes should be presented by asking the expert model to verbalize the solution steps including the rationale behind them. These verbalizations can then be replayed as a spoken audio overlay on an example video so that split-attention is avoided (e.g., Darabi et al., 2007; Van Gog et al., 2006).

Information on processes presented in an audio format include a flow in time. Van Gog et al. (2004) stress that in examples that include such information learners might not attend to the relevant information at the right time. On the other hand, it is crucial that learners attend to the relevant information in order to learn from the model (Bandura, 1977). However, adding cognitive process information does not guarantee that learners' attention is automatically guided to the relevant aspects of the examples. Hence, information on where to attend to, that is, information on perceptual processes should be included.

Perceptual processes might be, for instance, *verbalized* as well. There are at least three points, however, that speak against this presentation format. First, this presentation format will already be used for presenting cognitive processes. Presenting both processes via the same format might lead to interference. Second, experts' knowledge can be considered highly automated. This type of knowledge, therefore, is not easily accessible to conscious inspection. Thus, experts have been shown to have difficulties to verbalize their knowledge within their own domain (Chi, 2006a; Speelman, 1998). Third, verbalizations about perceptual processes refer to contents of working memory that initially exist in a non-verbal code (Ericsson & Simon, 1984). These contents need to be converted in a verbal code before they can be verbalized. Due to this procedure verbalizations of perceptual processes might be very difficult and incomplete (Ericsson & Simon, 1984). In sum, presenting perceptual processes by means of verbalizations seems not be an appropriate approach.

Another possibility to present information on perceptual processes would be to express this information via visual *cues*, where information that is being processed is highlighted in the order in which it is processed (also called signaling: Mautone & Mayer, 2001; B. J. F. Meyer, 1975). Although many studies exist on attention guidance via cueing in multimedia learning, these studies often failed to find positive effects on learning (for a review: De Koning, Tabbers, Rikers, & Paas, 2009). One reason might be that cues used in multimedia materials are usually based on what area a designer or domain expert *thinks* a learner should be looking at a particular moment. Experts, however, may not always know what the most important information for novices is, since they often have difficulties in correctly accessing the knowledge state of novices (Hinds, 1999; Nückles, Winter, Wittwer, Herbert, & Hübner, 2006; Wittwer, Nückles, & Renkl, 2010). Thus, simply asking experts about which information should be cued might not be the appropriate proceeding.

One cueing technique that did prove very successful is the one used by Grant and Spivey (2003), who developed a cue based on a comparison of eye movements of individuals who were and were not successful in solving an insight problem (Duncker's radiation problem). Successful problem solvers attended longer to a crucial area than unsuccessful problem solvers. Hence, in a follow-up experiment, individuals received the same problem, either without any cue, with this crucial area highlighted, or with a irrelevant area highlighted. Participants in the condition with the crucial area highlighted solved the problem significantly more often than participants in the other two conditions. Thus, Grant and Spivey showed that guiding attention based on successful problem solvers' eye movements could lead to enhanced performance. However, the problem used by Grant and Spivey (2003) consisted of a static visualization. Cues that are effective in static visualizations (e.g., arrows, color, highlighting), may not necessarily be effective in dynamic visualizations, because they may not be fully in synchronization with the movement that is displayed (Boucheix & Lowe, 2010). In this case, however, one might use the eye movements of successful individuals directly to guide other people's attention. For example, Velichkovsky (1995) showed that if an expert and a novice solve a puzzle together and the expert is not allowed to act, displaying the eye movements of the expert enhanced the novice's problem-solving performance. An interesting and open question, however, is whether attention guidance based on eye movements can not just influence performance of a specific task at hand, but also enhance *learning*. Learning refers to the resilient change in a person's knowledge about a task that enables him or her to independently perform that exact task or a similar but slightly novel tasks after practice (H. A. Simon, 1983).

In sum, it can be argued that displaying perceptual processes based on actual data of successful performers might be an appropriate solution to include perceptual processes into examples for instruction of perceptual skills. Thus, the expert model's *eye movements* will be presented to learners to provide information on perceptual processes in Study 2. next, it is important to consider how exactly the display of these eye movements of the expert model should be designed in modeling examples.

4.2.3 Instructing the Expert Model to Behave Didactical

As already described, the expert model should display a didactical behavior when presenting an appropriate solution procedure. To provide an *appropriate* solution procedure, the model has to be a domain expert. However, experts may face several challenges when asked to present this solution procedure in a didactical way. As shown in Study 1 of this thesis, experts often act in an automated manner and thereby do not always have conscious access to their processing (see also Chi, 2006b). Vast knowledge and experience in a domain enable an expert to skip parts of a solution procedure and use shortcuts. This procedure might neither be possible for novices to perform, nor useful to learn from. Moreover, due to knowledge encapsulation experts use technical terms as shown in Study 1 (see also Reimann & Rapp, 2008). Knowledge encapsulations are memory shortcuts that represent several solution steps as one unit. This is only possible for experts who have performed a task a vast number of times. In that case, experts can simply retrieve a technical term from memory and immediately have available the shortcut to the solution. This type of retrieval is obviously not possible for novices. Moreover, experts hardly can estimate how a novice would behave (Hinds, 1999; Nückles et al., 2006; Wittwer et al., 2010). To overcome those challenges, the following issues need to be considered.

The model should be an expert both in terms of the domain and in terms of *teaching*. Thus, the model would know from teaching experience where problems for novices occur. Even if the expert is a teaching expert, however, she or he should be to *be instructed* to behave even more didactical. For instructing experts to behave more didactical two research areas are of interest: research on gaze-voice span (Richardson & Dale, 2005; Griffin & Bock, 2000) and on expert-novice communication (Jucks, Schulte-Löbbert, & Bromme, 2007).

Findings from gaze-voice span research provide hints on how to make an expert behave in a didactical manner from a *perceptual* point of view. Richardson and Dale (2005) investigated the relationship between eye movements, speech production, and speech comprehension. To do so, the authors asked one participant ("the speaker") to verbally describe a stimulus while looking at it. They found that if a speaker is asked to describe a stimulus that s/he watched already beforehand, s/he will fixate all the elements in exactly that order in which they are mentioned, briefly before naming them. Such a tight gaze-speech coupling should also be achieved for the instructional method to be developed. If both gaze and speech refer to the same element at almost the same time, this should facilitate the integration of both processes information sources by the novice learners. In order to achieve this tight gaze-voice coupling the expert model was first acquainted to the stimulus and only after this phase s/he re-watched and explained it.

To make an expert behave in a didactical manner from a *cognitive* point of view findings from expert-novice communication research are of interest. Jucks et al. (2007) showed how to improve experts' written knowledge communication to novices. This was done by shifting the experts' focus from the content to the recipient, so that those experts produced texts that would be more meaningful to novices. The main task for the experts was to respond to a written inquiry from a fictitious layperson. First, the knowledge base of the novice recipient was clarified. Second, the expert wrote the text. Third, the focus of the expert was manipulated by asking her or him to confirm or deny a number of statements. Finally, the expert could revise the text. By doing so, the authors manipulated the focus of the expert, namely either content or recipient focus. Due to this manipulation either the recipient of the text or the content of the text was in focus (e.g., "It is important *that the reader knows* what the term XYZ means." vs. "It is important *to know* what the term XYZ means.").

Both approaches to help experts, who are meant to serve as models for novices, to behave in a didactic manner will be used in the EMME approach (for the concrete implementation see Section 5.1.2.2).

4.2.4 Eye Movement Modeling Examples (EMME)

This section presents an instructional method that together with colleagues I have developed to convey perceptual skills: eye movement modeling examples (EMME; Van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009).

As described in Section 4.2.1, this methods should consist of examples. As the task of classifying fish locomotion includes time critical events, these examples should be implemented as video recordings of the swimming fish. Besides the problem statement (the to-be-identified locomotion pattern) and the solution (technical term of the locomotion pattern), these examples should also include all steps of the solution procedure worked out. This should be done in that an expert model explicitly states each step of the solution procedure. Moreover, the processes underlying each solution step should be included. These processes are of both, perceptual as well as cognitive nature. The cognitive processes should be provided by asking the expert model to verbalize them and to replay these verbalizations to the learner in an audio format. As described in Section 4.2.2 perceptual processes can be presented by replaying eye movement recordings of the expert model.

More concrete guidelines on how EMME should be composed can be derived from the above presented instructional methods. The process information *must not be redundant*. Thus, process information is only meaningful for initial skill acquisition otherwise it hinders learning. Classifying locomotion patterns of reef fish has shown to be a difficult task for novices in Study 1. Thus, presenting process-oriented information to them for instructional purposes should not be redundant. Moreover, all three methods have shown that learning from skilled others – i.e., experts – seems to be an efficient way of learning, but this *expert model should behave in a didactical manner*. The processes of the expert may contain irrelevant information ("noise"), which may overwhelm learners. Hence, the expert should be instructed to produce as little noise in her/his process-information as possible. Furthermore, the process-oriented information must not be too complex, otherwise the process information will overwhelm learners leaving no space for learning to occur. For teaching perceptual skills to classify fish locomotion patterns this means that the expert should provide process-oriented information that is understandable to a novice. Hence, the expert will need to be instructed to do so.

In sum, for the construction of EMME, an expert model is asked to perform a task, while her / his eye movements are recorded (perceptual process information). Meanwhile the expert model verbalizes the solution steps including the explanation of why and how s/he chooses these steps (cognitive process information). Next, these recordings are replayed to the learner to study (see Figure 4.2). EMME not only merge the two approaches of cognitivistic (classic worked examples and process-oriented worked examples) and socio-constructivistic research (cognitive modeling), but also takes findings from research on attention guidance in problem solving into account (Grant & Spivey, 2003). Table 4.4 gives an overview of the EMME method.

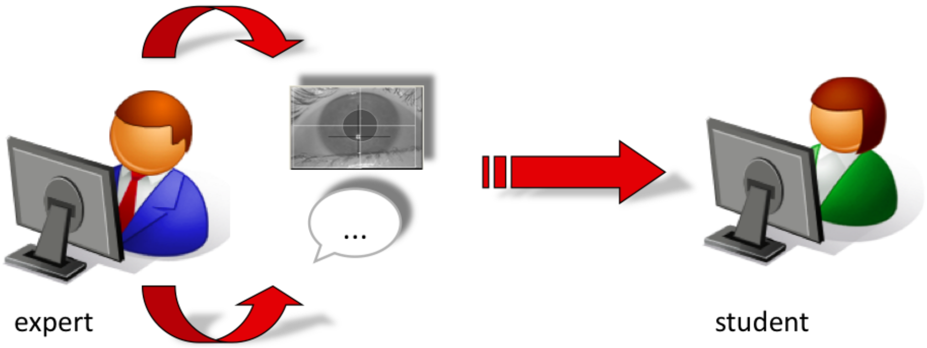


Figure 4.2: Recording and replaying of eye movement modeling examples.

Table 4.4:
Characteristics of Eye Movement Modeling Examples

Theoretical foundation	Cognitivistic research on instructional design, i.e., Cognitive load theory (Sweller, 1988; Sweller et al., 1998; Van Merriënboer & Sweller, 2005), socio-constructivistic research, i.e., social learning theory (Bandura, 1977, 1986), and research on attention guidance in problem solving (Grant & Spivey, 2003)
Definition	Video-based examples that include an expert model's verbalizations of solution steps and their rationale, and a visualization of his/hers attentional focus on the video
Process information	Perceptual: eye movements; cognitive: verbalizations
Instructional goal	Teaching skills that can only be executed in reference to a perceptual input (i.e., perceptual skills)
Target group	Novice learners
Target task	Perceptual task
Nature of the modeled behavior	Didactic
Constraints on learners' attentional processes	Instructional method may include a great deal of distracting information, there is much opportunity for learners to attend to irrelevant details (e.g., tone of voice, salient but irrelevant objects present in the environment or on the screen)

4.2.5 Constraints for the Implementation of EMME

The thorough analysis of the task and its requirements in the first part of this thesis revealed that the task of classifying fish locomotion patterns requires covert processes in form of perceptual skills. The current chapter argues that the EMME method might be suitable for conveying these particular skills. However, it can be assumed that EMME are useful only when the constraints introduced so far (with regards to the task and its requirements and with regard to the design of EMME) are satisfied. If the task or its requirements do not fit very well to the EMME approach and if the design of EMME does not take the considerations outlined in this chapter into account, this instructional method may even hamper learning as has been shown in a study by Van Gog et al. (2009).

In this study participants played a computer based leapfrog game, a structural analogue of the Tower-of-Hanoi game. Participants had to move six frogs, of which three sat to the right and three to the left side of the screen (Figure 4.3). The frogs on each side were looking in the others direction. All frogs had to switch sides by being only able to move in the direction they were looking at and moving only over one other frog at a time. First, participants watched an EMME where a successful player (i.e., an expert model) played the game once. Afterwards, they played it on their own. The authors initially assumed that showing perceptual processes of the model would reveal all frog moves considered by the model for each step of the task. In combination with the verbal explanations of the model, the students would thus acquire knowledge on which choices to make and why. This knowledge, in turn, was meant to lead to better learning outcomes. Results, however, showed that the combination of displaying the expert model's eye movements (i.e., perceptual processes) and verbalizations (i.e., cognitive processes) had detrimental effects on learning compared to either displaying perceptual processes only, cognitive processes only, or behavior modeling only.

In retrospect, it can be said that the task and its requirements did not fit the EMME approach, which in turn emphasizes the need to conduct a task analysis before implementing an instructional method. Moreover, the EMME design was not in line with the considerations outlined in this chapter.

Figure 4.3: A screenshot from the frog leap game as used in the study by Van Gog et al. (2009). The yellow dot between the frogs represents the perceptual processes of the model.

4.2.5.1 The Task and its Requirements

The task investigated in the study by Van Gog et al. (2009) did not contain complex visual information. That is, the representation was only modestly realistic, thus not many distracting irrelevant elements were present. Furthermore, the dynamism in this task was simple. Always only one element (i.e., frog) moved at a time (changed its position, i.e., translation). The direction (horizontal) and goal (empty stone) of the motion was predefined. No element disappeared (i.e., transitions) or changed form (i.e., transformation) during motion (Lowe, 2003, 2004). Thus, the difficulty of this task might not lie in the perceptual skills. The task required to classify fish locomotion patterns, in contrast, includes a highly realistic and dynamic perceptual input. Moreover, both the normative and the empirical task analysis revealed that this tasks causes severe difficulties on a perceptual level for novices.

Moreover, the perceptual processes might have been redundant for this task. Not the motion itself, but the order of the moved objects (i.e., frogs) had to be learned in this task. The definite order of the to be moved objects, however, could have been already inferred from *overt actions* as this was visible on the screen. The considered alternatives on the other hand could be inferred from the verbal explanations alone. Thus, the cognitive and the perceptual processes were redundant. Presenting redundant information is known to hamper learning from examples (Sweller, 2005). This interpretation is supported by the finding that presenting only verbal explanations was helpful for learning.

In contrast, the task of classifying fish locomotion consists only of *covert processes*. Thus, information on how to perform the task can only be delivered by uncovering covert processes. Unlike the frog-task, it cannot be assumed that presenting cognitive processes as spoken audio would be sufficient to learn this classification task. In this task the motion itself is of interest and is likely to be difficult to verbalize (see above). Moreover, the terms the expert uses may not suffice for a novice learner to know exactly what the expert is referring to. For example, students would need to know where a fish's dorsal, pectoral, or anal fins are located in order for the verbal explanations to suffice to guide their attention to these locations. Hence, the perceptual processes should not be redundant, but deliver important information to the learner.

4.2.5.2 EMME Design

In the study by Van Gog et al. (2009) the expert model was asked to behave ”didactically” (p. 3). However, research on expert-communication has shown that experts are not deliberately able to do so without a careful manipulation (cf. Jucks et al., 2007). Since Van Gog et al. (2009) did not use such a careful manipulation, it might have caused the negative effect of EMME on learning. The need to instruct the expert model to behave in a didactic manner was supported by the empirical task analysis for the task of classifying fish locomotion (the concrete procedure is described in Section 4.2.3). Thus, this instruction was improved in the studies reported in this thesis.

On the other hand, the way in which the expert model's eye movements (i.e., a representation of perceptual process information) were displayed in the study by Van Gog et al. (2009) may have been suboptimal. The eye tracker manufacturer's replay options were used, which showed fixations as solid colored dots that increased or decreased in size depending on fixation duration. A solid colored dot that enlarges or contracts might have the undesirable effect that the display itself either obscures relevant task features (in case of very long fixation durations) or attracts attention to itself, drawing attention temporarily away from the task feature the fixation was targeted at. Unfortunately, Van Gog et al. (2009) did not capture the learners' eye movements during example study, so it is unclear whether learners even followed the model's gaze. Moreover, this type of display *adds* information to the example, which might have overwhelmed the learners.

Thus, for the studies reported in this thesis, subtler cues were developed, which might be more effective. One option is a spotlight cue, in which the focused part of the video remains unaltered, whereas all other parts appear blurred. In line with this idea, research in computer science has recently shown that the contrast distribution on a stimulus predicts eye movements on it (Dorr, Vig, Gegenfurtner, Martinetz, & Barth, 2008). Complex stimuli, like realistic videos, can be manipulated in their spectral energy (i.e., contrast in space and over time). Spatio-temporal contrast determines the visibility of edges and motion, so that areas with lowered spectral energy show less pronounced edges and appear ”blurred”. By keeping the spectral energy on fixated areas unchanged, but lowering spectral energy on all the other parts (gaze-contigent manipulation of natural videos: Dorr et al., 2008; offline foveated video compression: Nyström & Holmqvist, 2008), this type of eye movement display may attract attention to the fixated areas without providing additional potentially distracting information or occluding relevant areas. Moreover, it might *reduce the amount of visual information* given in the video making it less visually complex.

As the negative influence of the type of display of the eye movements in the study by Van Gog et al. (2009) is only an assumption, Study 2 of this thesis compares two different eye movement display formats: a manufacturer-provided dot display (cf. Van Gog et al., 2009) and a spotlight display (Dorr, Jarodzka, & Barth, 2010)[3] to a control group without attentional guidance. Hence, in the following chapter it will be investigated whether EMME is a suitable method to teach perceptual skills in the task to classify fish locomotion patterns and which presentation format of perceptual processes is supposed to foster learning.

[3]The technical development of the spotlight display is beyond the scope of this thesis and content of a thesis in Computer Science.

Chapter 5

Teaching Perceptual Skills for Classification of Fish Locomotion via Eye Movement Modeling Examples: Study 2

Study 2 investigated how – and whether – different display types of eye movement modeling examples lead to differences in attention guidance and learning outcomes. These differences were measured by means of eye tracking and multiple-choice questionnaires. In order to gain insights into how to teach perceptual skills in terms of visually searching for relevant body parts and of interpreting the motion of these body parts. Furthermore, this study aimed at investigating how to guide attention of students successfully.

Research Questions - Hypotheses

This study investigated the effect of modeling examples that displayed cognitive and perceptual processes of the model to the learner. The perceptual processes were either displayed as a solid yellow dot (dot display) or by blurring non-attended areas (spotlight display). These conditions were compared to a condition without displaying perceptual processes (control group). Both experimental groups received a display of the perceptual processes of the model and, thus, had eye movement modeling examples at their disposal.

In this study, I hypothesized (Hypothesis 1) that students' visual attention – as measured by eye tracking – during example study would be guided by the eye movements of the expert model (e.g., Böhme, Dorr, Martinetz, & Barth, 2006; Dorr et al., 2008). That is, participants learning with both types of eye movement modeling examples would follow the expert model's eye movements, resulting in a scanpath that is highly similar to the expert model's scanpath, whereas the scanpaths from participants in the control group would be more diverse. Furthermore, if the assumption that the dot display in the study of Van Gog et al. (2009) had been suboptimal is correct, the difference between the students' scanpath and those of the model should be smaller in the spotlight display condition than in the dot display condition.

Learning was measured both in terms of the efficiency of visual search during a test, and in terms of the interpretation of the motion pattern. It was hypothesized (Hypothesis 2) that the eye movement modeling example groups would better learn what to attend to, thereby yielding more efficient visual search than the control group for the novel test videos, as indicated by attending faster and longer to relevant areas, while ignoring irrelevant areas. Since, the spotlight display guides the attention exactly on the relevant spots, instead of obscuring them slightly as the dot display does, visual search should be more efficient for the spotlight condition.

Moreover, I hypothesized (Hypothesis 3) that attention guidance during example study would result also in an enhanced interpretation performance with regard to the motion patterns compared to the control group. However, it is unclear, which experimental condition would be in favor with regard to this measure. On the one hand, as the dot in the dot display condition is solid, it obscures a part of the element the model looked at. Obscuring the relevant information during its verbal explanation might lead to problems for the learners in integrating both information sources. This, in turn, would imply that the spotlight display group would be expected to learn better.

On the other hand, the dot display may also lead to better interpretation performance than the spotlight display, because it acts as a cue. Cues draw attention to the relevant *elements* of the visualization. Attention guidance based on eye movements as used here does not highlight elements but gaze points, and thus, the student has to infer the relevant element behind the gaze point from the contextual information. In the dot display condition this is possible, because the entire visualization can be seen all the time, except for the small area that the model attended to and that is hence covered by the dot. In contrast, in the spotlight display condition, the areas that the expert model did not attend to are obscured. This however, might inhibit the learners to determine the entire object behind the display. In this case, the dot display group would be in favor.

5.1 Method

5.1.1 Participants and Design

Participants were 75 students of the University of Tuebingen. All participants had no prior knowledge of the task and had normal or corrected-to-normal vision. They were randomly assigned to the three conditions (*n* = 25 each): (1) control condition without attention guidance, (2) attention guidance by the original manufacturer-provided "gaze replay" of the model's eye movements (dot display), (3) attention guidance by blurring non-attended areas and leaving fixated areas displayed at a high resolution (spotlight display). Three participants from the spotlight display condition had to be excluded due to poor eye tracking data quality, leaving 72 participants (50 women, 22 men, $M_{age} = 22.83$ years, $SD = 4.04$). Participants were compensated €10 for their participation.

5.1.2 Apparatus and Materials

5.1.2.1 Eye Tracking Equipment

Participants' eye movements while studying the examples and while watching the test videos were recorded with a Tobii 1750 remote eye tracking system with a temporal resolution of 50 Hz on a 17" monitor using Clearview 2.7.0 software (www.tobii.com). This equipment was also used to create the modeling examples. The eye tracking data were analyzed with Clearview 2.7.0 software and self-programmed Java tools.

5.1.2.2 Modeling Examples

The modeling examples for the control group consisted of four digital videos in an audio interleave format (.avi), sized 720 $\times$ 576 pixels (corresponding to 9.6 $\times$ 7.68 inches). Each video depicted a single fish swimming, whereby each fish swam according to a different locomotion pattern. All videos included a spoken explanation of the locomotion pattern by the expert model. The duration of the videos was between 43 s and 114 s depending on the duration of the expert model's spoken explanation. The expert was a professor of marine zoology, with extensive experience in classifying and teaching to classify fish locomotion. Rather than using the expert's natural performance of these tasks as modeling examples it was decided to instruct the expert to behave didactically. The procedure for recording the expert model was as follows. First, according to Jucks et al. (2007) the knowledge base of the learner was clarified by showing the instruction for the student to the expert. Then, the following instruction was given: "A non-biology student asks you, what the relevant aspects of this locomotion pattern are.". Second, to achieve a small gaze-voice span according to Richardson and Dale (2005) the expert first viewed the video, then viewed it again and explained it. Third, the expert model reviewed and evaluated the recording. In that, he received the following statements he had to confirm or deny to shift his focus away from the content to the learner (according to Jucks et al., 2007; for details see Section 4.2.3)

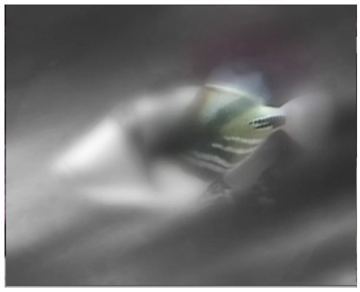
- It is important that the student knows what each term means.
- For this student, the locomotion pattern is explained in comprehensible enough terms.
- For this student, the locomotion pattern is explained in enough detail.
- All information that the student needs are contained.
- All contained information are important to the student.

The last two statements were added to ensure that all relevant and no irrelevant information was given. After reporting on all statements, the expert had the opportunity to revise the recordings. This procedure was repeated for all four learning videos. This recording procedure was chosen, because the Study 1 had shown that experts might use knowledge-based shortcuts due to automated processes (e.g., by recognizing the fish species very quickly). Additionally, they might use many technical terms that a novice student would not understand (Chapter 3.2).

In the dot display condition, participants received the same examples as the control group but those additionally included the expert's eye movements recorded during the final explanation of the video. These were created using the manufacturer-provided "gaze replay" (Tobii). The expert model's gaze points were filtered with the manufacturer implemented dispersion-based fixation detection algorithm (for an overview of disperion-based algorithms cf. Salvucci & Goldberg, 2000). The fixation definition was set at a minimum of 30 pixels of spatial range and a minimum of 100 ms duration. The fixations were displayed as a solid yellow dot (without gazetrail), which changed size (smaller, larger) according to gaze duration (shorter, longer).

In the spotlight display condition, potentially distracting features in the unattended areas were filtered out. That is, the focus of the expert's attention (with a radius of 32 pixels) was visible in an unaltered way, whereas the areas surrounding it were "blurred" by reducing local spatio-temporal spectral energy (i.e., contrast in space and over time) and removing color saturation from non-attended areas in a similar fashion (for a detailed description of the underlying procedure see Dorr et al., 2010). Table 5.1 shows a screen shot from each of the three conditions.

Table 5.1:
Screenshots From the Three Conditions Used in Study 2

Control	Dot display	Spotlight display
		

5.1.2.3 Tests

Participants' eye movements were captured during learning to estimate, whether participants were able to follow the attention guidance by calculating the similarity of their eye movements to the displayed eye movements of the model on the learning video. Moreover, two different measures of learning were obtained: the efficiency of visual search of relevant body parts and the interpretation of their movement during the inspection of new test videos. Visual search data were obtained while participants were shown four test videos of different fish (one video for each locomotion pattern). While watching the videos, participants' eye movements were recorded. From those recordings two different eye movement measures were obtained. First, time until participants first looked at relevant areas on the testing videos[1]. Second, time spent on the relevant areas on the testing videos. The interpretation of the motion patterns was measured by using multiple-choice questions. Participants received 22 additional videos (their eye movements were not recorded during this task). After watching each video once, participants had to answer two multiple-choice questions: (1) indicate from a list of six body parts, which was crucial for thrust propulsion (more than one option could be chosen) and (2) indicate how those body parts were moving (one of two options could be chosen). The questionnaire was presented via a html page.

[1]The relevant areas determined a priori by two domain experts (cf. Antes & Kristjanson, 1991; Charness et al., 2001).

5.1.3 Procedure

The experiment was run in individual sessions of approximately 45 min. At the beginning, participants filled in a questionnaire on their prior knowledge in this task and their demographic data. Then, they received a short introduction to the topic, stating very general information on fish anatomy and that fish may swim according to different locomotion patterns. Then, the learning phase started. The eye tracking system was adjusted to the individual features of the participant based on a nine-point calibration. Participants were told that they would subsequently receive four examples in which a biology expert explains each to-be-learned locomotion pattern. Participants in the dot display and spotlight display conditions were additionally informed that they would see what the expert was attending to, and an example of that was shown with an off-topic video. While studying the examples, participants' eye movements were recorded.

Then, participants completed the two tests. In the visual search testing phase the eye tracking system was recalibrated. Afterwards, a fixation cross appeared for two seconds followed by the testing video, which was replayed once for 4 s. Participants watched the video while their eye movements were recorded. Next, the video disappeared, resulting in a blank screen. This procedure was repeated for the remaining three videos. In the interpretation testing phase participants watched 22 new videos for 12 s each. After each video participants had to answer the multiple-choice questions. In this phase, participants eye movements were not tracked any more.

5.1.4 Data Analysis

5.1.4.1 Analysis of Eye Movement Data During Example Study

Similarity of participants' eye movements with the displayed eye movements of the expert model in the examples was analyzed by calculating the Euclidean distance between simultaneous gaze points of the expert and the learner over time (cf. Rao, Zielinsky, Hayhoe, & Ballard, 2002). Subsequently, the mean Euclidean distance was calculated for each participant.

5.1.4.2 Analysis of Learning Outcomes

To measure learning outcomes in terms of visual search, two eye movement measures were derived: (1) the time until participants first looked at the relevant areas for at least 100 ms, and (2) the total time spent on the relevant areas (total dwell time). These data were obtained by creating ”dynamic” areas of interest (AOI) for the four videos (by means of a very time-consuming manual method described in Section 3.2.2.4). The relevant AOIs were determined a priori by two domain experts in the field of study.

To measure learning outcomes in terms of interpretation, the multiple-choice answers were scored according to the categories necessary for describing a locomotion pattern (Lindsey, 1978): (1) correctly stating which part of the body is used to produce propulsion, (2) correctly stating how this part moves (undulating, i.e., wavelike or oscillating, i.e., paddlelike). This resulted in a maximum score of 1 point per video for both, the first category and the second category. However, points for the second category were only assigned if the first category had been described correctly. Hence, participants could receive a maximum of 2 points per video (1 point for each category).

5.2 Results

For all statistical tests reported here, a significance level of .05 is used. Means and standard deviations for each condition are given in Appendix Study 2, Table 3.

5.2.1 Eye Movements During Example Study

Results are summarized in Figure 5.1 (the exact values are given in Appendix Study 2, Table 3). An ANOVA was conducted with mean Euclidean distance as dependent variable and condition (control vs. dot display vs. spotlight display) as independent variable. The comparison of participants' eye movement data during example study with the expert's eye movement data, showed a main effect of condition, $F(2, 69) = 55.08, p < .01, \eta_p^2 = .62$.

Bonferroni post-hoc tests revealed that the spotlight display condition guided the students' attention most successful compared to the dot display condition, $p < .01$, and to the control condition, $p < .01$. Still, attention guidance was more effective in the dot display condition than in the control condition, $p < .01$.

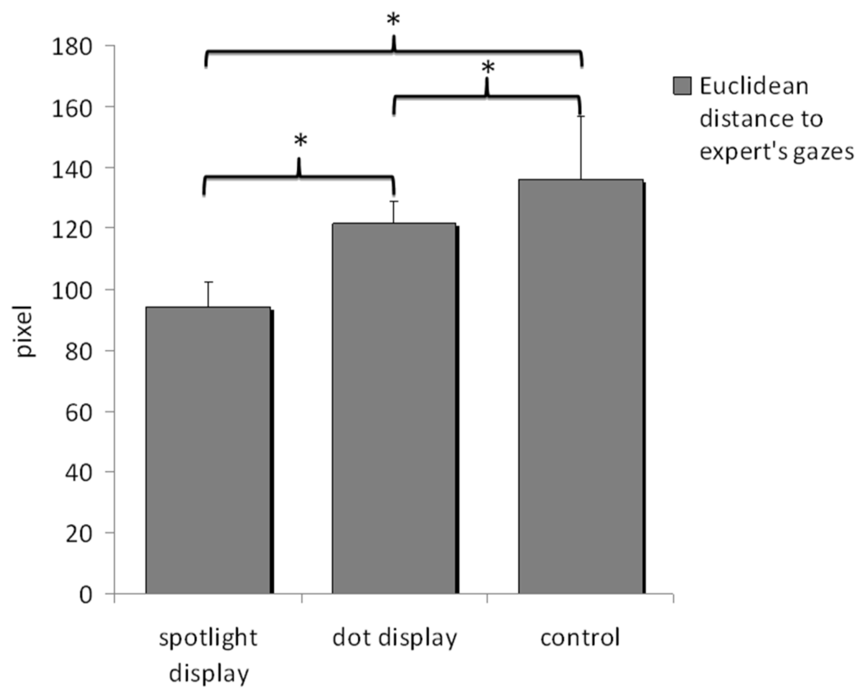


Figure 5.1: Means and standard deviations for Euclidean distance (in pixels) between expert's and students' gaze points over time during example study.

5.2.2 Learning Outcomes

5.2.2.1 Eye Movement Data During Testing

Results for the time before first looking at relevant AOIs are summarized in Figure 5.2 (the exact values are given in Appendix Study 2, Table 3). An ANOVA was conducted for this dependent variable and condition (control vs. dot display vs. spotlight display) as independent variable. The conditions significantly differed in the time participants took before looking at a relevant AOI for the first time, $F(2, 69) = 4.19, p = .02, \eta_p^2 = .11$. Bonferroni post-hoc tests indicated that the spotlight display group looked significantly faster at the relevant AOIs than the control group, $p = .02$. The dot display group did not differ significantly from the other two groups, ns.

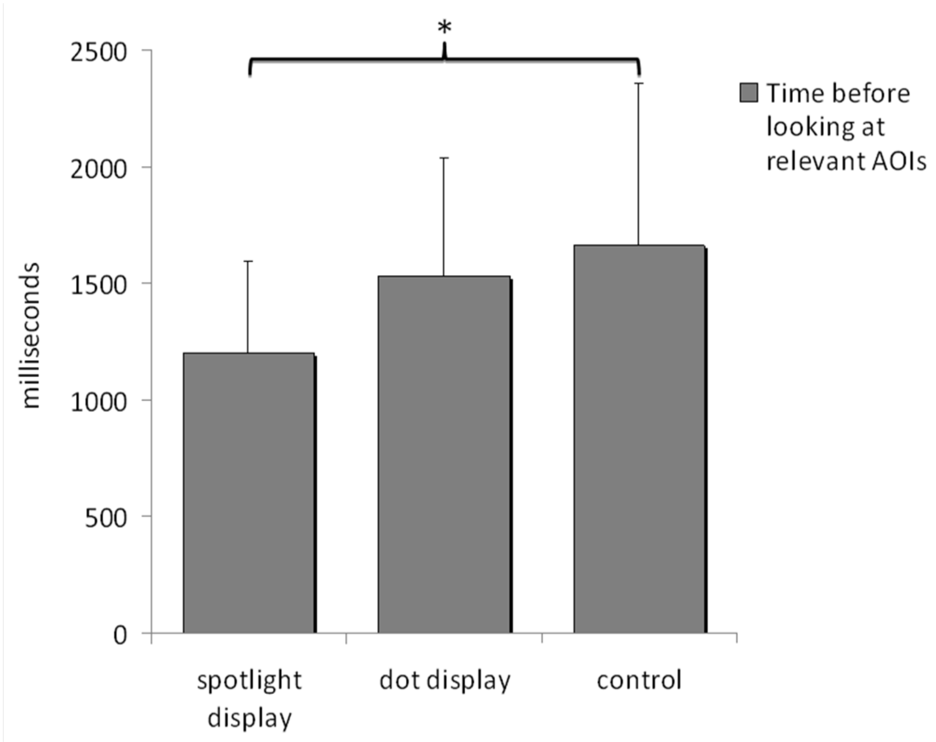


Figure 5.2: Means and standard deviations for time before looking on relevant AOIs for the first time (in ms) during testing.

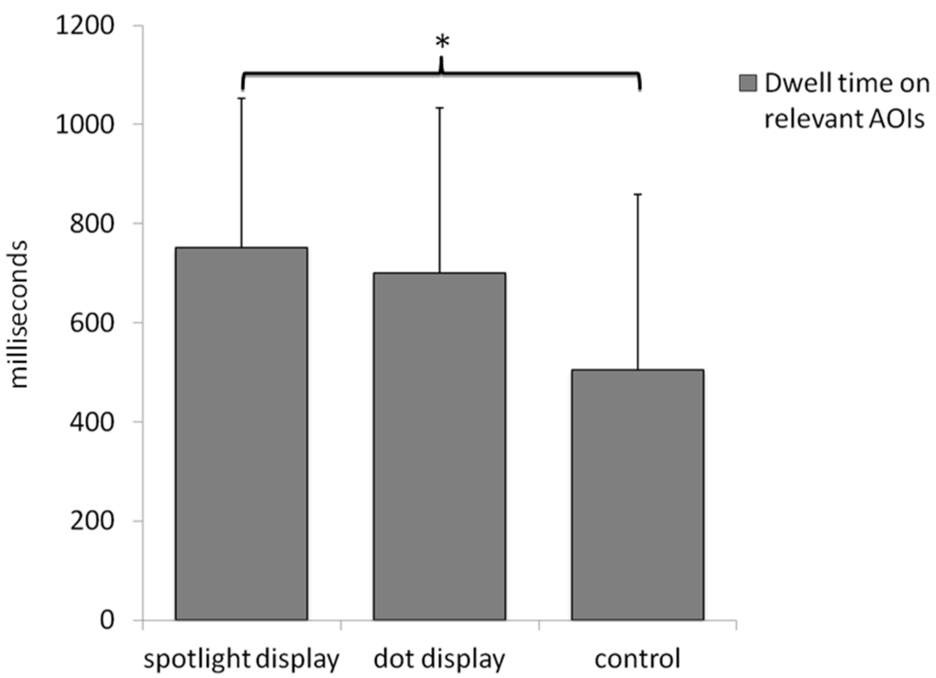


Figure 5.3: Means and standard deviations for dwell time on relevant AOIs (in ms) during testing.

Results for the dwell time on relevant AOIs are summarized in Figure 5.3 (the exact values are given in Appendix Study 2, Table 3). An ANOVA was conducted for this dependent variable and condition (control vs. dot display vs. spotlight display) as independent variable. Dwell time on relevant AOIs also differed significantly among groups, $F(2, 69) = 3.74, p = .03, \eta_p^2 = .10$. Bonferroni post-hoc tests indicated that the spotlight display group looked significantly longer at the relevant AOIs than the control group, $p = .04$. The dot display group did not differ significantly from the other two groups, ns.

5.2.2.2 Multiple-Choice Questionnaire

Results for the correctness scores in the multiple-choice questionnaire are summarized in Figure 5.4 (the exact values are given in Appendix Study 2, Table 3). An ANOVA for this dependent variable and condition (control vs. dot display vs. spotlight display) as independent variable showed significant differences between conditions in performance on the multiple-choice test, $F(2, 69) = 3.72, p = .03, \eta_p^2 = .10$. Bonferroni post-hoc tests indicated that the dot display condition outperformed the control group, $p = .05$, and marginally the spotlight display group, $p = .09$. The spotlight display group and the control group did not differ significantly from each other, ns.

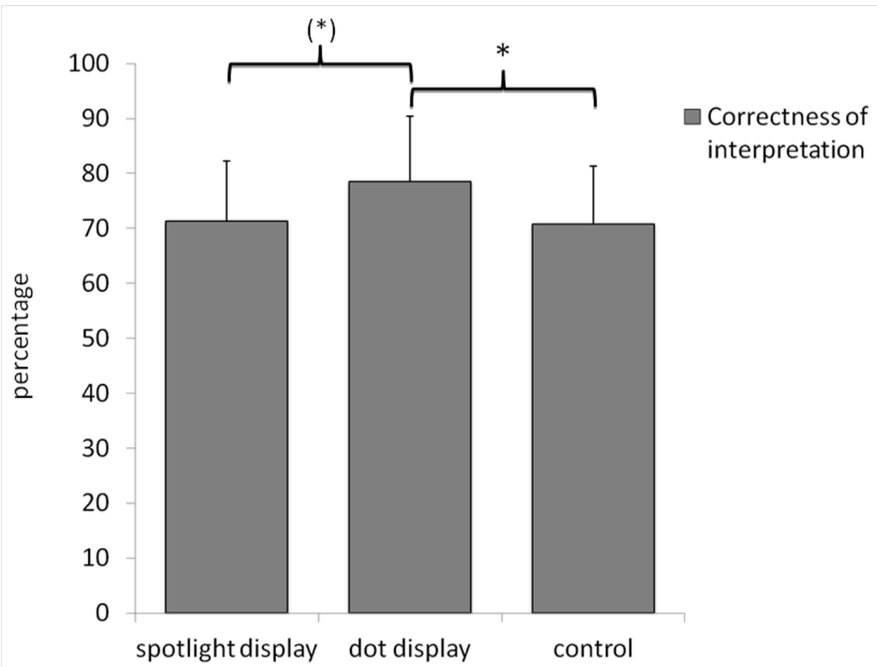


Figure 5.4: Means and standard deviations for correctness of interpretation score (in percent) for the test videos.

5.3 Discussion

Study 2 tested the effect of two versions of the instructional method EMME on supporting novice students – in particular on a perceptual level – in classifying fish locomotion patterns. This was done by developing modeling examples that included attention guidance to relevant aspects of the video based on eye movements of an expert model. Two ways of displaying the eye movements of the expert were compared to a control group without attentional guidance: a manufacturer-provided dot display (cf. Van Gog et al., 2009) and a spotlight display (Dorr et al., 2010). It was hypothesized (Hypothesis 1) that attention guidance during example study would result in a scanpath that is more similar to the expert model's scanpath in the spotlight display condition compared to the dot display condition. The condition without attention guidance would have the least similar scanpath to the expert model. Furthermore, it was hypothesized (Hypothesis 2) that the eye movement modeling example groups would better learn what to attend to, thereby yielding more efficient visual search than the control group, in particular for the spotlight condition. Moreover, I hypothesized (Hypothesis 3) that attention guidance during example study would result also in an enhanced interpretation performance of locomotion patterns compared to the control group. However, it was unclear, which experimental condition would be in favor.

In line with Hypothesis 1, analyses of eye tracking data during example study indicated that both displays of the expert's eye movements were successful in guiding the students' attention, whereby the spotlight display guided the students' attention better. Analyses of eye tracking data during testing revealed that participants in the spotlight condition attended faster and longer to the relevant areas of test videos than the control group, which was in line with Hypothesis 2. Participants also differed in their performance on the multiple-choice questions concerning the fish locomotion patterns shown in test videos. In particular, participants in the dot display group outperformed participants in the control group. This result was only partly in line with Hypothesis 3, because the spotlight display group did not differ from the control group. This might be due to a ceiling effect in the interpretation performance data. As can be seen from the interpretation performance results, the task of interpreting fish locomotion was – contrary to prior expectations – not very difficult for the students as all scored over 70% correct. It seems that the task might have not been sufficiently difficult for learners, once the principle was understood. This might be due to the fact that the decision of which body part is relevant for producing propulsion has not been too difficult for learners, because the relevant parts are always moving, and thus, are salient. Thus, although not all salient parts are relevant, all relevant parts are salient. The second decision of how the relevant body parts are moving, might have also been unexpectedly easy, because it is a dichotomous decision. As a consequence, a task that has a more complex underlying decision tree, namely diagnosing epileptic seizures in infants will be used in Study 3. It is expected that this task will benefit more from EMME then the task to classify fish locomotion patterns, because it is not only more difficult from a perceptual perspective (relevant body parts are often less salient than irrelevant body parts), but also has a more difficult underlying procedure (more decisions need to be taken based on the video examples than in the fish locomotion classification task). This task will be described in detail in the following chapter.

On the other hand, the finding that the spotlight groups' attention was more focused during learning and showed a more efficient visual search performance on the test than the control group, whereas it was the dot display group showed a better locomotion classification performance than the control group, might also be due to the fact that the non-fixated areas were too much obscured in the spotlight condition. Participants in the dot condition saw the entire, visually complex material during example study. Whereas this may have enticed them to look away from the expert's gaze to inspect other parts of the video once in a while (as suggested by the fact that the spotlight group followed the expert's gaze more closely), it also may have allowed them to perceive the relevant body parts together with their context and thus, to probably better understand their motion. In contrast, participants in the spotlight condition saw only the attended area, but not its context. This, in turn, might have hampered the understanding and thus, the later interpretation of the motion pattern in the test video.

Consequently, in Study 3 both display types will be modified slightly in order to improve their effect on learning. Although the *spotlight display* was successful to guide learners' visual attention and positively affected visual search, this was not reflected in learners' ability to interpret the locomotion pattern. As described above, this result might be due to the fact that the spotlight strongly occluded context information during training rendering it impossible to gain a holistic overview of the overall motion pattern. Thus, the spotlight display might be better designed in a less intrusive way, so that it guides attention in a more subtle way and at the same time preserves some overall impression of the motion (cf. Nyström & Holmqvist, 2008). This will be done in Study 3 by the so called ”foveation” technique (Nyström, 2008). In that, the spatial contrast is lowered on unattended areas, while color and temporal contrast remain unaltered. Thus, the overall motion is not blurred. As a consequence, a holistic overview over the entire scene – including motion – is possible.

In the *dot display* condition participants' ability to interpret locomotion patterns in test videos was better than in the control group, however, their visual search was not improved compared to the control group. This might be due to the fact that the solid dot occluded the exact spot the expert model looked at. Thus, participants learning with the dot display were guided ”around” the relevant areas during learning instead being guided directly on the spot the expert model looked at. Hence, they might have experienced problems to correctly locate those areas quickly during testing. This would be even more severe in the task of describing epileptic motion in Study 3, because the relevant areas in this task are sometimes very small (e.g., an eye) and would be fully occluded by a dot. Thus, the dot display, might be improved by displaying gazes not as a solid dot that enlarges and contracts, but instead as a transparent dot that does not change in size and that does not occlude relevant details. Thus, in Study 3 a circle display will be used instead of a dot display.

In sum, the experiment showed not only that attention can be guided based on experts' eye movements, but also that this type of attention guidance fosters learning in terms of improved visual search and interpretation of relevant body parts. None of the display designs, however, could improve both measures of learning – visual search and interpretation – simultaneously.

Chapter 6

Analyzing the Task of Diagnosing Epileptic Seizures in Infants

This chapter presents a second task besides fish locomotion patterns, where perceptual skills are likely to play a crucial role in the interpretation of motion patterns, namely the semiological diagnosis of epileptic seizures in infants. Stating a correct and detailed diagnosis of infant epilepsy is a crucial medical problem: it is important to provide a prognosis, to choose a medical treatment (pharmacological and therapeutic), and last but not least to allow researchers and practitioners in different fields to communicate so that they can develop new knowledge on causes and treatments of epilepsy in infants (Dreifuss & Nordli, 2000).

To diagnose an epileptic seizure several sources of information may be taken into account: the age of the onset of the symptoms, the etiology (i.e., the cause), the patients' anatomy, precipitating factors (i.e., the trigger of the symptoms), and many more. Using these different information sources for diagnostic purposes, however, often leads to two errors: unspecific diagnoses or underdiagnosis (i.e., either diagnosed as another disease or as healthy). On the one hand, when using the above mentioned criteria, epilepsy is often diagnosed as "unclassified epileptic seizures". Such a classification, however, is not helpful, because it does not deliver valuable diagnostic and prognositc information. On the other hand, certain seizure types are frequently not diagnozed as such. For instance, focal seizures in infants are underdiagnosed, because their symsptoms are too subtle (Nordli, 2002) or because other behavioral or motion patterns are more salient (Nordli et al., 1997). Thus, diagnosing epileptic seizures based on the above mentioned information sources has been criticized as not being helpful in clinical praxis (Dreifuss & Nordli, 2000).

To avoid these problems, a new diagnostic system has recently been introduced. In this diagnostic system epileptic seizures are diagnosed according to a description of the semiology of ictal (i.e., seizure induced) events, that is, by describing the behavioral and motion patterns during an epileptic seizure (cf. International League Against Epilepsy, 2010; Loddenkemper et al., 2005; Lüders et al., 1998). Based on the observed motion patterns and verbal descriptions of the seizure by the patient, epileptic seizures can be diagnosed according to which parts of the brain are involved, to which degree the state of consciousness is impaired and with regard to the type of movement that is caused by the seizure (e.g., International League Against Epilepsy, 2010; Cavazos & Spitz, 2009). Diagnosis based on semiology showed to be more reliable than standard clinical observations and interictal EEGs (Nordli et al., 1997). Moreover, semiological diagnosis is practical and easy to apply (Hamer, Wyllie, Lüders, Kotagal, & Acharya, 1999).

However, the semiological diagnostic system was mainly derived from a large number of observations of seizures in adults (Nordli, 2002). Starting with the age of about 5 or 6 years infants' epileptic syndroms are similar to those of adults and the semiological diagnostic systems can easily be used (Nordli, Kuroda, & Hirsch, 2001). Before that age, however, this type of diagnostic systems is not easy to apply. For example, very young infants cannot provide a verbal description of their experiences and state of consciousness during the seizure or perceptual disturbance experienced before a seizure (so-called auras), which in turn makes diagnosis very difficult (Dreifuss & Nordli, 2000). Hence, Nordli and colleagues (Dreifuss & Nordli, 2000; Nordli, 2002; Nordli et al., 1997) recommend to describe only the ictal motion patterns when diagnosing epileptic seizures in small infants. For instance, focal seizures are characterized by specific movements of the hand or arm (jerking), specific motions within the face (lateral version of the eyes), and the inability to stop this motion by touching the infant (Nordli, 2002). In particular, as these types of behaviors exhibited by infants during seizures are easy to understand (Acharya et al., 1997), diagnosis based on semiology should be applied to infants that are suspected to suffer from epilepsy.

As this type of diagnosis is so important for the clinical practice, it is crucial that medical students learn to perform it. Hence, the question is how to convey the crucial knowledge and skills for this task. As this task heavily relies on perceptual input, EMME might be an appropriate instructional method in this context. Before implementing EMME for this task, however, a task analysis will be conducted to investigate whether EMME fit to the specific requirements of this domain.

6.1 Analyzing the Task of Diagnosing Epileptic Seizures in Infants in Semiologic Terms

As described in Chapter 2, a task analysis is a prerequisite for a successful instructional design. Similar to the fish locomotion classification task, the task of diagnosing epileptic seizures in infants will be analyzed in accordance with the approach of Schaafstal and Schraagen (1992): The next section will provide a normative task analysis (describing how the task *should* be performed), whereas Section 6.1.2 will provide a task analysis based on empirical findings (describing how the task *is* performed).

6.1.1 Normative Task Analysis Based on Theoretical Considerations

According to Schaafstal and Schraagen (1992) a normative task analysis (cf. Section 2.2) consists of a description of the task procedure (i.e., of the steps that need to be performed to solve the task) and the specification of the required knowledge and skills to execute each step. This results in a so-called task model, which can be presented in a table format. Next, this task model is used to predict difficulties likely to occur while performing this task. In a final step, instructional goals are derived from the task model.

6.1.1.1 Analyzing the Task Procedure

The International League Against Epilepsy (ILAE, International League Against Epilepsy, 2010) recommends to diagnose epileptic seizures according to the source of the seizure in the brain hemispheres, the degree of impairment of patients' consciousness, and the type of the movements caused by the seizure[1]. The transformation and extension of these diagnostic guidelines into a concrete task procedure based on semiological descriptions of ictal motion patterns was conducted with the help of a domain expert and teacher in the field of pediatric neurology (T. Balslev, practicing neurological pediatrician and teacher in this domain, personal communication, September, 2009). To identify the source of the seizure in the brain, it has in a first step to be decided which parts of the body move in a suspicious manner. The motion can either be lateral or bilateral and involve different body parts. Second, it has to be decided upon how these body parts move. These motions can be jerky, spasmic, myoclonic, etc. Third, the infant's level of consciousness has to be determined. Fourth, it has to be decided whether the face is moving in a suspicious manner (this allows to differentiate a seizure from normal behavior, like benign sleep myoclonus, Egger, Grossmann, & Auchterlonie, 2003). Fifth, it has to be decided wheather the movements change after touching the infant (this allows to differentiate a seizure from normal behavior, like infantile masturbation, Hansen & Balslev, 2009). Sixth, based on these decisions diagnostic codes can be assigned (e.g., according to the diagnostic systems of ICD-10 by the World Health Organisation, 2006 or the ILAE system, see International League Against Epilepsy, 2010). Table 6.1 provides an overview of the steps that need to be taken to perform this task.

6.1.1.2 Identifying the Knowledge and Skills Required to Perform the Task

The task investigated in Study 3 is a form of medical diagnosis. According to Van Lehn (1989, p. 528) medical diagnosis is a **knowledge-rich task**. A knowledge-rich task is a task

[1]Moreover, the ILAE recommends to take age of seizure onset and EEG measures into account. I decided to forgo these further diagnostic criteria in the task description, because they do not refer to directly observable behavior of the patient, which is in focus of the semiological diagnosis of epileptic seizures in infants (Dreifuss & Nordli, 2000; Nordli, 2002; Nordli et al., 1997).

Table 6.1:
Procedure of the Task to State an Initial Diagnosis of Epileptic Seizures in Infants

Steps to be performed	Possible outcomes
Affected body parts	One or both arms, one or both legs
Motion pattern	Spasm, jerk, twitch, myoclonia, muscle tense
State of consciousness	Awake and conscious, sleepy, drowsy, unconscious
Involvement of face	Yes or no
Change after touching	Yes or no
Diagnostic code	not diseased or G40.0 - G41.9 (according to World Health Organisation, 2006, for the coding of the ILAE system see International League Against Epilepsy, 2010)

in which not all information needed to solve this task are given in the problem statement. Instead, to perform this task either prior knowledge or training is required (Van Lehn, 1989). Hence, this also will be the case for the medical diagnosis investigated in Study 3. The question arises *what* to train. Since this task requires the inspection and examination of a perceptual input – without it the task cannot be performed – it is a **perceptual task** according to the terminology used in this thesis (cf. Chase & Chi, 1981; Chi, 2006a). In the following, the knowledge and skills required to perform this perceptual and knoweldge-rich task will be outlined.

The procedural task analysis above indicates that possessing conceptual knowledge (e.g., on different diagnostic codes) and being able to apply it is necessary but not sufficient for successfully performing the task of stating a semiological diagnosis of an epileptic seizure. A prerequisite for applying this conceptual knowledge is to possess the appropriate perceptual skills (i.e., a rapid visual search and identification, as well as a visual inspection and interpretation of the relevant areas of a visual stimulus, cf. Chi, 2006a; Gibson & Pick, 2000; Manning et al., 2005).

Table 6.2 provides an overview of the task procedure and the required knowledge and skills in the form of a task model.

Table 6.2:
Task Model for the Task to Diagnose Epileptic Seizures

Task steps	Required knowledge and skills
Perceptual input mandatory, based on **perceptual skills**:	
1. Specifying *body parts* that might be affected by the disease	Visual search and identification of uncommonly moving body parts
2. Specifying the *motion pattern* of these body parts	Visual inspection of suspected body parts and interpretation of their motion
3. Specifying infant's state of *consciousness*	Visually inspecting and interpreting facial indicators for consciousness
4. Indicating the involvement of the *face*	Visual search and identification of uncommon motion or lack of it within the face
5. Indicating a *change* in motion after touching	Visual search and identification of change in motion
Perceptual input NOT mandatory, based solely on **conceptual knowledge**:	
6. *Diagnosis* of the disease	Assignment of observations to the correct diagnostic code

The current task is – as the task to classify fish locomotion patterns – a knowledge-rich task and requires schemas to be executed (Cooper & Sweller, 1987). The top-level schema for this task can be considered to consist of five slots that need to be instantiated with example specific information. Lower-level schemas might represent different types of suspicious movements of different body parts, different motion patterns of body parts, like spasms, facial indicators for consciousness, or different diagnoses). The decision on which body parts move in a suspicious manner and whether the face might be involved requires the visual search and identification of the relevant body parts (arms, legs, eyes, ...). The question whether changes occur after touching the child also requires visual search and identification of the relevant body parts at the right time. The decision on the type of the movement as well as the decision on the infant's level of consciousness both require the visual inspection and correct interpretation of the relevant body parts. Hence, the first five decisions necessary for the task procedure are based on the examination of perceptual input and thus, require "perceptual skills". After these slots of the top-level schema are instantiated, a schema specific solution procedure can be executed (Van Lehn, 1989). In the current task, this would be the assignment of the observations to a diagnostic code. This final step of the task can be performed without perceptual input and is thus, purely based on conceptual knowledge.

In sum, diagnosing epileptic seizures based on ictal motion patterns requires perceptual skills (i.e., skills that cannot be performed without a perceptual input) that are composed of the following subskills:

1. Detecting affected body parts, involvement of the face, and changes in motion: Visually searching for the relevant areas (i.e., quickly looking at them) and identifying the relevant areas (i.e., denoting them as being relevant).

2. Detecting how these body parts move and identifying the state of the infant's consciousness: Visually inspecting the relevant areas (i.e., looking at them for a sufficiently long time) and correctly interpreting the motion of those relevant areas (i.e., describing their motion correctly).

6.1.1.3 Predicting Potential Difficulties of the Task

The last section provided a task model for the task to diagnose epileptic seizures in infants in semiological terms. Based on this task model this section will provide a prediction of difficulties that are likely to occur when performing this task. In that, difficulties may be caused by the procedure of the task itself and by characteristics of the perceptual input. In clinical practice the actual infant patient would be observed. To create a standardized set of stimuli, video recordings of real infant patients during an epileptic seizure were used in the course of this thesis (so called patient video cases, PVCs, e.g., Dequeker & Jaspaert, 1998).

Difficulty Caused by the Task Procedure. To perform the task of diagnosing epileptic seizures in infants *five* perception-based steps need to be conducted (cf. Table 6.2). Since the procedure of this diagnostic task requires more steps to be performed, it is more difficult then the task used in Study 1 and Study 2, which required only two perception-based steps (cf. Table 3.1).

Within each of these steps, unexperienced people (i.e., novices) may make mistakes. In the first step, novices may not name body parts that were affected by the disease, but body parts showing a salient movement. In the second step, novices may describe the motion of those body parts in a wrong way. For instance, the infant in the test PVC depicted in the lower part of Figure 6.1 displays a salient motion by waving around her arms and legs. This, however, is not that motion that is indicative for an epileptic seizure. Now and then she displays more subtle spasmic movements of the same body parts that are a symptom of this disease. A novice may pay attention and describe the salient, but not relevant motion type, and ignore the subtle but relevant motion. In the third step, a novice may not be able to interpret the infant's state of consciousness correctly. In the fourth step, a novice may not realize that the face is also involved in the disease. In the fifth step, a novice may miss to recognize whether the motion changes after the infant is being touched. In the last step, the five decisions may be assigned to a wrong diagnostic code.

As demonstrated, unexperienced people may fail at all six steps of this task. The first five steps, however, are a necessary prerequisite to approach the sixth step. Without conducting the first five steps correctly one will fail the last one. Hence, the first five steps are of particular importance in this task. As already mentioned these steps cannot be executed without a perceptual input. Thus, in the next section difficulties that might be caused by the perceptual input are described in greater detail.

Difficulties Caused by the Perceptual Input. Under real-life task performance conditions, the perceptual input in this task would be an infant patient lying in a pediatric ward in hospital. To standardize these conditions, video recordings of these patients were used (PVCs). However, it is important to note that the core characteristics of both types of perceptual input (i.e., living patient vs. video recordings) are similar and so are the difficulties that they may cause for the observer. This is the case, because the video camera that is controlled and angeled by a person from clinical staff in a pediatric ward of a hospital substitutes for the observer him- or herself. Three exemplary screenshots of the PVCs used in Study 3 are presented in Figure 6.1. In the following, to analyze the characteristics of these stimuli, research on learning from dynamic visualizations is consulted as in this research challenges that may occur when inspecting highly information-dense and dynamic input have already been well investigated. Accordingly, the findings from this research can be well applied to real-life scenarios, where the perceptual input is characterized by the same features.

The stimuli in this task transport a large amount of irrelevant information. This characteristic is likely to cause difficulties for visual search and identification of relevant information (i.e., Step 1, 4, & 5). As discussed in Section 3.1.3, research on learning from highly realistic representations has shown that irrelevant features are challenging for learners (Scheiter et al., 2009), since they may attract attention and guide attention away from more relevant elements (Dwyer, 1976). Moreover, relevant elements may be difficult to detect. As can be seen in Figure 6.1, the examples used in Study 3 include many irrelevant information (everything outside of the yellow marking is irrelevant) that might potentially be distracting, thus, focusing on the relevant areas and interpreting them correctly is likely

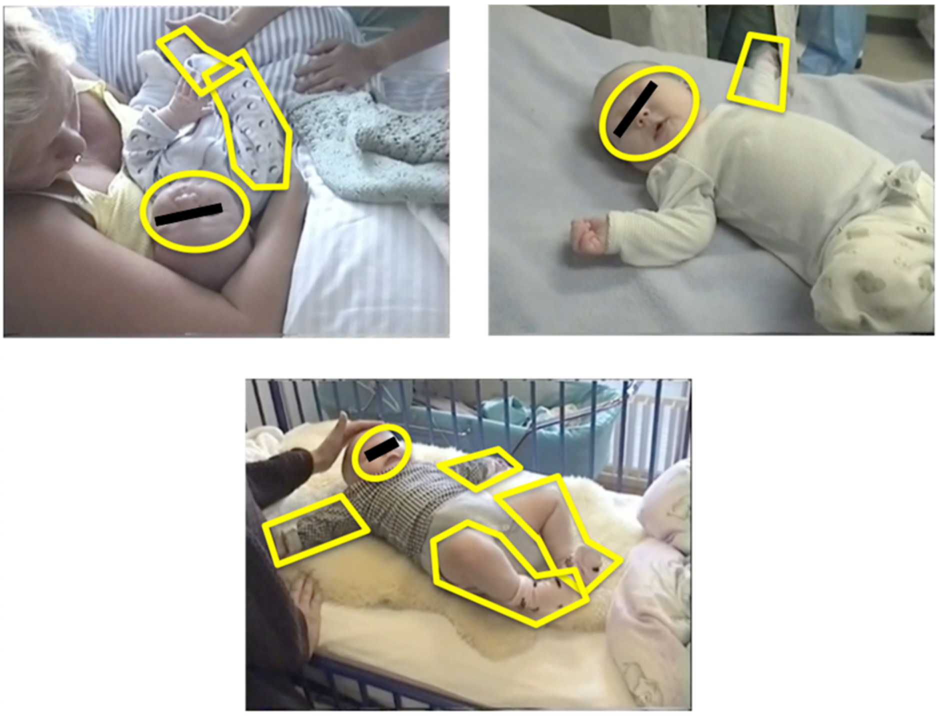


Figure 6.1: Screenshots from three patient video cases used as examples of seizures in Study 3 to test the participants' performance. The areas, for which movement is relevant for description and needs to be interpreted, are marked yellow.

to be challenging for novices (cf. Study 1). Furthermore, the infant patients suffering from a seizure are *dynamic* stimuli (cf. definition of Schnotz & Lowe, 2008). This characteristic is likely to cause difficulties for visual inspection and interpretation of relevant information (i.e., Step 2 & 3). The videos of infant patients used in Study 3 display temporal change (cf. Lowe, 2003, 2004) in terms of translations (i.e., position changes of single elements), transitions (i.e., visual appearances and disappearances of single elements), and transformations (i.e., visual changes of size, shape, color, and texture of single elements). For instance, when an infant's arm jerks in a PVC, its shape, size, and the position on the monitor change. Extracting the crucial information from dynamic stimuli can be challenging for learners (Lowe, 2008). Moreover, the information in these dynamic stimuli is *transient* (i.e., only temporary available) which means that once an information is passed or changed it cannot be accessed any more by the observer to compare it to the current status (Hegarty, 1992). Hence, the observer has to keep all passed relevant information in mind to relate it to the current status. This, however, may be very difficult, in particular, for novices. Furthermore, this dynamic information is time-critical, hence, the observer has to attend to the relevant information in a certain point in time, otherwise, s/he would miss it. Another difficulty in this task is that the visual saliency of single elements of the patient's body is not necessarily related to their relevance for task performance. The absence of movement or very subtle movements are often relevant (e.g., subtle motions in the face). Such a weak relation between visual saliency and task relevance, however, has shown to impede task performance (Lowe, 2003, 1999).

All these described characteristics of the infant patients in the focus of Study 3 are equal to those of the swimming fish investigated in Study 1 and 2. In addition to that, however, the infant patients also display a motion of several elements at the same time according to *different types of motion*[2], which may cause a split-attention effect (Lowe, 2003, 2004). Moreover, the motions displayed by the infants did not re-occur as the fish locomotion pattern did. Thus, the infants' motion patterns are even more time critical. Thus, the stimuli used in Study 3 involve a more complex dynamism and thus may cause more difficulties for novices than the video examples used to classify locomotion patterns of fish.

6.1.1.4 Instructional Goals Derived From the Normative Task Analysis

The normative task analysis revealed that training the ability to inspect and examine the perceptual input (i.e., training of perceptual skills) seems to be of prior importance for this task. This is supported by three findings.

First, five out of six steps of the task to diagnose epileptic seizures in infants require perceptual skills. As can be clearly seen from the description of the task procedure, the larger part of the task is based on making the correct observations, and thus, based on perceptual input. Only the final step of the task can be performed without the perceptual input. Hence, making correct observations from the perceptual input plays a dominant role in this task. Thus, according to the procedural task analysis it is an important instructional goal to provide learners with a training that allows them to correctly handle the perceptual input.

Second, the steps of the task that are bound to the perceptual input require the follow-

[2] Although some of the fish motion patterns in focus of Study 1 and Study 2 included the motion of two fins, these fins always moved in exactly the same way.

ing perceptual subskills: visually searching and identifying relevant body parts as well as visually inspecting and interpreting the motion of these body parts. Accordingly, training these subskills should be a main instructional goal.

Third, the difficulties that may occur in this task mostly affect perceptual skills. The perceptual input – PVCs – involves a large amount of irrelevant information due to its realism. As a consequence, relevant information may get out of focus and may be difficult to detect. Furthermore, the PVCs include dynamism of various types: they include transformations, translations, and transitions. All those types of dynamism may occur for several elements at a time. Guiding novices' attention away from irrelevant areas to relevant areas at the right time might facilitate the processing of these PVCs. Beyond the difficulties caused by the perceptual input the task procedure itself may potentially cause difficulties (cf. Section 6.1.1.3). Again, the difficulties from the task procedure mostly affect perceptual skills, since most of the steps require the inspection of a perceptual input, the difficulties that may occur mostly affect perceptual skills. Thus, perceptual skills have a high danger of being jeopardized in this task. Consequently, based on the analysis of the difficulties that are likely to occur in this task, it is important to place a high emphasis on the training of perceptual skills.

According to the normative task analysis, it can be expected that more perceptual difficulties may occur in this task than in the fish locomotion classification task in Study 1 and 2. Hence, perceptual skills might be even more important in the task to diagnose epileptic seizures. Furthermore, the task procedure is also more complex and requires the observer to visually search and identify as well as visually inspect and interpret more elements than in the fish locomotion classification task. Again, the perceptual skills seem to be more crucial for pursuing the task procedure than in the fish locomotion classification task.

In sum, based on the normative task analysis a main instructional challenge and goal for this task is to convey perceptual skills. The following section discusses which instructional goals for this task may additionally be drawn based on empirical findings.

6.1.2 Empirical Task Analysis Based on Empirical Evidence on Task Performance

Following Schaafstal and Schraagen (1992) a task analysis should consist of a normative analysis, which was conducted in the last section, and an empirical analysis, which will be provided in the current section. An empirical analysis is usually done by investigating how experts and novices perform a given task (cf. Section 2.2.3). However, this is not always necessary as long as the aim for performing an empirical task analysis can be achieved alternatively, for instance, if enough information from empirical evidence on how the task under investigation would be performed is already given in the existing literature[3]. The aim for performing an analysis of the current task is to investigate whether EMME are appropriate as instructional method to teach this task. Thus, it might not be necessary to analyze the task in the greatest possible detail. Instead it might be sufficient to collect empirical evidence that this task and its requirements fit to the constraints elaborated for the EMME approach in Section 4.2.5. To conduct an empirical analysis without actually performing an expertise difference study, Schaafstal and Schraagen (2000) propose to conduct a literature study and to compare the task to a similar task for which an empirical task analysis was already conducted. This procedure can also be chosen for the current task of diagnosing epileptic seizures in infants. As pointed out in the normative task analysis, both tasks used in this thesis are rather similar in that they both require a careful examination of a perceptual input. In both cases, the perceptual input is related to movement patterns of living beings that are either inspected live or on video examples, both of which are characterized by realism and dynamism. As a thorough empirical task analysis was already conducted for the task to classify fish locomotion patterns, existing literature from medical domains will be reviewed and conclusions that can be drawn from that will be compared to those drawn for the empirical task analysis of the fish locomotion classification task. This will be done with regard to the task procedure, the knowledge and skills required to perform the task, the difficulties that occur when performing this task, and the instructional goals that can be derived.

[3]The reason for Schraagen to take this approach was because ”this kind of study exceeded the projects budget” (Schraagen, 2006, p. 197).

6.1.2.1 Analysis of the Task Procedure

Instead of performing a task step by step, experts use shortcuts or heuristics (e.g., Hoffman, Coffey, & Ford, 2000; Hutchins, 1995; Lave, 1988; Newell & Simon, 1972; Schraagen, 1993). Study 1 was – to my knowledge – the first one to directly confirm the use of shortcuts in experts on a perceptual level by means of eye tracking. The use of shortcuts by experts was also found by a large body of research in the medical domain (e.g., Boshuizen & Schmidt, 1992; H. G. Schmidt & Boshuizen, 1992, 1993). It was found that with increasing experience, physicians do not diagnose based on scientific facts (so called biomedical knowledge), but rather based on experience with former patients to diagnose similar patients (so called illness scripts). These findings for medical diagnosis, however, do not refer to tasks that require perceptual input, but to tasks that are based on verbal and numerical materials. Nevertheless, since experts' use of shortcuts was demonstrated for perceptual tasks (Study 1) and since it was also shown for medical diagnosis (e.g., Boshuizen & Schmidt, 1992; H. G. Schmidt & Boshuizen, 1992, 1993), it cannot be excluded that shortcuts will be used by experts in the diagnosis of epileptic seizures in infants which is based on perceptual input. For instance, instead of analyzing the diagnostic features suggested by the International League Against Epilepsy (2010), an experienced physician might draw conclusions based on the age and gender of an infant patient towards a certain diagnosis. However, as the aim of the current task analysis is to draw conclusions with regard to efficient instructions for this task and not to provide an as detailed description of the task as possible, it is not necessary to analyze these potential shortcuts in depth as they will not be taught, but avoided. Nevertheless, it has to be ensured that an expert used to generate EMME for this task does not use shortcuts in this task, for instance, by using a suitable instruction.

6.1.2.2 Analysis of the Required Knowledge and Skills

Many studies show that experts possess sophisticated perceptual skills that allow them to attend fast and enduring to relevant areas of a perceptual input as compared to novices (e.g., Canham & Hegarty, 2010; Charness et al., 2001; Haider & Frensch, 1999; Lowe, 1999;

Underwood et al., 2003; Van Gog, Paas, & Van Merriënboer, 2005). Although these studies only investigated static representations, Study 1 showed that both facts also hold true for dynamic perceptual input. A question might be how these findings are expandable to the medical domain. There is already a large amount of research stating that experts exceed novices on a perceptual level in diagnosing from static representations, for instance, X-rays, microscopic slides, and mammograms (e.g., Krupinski, 2005; Krupinski et al., 2006; Kundel, Nodine, Krupinski, & Mello-Thoms, 2008; Lesgold et al., 1988; Nodine et al., 1996). Taken together, experts have shown to have superior perceptual skills than novices in medical domains using static representations and in Study 1 of this thesis in a biological domain using dynamic representations. Thus, it can be assumed that this finding will also hold true for a medical domain using dynamic representation, namely in diagnosing epileptic seizures in infants. Moreover, preliminary results provide first evidence in favor of such expertise effects for this task (Balslev et al., 2011).

6.1.2.3 Analysis of Potential Difficulties of the Task

Novices get easily distracted by salient but potentially irrelevant elements of a perceptual input, while missing relevant information (Lowe, 1999). Study 1 confirmed this finding for dynamic perceptual input directly by means of eye tracking. Similar findings based on eye tracking and other methods were also found in the medical domain, for instance, in diagnosing based on static and more abstract representations, like microscopic slides (e.g., Krupinski et al., 2006) as well as in diagnosing based on realistic representations, like photographs of patients (Brooks, LeBlanc, & Norman, 2000). Since similar difficulties were shown directly for dynamic, realistic stimuli (Study 1) and in the domain of medical diagnosis based on static perceptual input (e.g., Krupinski et al., 2006), it can be assumed that they also will occur for medical diagnosis based on PVCs.

In particular, these difficulties are very likely to occur in the task of semiological diagnosis of epileptic seizures in infants. According to Nordli et al. (1997) infantile seizures may be so subtle that even skilled observers may fail to identify them correctly. Moreover, some of these motion patterns occur only occasionally, are short-term, subtle, time-sensitive, and not salient compared to other movements or characteristics of the infant (Hansen & Balslev, 2009)[4]. In that, the subtle motion patterns may be either not recognized at all or attributed to other diseases (Nordli, 2002, for a list of so-called epilepsy ”imitators” for children in different age groups see Prensky, 2000, p. 100). For instance, epileptic seizures may easily be mistaken for normal behavior (Egger et al., 2003; Hansen & Balslev, 2009).

In sum, all conclusions required for the current purpose of this thesis can be answered based on findings from Study 1 of this thesis and on existing research. Thus, conducting a study on expertise differences in the domain of diagnosing epileptic seizures in infants for task analytical purposes might not be necessary[5].

6.1.2.4 Instructional Goals Derived from the Task Analysis

The aim of this empirical task analysis was to ensure that the task of diagnosing epileptic seizures in infants fits to the constraints elaborated for the EMME approach, so that EMME can be expected to be a successful instructional approach for the current task. This is the case as can be seen by three main conclusions that can be drawn from the last section: First, given the already existing literature, both in medical and non-medical domains, and given the findings from Study 1 of this thesis (cf. the analysis of potential difficulties of the task), it can be assumed that novices will require training in their perceptual skills to successfully perform the task of diagnosing epileptic seizures in infants. Second, prior research has already shown with regard to static representations in medical and non-medical domains and Study 1 of this thesis confirmed this for dynamic representations, that experts outperform novices with regard to their perceptual skills (cf. the analysis of the required knowledge and skills). Thus, I assume that experts' perceptual (and also cognitive) processes in diagnosing epileptic seizures in infants would be also in principle useful as a model for novices. Re-

[4]With age, these symptoms become more prominent and are thus, easier to assign to an epileptic diagnosis (Nordli et al., 2001).

[5]Nevertheless, a more detailed empirical analysis of this task would be interesting to get a better understanding of expertise effects in this task. Therefore, such a study was conducted and the results are under analysis (Balslev et al., 2011). This research question, however, goes beyond the development of an appropriate instructional design and is thus, out of the scope of this thesis.

search on medical education has shown, however, that experts use shortcuts in performing a task. This finding could also be confirmed directly on a perceptual level in Study 1 (cf. analysis of the task procedure). Thus, I assume for the current task that experts, before using them as models, should be instructed to act in a didactical manner, as was done for Study 2.

6.2 General Conclusions

This final section summarizes the instructional goals deducted from both task analyses and presents how these are met by means of the EMME approach. As described in the beginning of this section, epileptic seizures are a serious disease that can easily be mistaken for normal behavior or mis-diagnosed. Medical experts recommend to diagnose epileptic seizures based on semiology to reduce this danger (e.g., Dreifuss & Nordli, 2000; Nordli, 2002; Nordli et al., 1997). Thus, it is crucial that medical students learn to diagnose this disease in a semiological manner.

The last sections provided a task analysis to determine how this type of diagnosis should be taught. Table 6.3 summarizes conclusions from the above executed task analyses and their consequences for instruction. The main conclusion is that perceptual skills play a crucial in performing this task. Besides, conceptual knowledge is required to perform a final step in this task procedure. Medical education has focused already very early on teaching conceptual knowledge, thus, many successful instructional approaches to convey this type of knowledge are already established (Feltovich & Barrows, 1984; Feltovich, Johnson, Moller, & Swanson, 1984; Kuipers & Kassirer, 1984; Lesgold, 1984; Lesgold et al., 1988). Teaching perceptual skills, on the other hand, has hardly been in the focus of medical education research. Moreover, the instructional methods already in use to convey this kind of knowledge (e.g., PVCs) can be presumably improved. Thus, there is an obvious need for more research on conveying perceptual skills in medical education.

EMME has already been demonstrated to be adequate for teaching perceptual skills in the task of classifying fish locomotion patterns in Study 2. As can be seen from Table 6.3, the instructional goals of the task to diagnose epileptic seizures in infants are complementary to the EMME approach. Thus, it is very likely that EMME would be appropriate to teach the perceptual skills involved in diagnosing epileptic seizures in infants. Study 3 will examine this claim empirically.

Table 6.3:
Summary of Conclusions Drawn from Task Analysis for Instructional Requirements and how EMME Meets Them

Conclusions from...	Consequence for instruction	EMME meet this issue by ...
Normative task analysis		
The task *procedure* is composed of six hierarchical steps: the first five are based on perceptual input, the sixth is solely based on conceptual knowledge	1. Correctly processing the perceptual input is a prerequisite to successfully perform the task	Aiming at teaching to process perceptual input
The task requires perceptual *skills* (i.e., detecting the relevant body parts and interpreting their movement) and conceptual *knowledge* (assigning the observations to a diagnostic code)	2. Conceptual knowledge can be trained with traditional methods, however, a method to convey perceptual skills is required	Aiming at teaching perceptual skills
Difficulties in this task would occur mainly on a perceptual level. Moreover, the PVCs are highly realistic (i.e., they include irrelevant information) and dynamic (i.e., the information is transient and time critical)	3. Novices' visual attention should be guided	Visually guiding the novices' attention
Empirical task analysis		
Task procedure: Experts may use shortcuts	1. Experts' process data might not be useful without instructing them to behave in a didactical manner	Instructing the expert model to behave in a didactical manner
Knowledge and skills: Experts outperform novices in their perceptual skills	2. Experts' process data is in principle useful for training purposes	Using an expert's perceptual processes as a model for novices
Difficulties: Novices get distracted by irrelevant elements	3. Novices' visual attention should be guided away from irrelevant to relevant information	Guiding novices' visual attention towards relevant elements

Chapter 7

Teaching Perceptual Skills for Diagnosing Epileptic Seizures in Infants via Eye Movement Modeling Examples: Study 3

As outlined in the last section, epileptic seizures in infants should be described according to the infants' motion patterns (i.e., semiological diagnosis of ictal patterns). These patterns, however, may be difficult to recognize (cf. Section 6.1.2.3). Hence, a thorough training of the ability to detect and describe infants' motion patterns during epileptic seizures is important.

The focus in medical education of epileptic seizures, however, does not lie specifically on teaching such skills. Instead, motion patterns of a epilepsy are often paraphrased in textbooks. For an example see the textbook on epilepsy by D. Schmidt and Schachter (2000, p. 10):

> ”The patient opens his eyes with a frightened expression; the right arm is rotated inwards and stiff. He seems to be trying to ward off something menacing. He is screaming loudly throughout the episode, lasting 1 min.”

Only in some books these descriptions are accompanied by static pictures (see Figure 7.1, D. Schmidt & Schachter, 2000, p. 10). Research on medical education has shown that this form of teaching is not adequate to acquire the ability to perform semiological diagnosis (Balslev, De Grave, Muijtjens, Eika, & Scherpbier, 2009; Balslev, De Grave, Muijtjens, & Scherpbier, 2005).

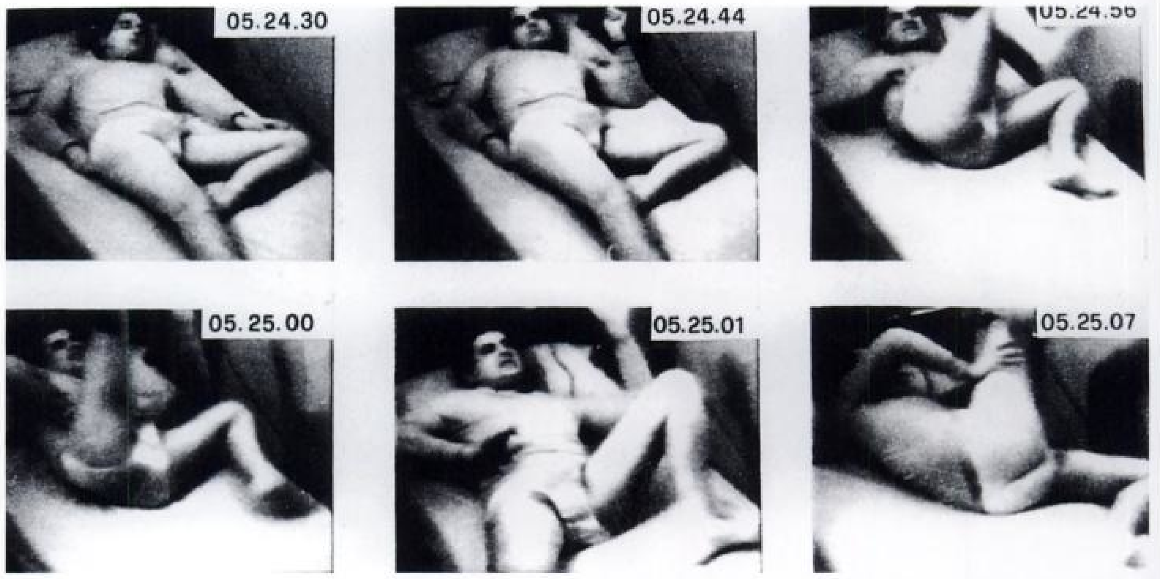


Figure 7.1: Static depiction of a seizure motion. Taken from "Epilepsy: Problem solving in clinical praxis", by D. Schmidt & S. C. Schachter, 2000, p. 10.

Consequently, medical education often makes use of PVCs (e.g., Dequeker & Jaspaert, 1998). Using PVCs for educational purposes has shown to improve diagnostic processes in medical students compared to traditional written descriptions of a patient case (Balslev et al., 2009, 2005; De Leng, Dolmans, Van de Wiel, Muijtjens, & Van der Vleuten, 2007; Kamin, Deterding, Wilson, Armacost, & Breedon, 1999; Kamin, O'Sullivan, Deterding, & Younger, 2003). Unfortunately, none of the studies on PVCs conducted so far focuses on individual learning, which is the focus of the current thesis. Initial analyses of an ongoing study suggest (Balslev et al., 2011), however, that novices have severe problems on a perceptual level when trying to conduct a medical diagnosis based on PVCs. Similar results have already been shown for diagnosing based on static, abstract medical images (e.g., Krupinski, 2005; Krupinski et al., 2006; Kundel et al., 2008; Lesgold et al., 1988; Nodine et al., 1996) and even for realistic photographs of patients (Brooks et al., 2000). Furthermore, analogies to learning from video examples can be drawn from adjacent domains: For instance, in the biological task of classifying fish locomotion patterns, Study 1 of this thesis showed that novices have difficulties in dealing with video examples.

In sum, learning with PVCs is a promising approach, however, novices may still have difficulties to process them. Since EMME have been shown to improve learning from video examples in the task of classifying fish locomotion patterns, it can potentially be used for improving teaching from PVCs as well. Hence, Study 3 of this thesis will investigate whether EMME can be successfully implemented for conveying perceptual skills in diagnosing epileptic seizures in infants. Two types of EMME are compared, an improved spotlight display, and an improved dot display.

Research Questions - Hypotheses

The following hypotheses were tested in Study 3: First, I hypothesize to replicate the effect from Study 2, that displaying the eye movements of an expert in example-based learning guides attention. That is, it is predicted that participants learning with both types of EMME would follow the expert model's eye movements (resulting in a scanpath that is highly similar to the expert model's scanpath) than participants in the control group (whose scanpaths are expected to be more diverse). Since the spotlight display is less intrusive than in Study 2 and the circle display does not obscure elements any more, both should guide learners' visual attention equally well (Hypothesis 1).

As Study 2 has shown that EMME fostered learning of perceptual skills in a biological classification task, I hypothesize that EMME would foster learning of perceptual skills also in medical diagnosis. Again, learning will be measured both in terms of the efficiency of visual search during a test (by means of eye tracking) and in terms of interpreting the motion patterns observed (by means of a multiple choice questionnaire).

Based on the findings from Study 2, it was hypothesized (Hypothesis 2) that the EMME groups would better learn what to attend to, thereby yielding a more efficient visual search than the control group for novel test videos, as indicated by attending faster and longer to relevant areas, while ignoring irrelevant areas. Since, both displays should guide learners' attention equally well to the relevant areas, visual search should be improved to a similar extent by both displays.

Moreover, I hypothesized (Hypothesis 3) that EMME would result also in an enhanced interpretation performance with regard to motion patterns shown in novel test videos compared to the control group. Since the spotlight display was improved in this study, so that it should allow for a holistic overview over the entire motion pattern as the circle display does, I assume both EMME conditions to be equally successful in fostering interpretation performance.

7.1 Method

7.1.1 Participants and Design

Participants were 60 medical students in their final year of the University of Aarhus (41 women, 19 men, $M_{age} = 26.57$ years, $SD = 2.03$). All participants had no prior knowledge on the task and had normal or corrected-to normal vision. They were randomly assigned to one of three conditions (*n* = 20 each): (1) control condition without attentional guidance, (2) attentional guidance by a manufacturer-provided "gaze replay" (SMI) as a circle on fixated areas based on the model's eye movements (circle display), (3) attentional guidance by blurring non-attended areas and leaving fixated areas displayed at a high resolution (spotlight display). Participants were compensated €10 for their participation.

7.1.2 Apparatus and Materials

7.1.2.1 Eye Tracking Equipment

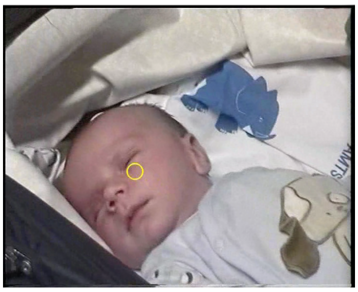
Participants' eye movements were recorded with a SMI High Speed eye tracking system with a temporal resolution of 240 Hz on a 17" monitor and the iView X 2.2 software. The stimulus material was presented via Experiment Center 2.2. The eye tracking data were analyzed with BeGaze 2.3 software (www.smivision.com) and self-programmed MatLab algorithms.

7.1.2.2 Modeling Examples

Diagnosing epileptic seizures in infants may result in many different diagnostic classes. This thesis, however, does not focus on the final diagnostic step, but rather on the perceptual skills underlying this step (as explained in Section 6.1). Hence, it was not the aim to train all possible seizure types, but rather, to train the procedure of how to go about collecting symptoms that a diagnosis can be based upon. Thus, for the purpose of the current study, two exemplary patient cases of focal seizures were chosen to train the perceptual skills. These modeling examples for the control group consisted of two digital PVCs in an audio interleave format (.avi), sized 720 × 576 pixels and presented in fullscreen on a 1280 × 1024 pixels resolution (corresponding to 17.07 × 13.65 inches). Each PVCs depicted a single infant (3 weeks and 7 months old), whereby both infants deployed motion patterns corresponding to an epileptic seizure. The duration of the videos was 71 s and 73 s. Both videos included a spoken description and diagnosis of the motion patterns by the expert model. The original sound was removed from the videos, because parents and clinical staff were talking, which would disturb the use of verbal explanations. The expert was a physician of pediatric neurology, with extensive experience in diagnosing and teaching to diagnose epileptic seizures. Rather than using the expert's natural performance of these tasks as an example, the expert was instructed to behave didactically, that is, to explain to novice students what the relevant aspects of the motion pattern shown in each PVC are. Each recording was replayed to the expert so that he could self-evaluate the replay data based on a number of statements (e.g., for a novice student, the disease is explained in enough detail, in comprehensible terms, et cetera; cf. Jucks et al., 2007), and if necessary, he could re-record it (for a detailed description of this procedure see Section 4.2.3).

In both experimental conditions, participants received the same examples as the control group but those additionally included the expert's eye movements. The displays for the circle display condition were created using the manufacturer rendered "fixation scanpath display function" (SMI). Instead of displaying the entire raw eye movement data set, which is very dense and would very likely be too difficult to follow for an observer, the expert model's

Table 7.1:
Screenshots From the Three Conditions Used in Study 3

Control	Circle display	Spotlight display
		

gaze points were filtered. The filter used was the manufacturer-implemented ”velocity-based saccade detection algorithm” (Smeets & Hooge, 2003). As a result, only the fixations of the expert model were presented (for a detailed description of this issue see Section 2.3.2). The settings for the velocity-based saccade detection algorithm were set at a peak velocity threshold of 40°/s. The fixations were displayed as yellow circles with a line thickness of one pixel and a gaze trail for a temporal window of 1 s.

In the spotlight display condition, potentially distracting features in the unattended areas were filtered out. That is, the focus of the expert's attention (with a radius of 32 pixels) was visible as usual, whereas the areas surrounding it were ”blurred” by off-line foveation via video compression on non-attended areas (for a detailed description of the proceeding cf. Nyström, 2008; Nyström & Holmqvist, 2007). Table 7.1 shows a screen shot from each of the three conditions.

7.1.2.3 Tests

In order to estimate, whether participants were able to follow the attention guidance their eye movements were recorded during learning and the similarity of their eye movements to the eye movements of the model displayed on the learning video was calculated. Moreover, two different measures of learning were obtained: the efficiency of visual search during a test, and the ability to interpret important areas for the diagnosis of epileptic seizures. Visual search data were obtained during the test, in which participants were shown three new patient video cases for a mean duration of 29.67 s each (*SD* = 17.79) of different children displaying a type of seizure or spasm. The duration of the testing video depended on the duration of the seizure. While watching the videos, participants' eye movements were recorded. From those recordings two different eye movement measures were obtained, namely the time until participants looked at all relevant areas on the testing videos and the time spent on the relevant areas on the testing videos[1]. Afterwards, their interpretation performance was assessed by answering the following multiple-choice questions on those videos: (1) indicating from a list of body parts, which was moving suspiciously, (2) indicating from a list the type of the movement of these body parts, (3) indicating, whether touching the child would change the movement, (4) indicating, whether the face was involved and whether or not this was important for the diagnosis, (5) indicating the child's level of consciousness. The questionnaire was presented with the e-prime 2.0 software.

7.1.3 Procedure

The recording was conducted in individual sessions of approximately 30 min each. At the beginning, participants filled in a questionnaire on their prior knowledge in this task and their demographic data. Then, they received a short introduction to the topic, stating very general information on seizures and the importance to distinguish them from normal behavior. Then, the learning phase started. The eye tracking system was adjusted to the individual features of the participant based on a 13-point calibration. Participants were told that they would subsequently receive videos of the to-be-learned disease, where an expert explains the motion pattern visible in the video. Depending on the condition, they were told that they would additionally see where the expert's attention was attracted to on the video. Before watching the learning videos, participants received information on the age, gender, and a short problem description of the patient. While watching the corresponding learning video, the eye movements of the participants were tracked.

[1]The relevant areas determined a priori by two domain experts (cf. Antes & Kristjanson, 1991; Charness et al., 2001).

In the testing phase a fixation cross appeared for two seconds followed by the testing video, which was replayed once. Participants watched the video while their eye movements were recorded. Afterwards, the video disappeared, resulting in a blank screen. Then, the participants had to answer the multiple-choice questions as described in the section above. This procedure was repeated for the remaining two patient video cases in the test.

7.1.4 Data Analysis

7.1.4.1 Analysis of Eye Movement Data During Example Study

Similarity of participants' eye movements with the eye movements of the expert was analyzed by calculating the Euclidean distance between simultaneous gaze points of the expert and the participant over time (cf. Rao et al., 2002). Subsequently, the mean Euclidean distances were calculated for each participant (cf. chapter 5.1.4).

7.1.4.2 Analysis of Learning Outcomes

To analyze participants' learning outcomes in terms of their visual search skills, two eye movement measures were derived: (1) time until first having looked at all relevant areas of interest (AOI) and (2) time spent on relevant AOIs (total dwell time). Each video included several AOIs that were not redundant so that each of them had to be looked at in order to describe the infant's movement pattern exhaustively. In order to obtain the time for visually searching all of the relevant AOIs only the time that participants spent ”outside” of any relevant AOI until finally each AOI was looked at was taken into account. Thus, time until looking at *all* relevant areas was calculated by the time t_1 when all relevant areas have been looked at minus the dwell time on relevant AOIs before t_1. Since the testing videos were of different durations both eye tracking measures were normalized by the video length. In that, each value for each participant was divided by the value for the length of the respective video[2]. The relevant AOIs were determined a priori by two domain experts in the field of study.

The construction and scoring of the learning outcomes in terms of interpretation performance (as assessed by the multiple-choice questionnaire) was informed by the task analysis and by the help of domain experts. Accordingly, to diagnose an epileptic seizure, a partial credit of one point was assigned for each of the following solution steps: (1) correctly stating which part of the body is involved in the movement, (2) correctly stating how this part moves, (3) correctly stating, whether touching the child would change the movement, (4) correctly stating, whether the face was involved and whether or not this is important for the diagnosis, (5) correctly stating the child's level of consciousness. Hence, participants could receive a maximum of five points per video (1 point for each solution step).

7.2 Results

For all statistical tests reported here, a significance level of .05 is used. Means and standard deviations for each condition are given in Appendix Study 3, Table 4.

7.2.1 Eye Movements During Example Study

An ANOVA was conducted with mean Euclidean distance as dependent variable and condition (control vs. circle display vs. spotlight display) as independent variable. The comparison of participants' eye movement data during example study with the expert's eye movement data, showed a main effect of condition, $F(2, 57) = 3.60, p = .03, \eta_p^2 = .11$. Bonferroni post-hoc tests revealed that the spotlight display condition guided the students' attention marginally more successful than in the circle display condition ($p = .07$) and in the control condition ($p = .08$). Circle condition and control condition did not differ significantly in

[2]For instance, a participant looked at a relevant AOI A for 200 ms in video A which is 400 ms long. If the same participant looked at a relevant AOI B for 300 ms in video B which is 1200 ms long, this value needs to be normalized to be comparable to the value from video A (i.e., in this case divided by 4). Hence, the resulting normalized value is lower, namely 100 ms, even if the absolute value is higher.

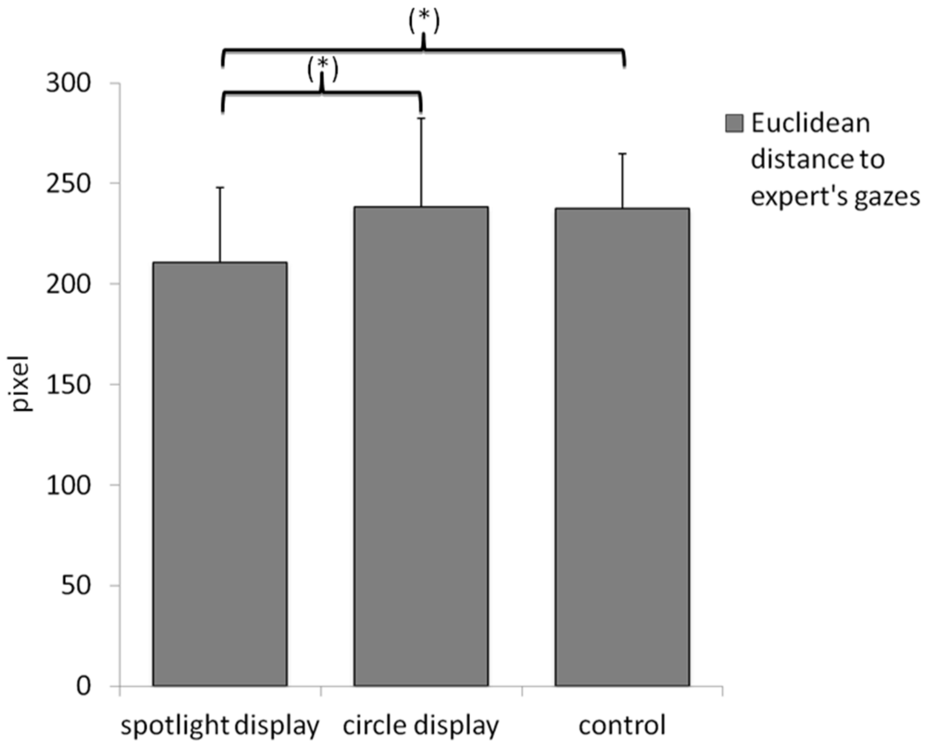


Figure 7.2: Means and standard deviations for Euclidean distance (in pixels) between expert's and students' gaze points over time during example study.

their potential to guide attention, ns. Results are summarized in Figure 7.2.

7.2.2 Learning Outcomes

7.2.2.1 Eye Movement Data During Testing

An ANOVA was conducted with time before first having looked at all relevant AOIs while watching the test videos as dependent variable and condition (control vs. circle display vs. spotlight display) as independent variable. The conditions significantly differed in time participants took before having looked at all relevant AOIs for the first time, $F(2, 57) = 3.41, p = .04, \eta_p^2 = .11$. Bonferroni post-hoc tests indicated that the spotlight display group has looked marginally earlier at all relevant AOIs than the control group, $p = .06$. The control condition and the circle condition did not differ significantly from each other, ns. Results are summarized in Figure 7.3.

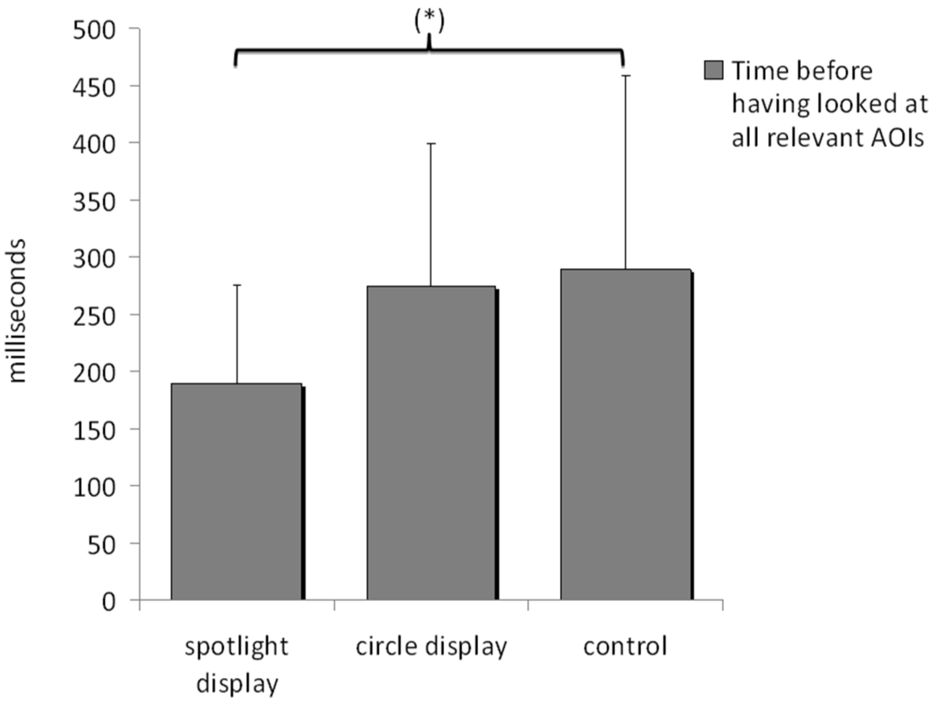


Figure 7.3: Means and standard deviations for time before having looked at all relevant AOIs (in ms) during testing.

An ANOVA with dwell time on relevant AOIs as dependent variable and condition (control vs. circle display vs. spotlight display) as independent variable showed significant differences between conditions in dwell time on relevant AOIs, $F(2, 57) = 3.64, p = .03, \eta_p^2 = .11$. Bonferroni post-hoc tests indicated that the spotlight display group looked marginally longer on the relevant AOIs than the control group, $p = .07$, and than the circle group, $p = .07$. The circle and the control group did not differ significantly, ns. Results are summarized in Figure 7.4.

7.2.2.2 Multiple-Choice Questionnaire

An ANOVA with correctness scores in the multiple-choice questionnaire as dependent variable and condition (control vs. circle display vs. spotlight display) as independent variable showed significant differences between conditions in interpretation performance, $F(2, 57) = 4.71, p < .01, \eta_p^2 = .14$. Bonferroni post-hoc tests indicated that the spotlight display condition outperformed the control condition, $p = .01$, but did not differ significantly from the circle display condition, $p = .11$. The control condition and the circle condition did not differ significantly from each other, ns. Results are shown in Figure 7.5.

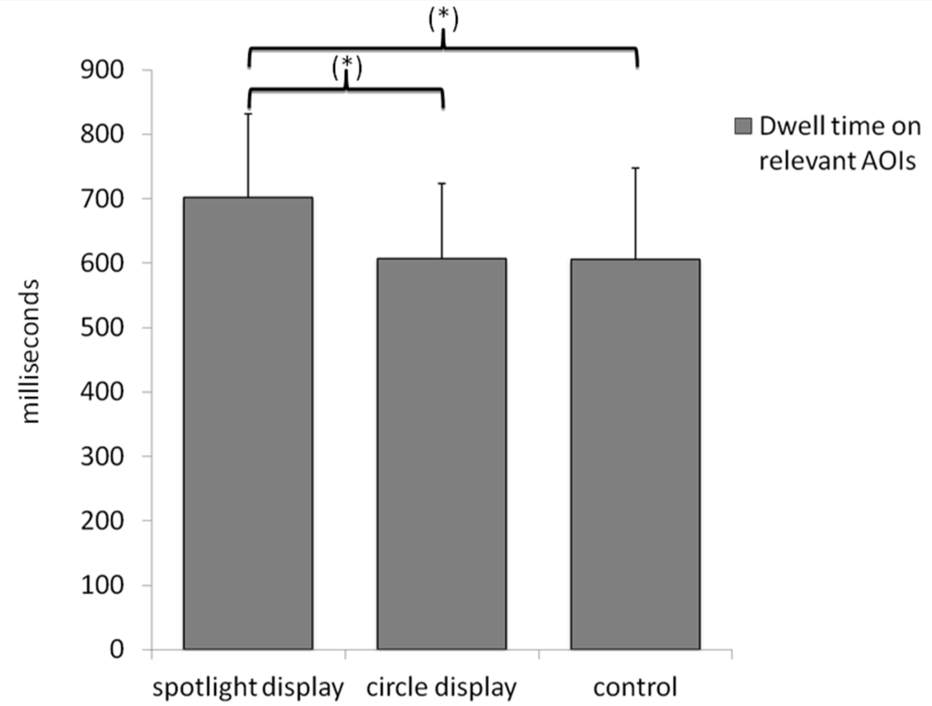


Figure 7.4: Means and standard deviations for dwell time on relevant AOIs (in ms) during testing.

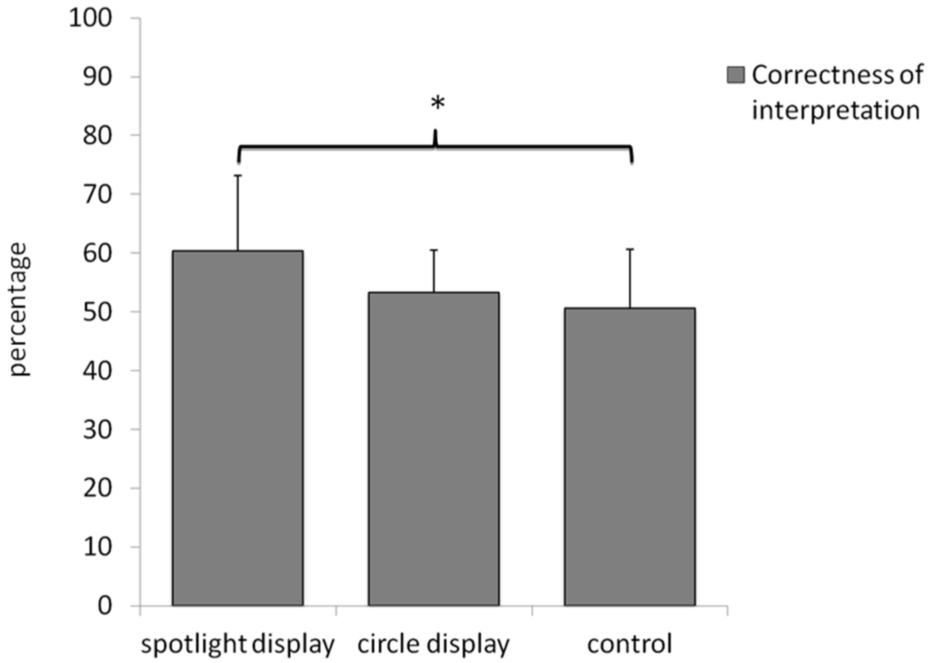


Figure 7.5: Means and standard deviations for correctness of interpretation score (in percent) for the test videos.

7.3 Discussion

The third study aimed at teaching perceptual skills necessary for diagnosing epileptic seizures in infants according to semiological terms. Based on the findings from Study 2 this was again done by EMME. This time, however, an expert explained symptoms of epileptic seizures on PVCs. Moreover, two improved ways of displaying the experts' eye movements were compared: a manufacturer-provided circle display (SMI) and an improved spotlight display (Nyström, 2008; Nyström & Holmqvist, 2007). Since Study 2 revealed that displaying the eye movements of an expert in example-based learning guides attention, I hypothesized to replicate this effect, with an equally strong attention guidance for both EMME displays, and a less strong one for the control condition (Hypothesis 1). Moreover, I hypothesized that EMME would foster learning of perceptual skills not only in a biological classification task, but also in medical diagnosis. Since both EMME displays were improved, I assumed them to enhance learning, in terms of a more efficient visual search on the videos (Hypothesis 2) and in terms of a better interpretation performance (Hypothesis 3) in comparison to the control condition.

Results of this study showed that learners' attention can be guided by displaying expert's eye movements in PVCs. Analyses of participants' eye movements during example study indicated that the spotlight display of the expert's eye movements guided students' attention during learning stronger than the circle display or no attention guidance. Moreover, the results showed that this type of attention guidance fostered learning, both in terms of improved visual search (as measured by means of eye tracking) and interpretation ability (as measured by means of multiple choice questionnaires). Participants in the spotlight display group outperformed both, the circle display group and the control condition with regard to both measures. Surprisingly, the circle display condition did not differ from the control group in terms of neither attention guidance, nor learning.

The reason for the failure of the circle group might lie in the fact that in this condition the expert model's perceptual processes were presented as *additional* information. Learning from representations that include a large amount of information require a large amount of students' cognitive capacities (Sweller et al., 1998). Accordingly, if more information is included this might impose additional load on students' cognitive capacities. Overloading learners' cognitive capacities, however, has shown to hamper learning (Sweller et al., 1998). As shown in Section 6.1.1.3 the PVCs used in this study are very demanding. Hence, it is not unlikely that learning with those PVCs reaches the limits of novice students' cognitive capacities. If one adds additional information to these videos, it might occur that this information exceeds students' cognitive capacities, which in turn, would explain why participants in this group would not learn better than in the control group. The spotlight display, in contrast, *reduced* the amount of presented information during learning. In that, the eye movements of the expert model are not added onto the video, but the information provided with the video is reduced according to the expert model's eye movements (i.e., the information provided in areas the expert model did not look at was reduced by blurring it). This manipulation seems to free cognitive capacities of the learners. These freed capacities in turn seem to be used for cognitive activities supporting learning.

These results show that guiding students' attention by means of EMME may foster acquisition of perceptual skills in medical diagnosis of epileptic seizures in infants, if implemented in a spotlight manner. Additionally, this study underlines the need to explicitly teach perceptual skills in instruction on medical diagnosis based on PVCs. EMME seem to provide an appropriate instructional format to address this need, provided that the way of displaying eye movements is optimized with regard to the processing demands imposed onto the learner.

Chapter 8

General Discussion

8.1 Summary and Overview of the Findings of this Thesis

The main research question of this thesis was how to convey perceptual skills in biological and medical domains. This research question was addressed with two examples, one within each domain: classifying fish locomotion patterns in biology and diagnosing epileptic seizures in medicine. The thesis is divided in two parts: the first part introduces task analysis as a framework for instructional design. The second part of this thesis deals with teaching perceptual skills.

8.1.1 Task Analyses Revealed the Importance of Perceptual Skills

In the first part of this thesis, a task analysis was performed for the task of classifying fish locomotion patterns. The task analysis was conducted for the purpose to develop an appropriate instructional method to teach this task in Part 2. First, a normative task analysis based on textbooks from the domain of marine zoology and with the help of domain experts identified how this task *should* be performed by analyzing the task procedure, the knowledge and skills required to perform this task, and difficulties likely to occur in this task. In general, perceptual skills occurred to be most important for a successful task performance.

Second, an empirical task analysis was conducted based on a review of existing literature in order to reveal how this task actually *would* be performed. This review revealed that the existing literature would not allow to perform an exhaustive empirical task analysis. The overview of available of empirical findings showed that studies investigating perceptual processes directly via eye tracking with dynamic stimuli are scarce and that experts' strategies were hardly investigated on a perceptual level so far. Thus, Study 1 was conducted to analyze the task of classifying fish locomotion pattern in detail from an empirical point of view.

Study 1 compared experts and novices while classifying fish locomotion patterns based on four example videos. In this study, not only task performance, but also the processes underlying task performance were investigated. Cognitive processes were investigated via cued retrospective reports, while perceptual processes were investigated via eye tracking. In line with Hypothesis 1, results showed that experts outperformed novices in this task by performing faster and more accurate. This finding might have been in part due to the fact that experts used shortcuts in their task performance, confirming Hypothesis 3. Investigating the processes underlying the task performance revealed that experts attended more to relevant parts of the fish body, while novices got distracted by salient, but irrelevant features, which confirmed Hypothesis 5. These results were found on both, a cognitive as well as a perceptual level. The fact that novices' attention was rather guided by saliency than relevance resulted in more homogenous perceptual strategies in comparison to experts, who were rather guided by what occurred to be relevant according to their individual learning history as predicted by Hypothesis 4. Hypothesis 2, however, was confirmed only in part. While experts used more technical terms as hypothesized, they did not verbalize less then novices. This may be due to two reasons: either experts do not have verbalization difficulties in this task at all – which is not very likely given a large body of literature in favor of this assumption (e.g., Chi, 2006a; Speelman, 1998) – or the cued retrospective reporting method helped experts to overcome their difficulties to verbalize. Hence, the later reason is more likely.

The consequences drawn for instructional design were as follows. Since novices performed poor on this task, they obviously require training. In particular, novices get distracted by irrelevant elements, do not attend to relevant body parts, and cannot describe their motion correctly. Thus, novices require training on their perceptual skills. Experts, on the other hand, outperform novices. Thus, their task approach is in principle useful to draw conclusions on instructional design upon. However, experts use shortcuts and technical terms. This, however, is not useful for an initial training for novices. Thus, while the original task approach of experts seems not to be useful directly as a basis for instructing novices it might nevertheless be possible to instruct experts to behave in a more didactic manner before their approach is conveyed to novices. Finally, as experts have diverse strategies in approaching the task, it is more useful to use one carefully chosen expert for instructional purposes than a mean expert approach.

8.1.2 EMME Conveyed Perceptual Skills

8.1.2.1 Example 1: Fish Locomotion Classification in Marine Zoology

Part 2 of this thesis addressed the issue of how to teach perceptual skills. Current instructional approaches for conveying knowledge on motion patterns in marine zoology do not focus on perceptual skills. Due to the lack of existing methods to convey perceptual skills, I tried to adapt well-known methods for the early acquisition of cognitive skills: classical worked examples (Atkinson et al., 2000; Sweller et al., 1998), process-oriented worked examples: (Van Gog et al., 2004), and cognitive modeling (Collins et al., 1989). As a result of these considerations and of the prior executed task analysis, in collaboration with other colleagues I developed the EMME approach (eye movement modeling examples Van Gog et al., 2009). EMME are the recordings of an expert model, whose verbalizations (i.e., cognitive processes) and eye movements (i.e., perceptual processes) are recorded while s/he performs a task. These recordings are replayed to the learner by superimposing them on a video displaying an example problem (i.e., to-be-classified swimming fish or to-be-diagnosed moving infant) solved by an expert.

Contrary to the expectations, a first implementation of EMME for a Tower-of-Hanoi-like problem solving task – which is not part of this thesis – even hampered learning (Van

Gog et al., 2009). In retrospect, it can be said that the task and its requirements did not fit the EMME approach, that the expert serving as a model might have not been instructed enough to behave in a didactic manner, and that the EMME design was not in line with the considerations outlined in this thesis. Based on this study, I assumed these three issues to be important to reassure before implementing EMME. First, EMME might only be suitable for teaching tasks that strongly depend on perceptual skills. Classifying fish locomotion patterns requires the examination of a perceptual input, hence, it is a perceptual task (Chase & Chi, 1981; Chi, 2006a). Moreover, the information given in the task is not sufficient to solve it, instead prior-knowledge and training is needed, hence, classifying fish locomotion patterns is a knowledge-rich task (Van Lehn, 1989). Second, the expert model probably needs to be instructed more thoroughly to behave didactical instead of displaying original expert behavior. This can be achieved by ensuing that the expert focuses on the novice recipient instead on the task content (cf. Jucks et al., 2007) and by reducing the temporal gap between looking at an area and describing it verbally (cf. Richardson & Dale, 2005). Third, the type of display used to present the perceptual processes of the expert model to the learner might be crucial. It might be adverse to present experts' eye movements as additional information superimposed onto the video. Presenting eye movements by blurring non-attended areas, instead, might be better suited as it reduces the overall amount of presented information (Dorr et al., 2010). This assumption was investigated in Study 2.

Study 2 compared the influence of EMME on learning perceptual skills. Participants received example videos of a swimming fish verbally explained by an expert (control group). The EMME groups additionally received a presentation of the expert model's attentional focus: while a dot group saw a yellow dot on the areas the expert model attended to (information added), a spotlight group saw the areas attended by the expert unaltered, whereas the remaining of the video was blurred (reduced amount of information). Results showed that – in line with Hypothesis 1 – that participants learning with the spotlight display followed the eye movements of the expert model most closely compared to the dot display group (captured via eye tracking), who followed the expert model's eye movements more closely than the control group. Learning of perceptual skills was assessed by measuring the ability to visually search the fish's relevant body parts on new test videos (captured via eye tracking) and the ability to interpret the movement of these body parts. In line with Hypothesis 2, the participants in the spotlight display group looked faster and longer on the relevant body parts, compared to the other two groups. The difference between both EMME groups is most likely due to the fact that the spotlight display guides the learners' attention directly to the model's focus of attention, while the dot display group guides the learners' attention slightly around the model's focus of attention. Thus, the spotlight display group follows the expert model's eye movements more closely, which allows them to more efficiently visually search relevant body parts on new videos. Hypothesis 3 postulated a positive effect of EMME on the ability to interpret the observations of the video more directly. This was confirmed only for the dot display group, which outperformed the other two groups. This result might be due to an unexpected ceiling effect in this study, which might have resulted in little variance of the learners' outcomes and thus, might have overshadowed a possible EMME effect. On the other hand, this difference between both EMME groups may also be due to the fact that participants studying the dot display saw the entire, visually complex material. Although this might have enticed them to look away from the expert's focus once in a while, it might have allowed them to perceive the relevant body parts together with their context and thus, to probably better understand their motion. In contrast, the spotlight display might have obscured too much of the rest of the video, and thus, participants studying these examples might have seen only the area attended by the expert model, but not its context, which in turn might have hampered their understanding of the motion pattern. If this interpretation of the results is true, it might be better to manufacture a less intrusive spotlight display. Another reason might have been that the interpretation of the fish locomotion pattern was too easy, because all three groups scored between 70% and 80% correct. This ceiling effect might have caused a loss in variance between groups. Thus, EMME mighty not unfold their full potential in too easy tasks.

In sum, EMME have shown to be suited to teach perceptual skills in the task to classify fish locomotion in contrast to video examples that included a verbal explanation only without attentional guidance. Nevertheless, neither of the presentation types (dot or spotlight) of the expert model's eye movements stands out and was able to improve both measures of learning (visual search and interpretation of motion pattern) simultaneously.

8.1.2.2 Improvement of the Implementation of EMME for Teaching Perceptual Skills

As stated above, the first implementation of EMME (cf. Section 4.2.5) revealed that three issues need to be taken into account for a successful implementation of the EMME approach: the requirements of the to-be-learnt task, the nature of the model, and the presentation of the model's perceptual processes. Study 3 investigated an implementation of EMME that was carefully designed with regard to theses aspects.

Task. Section 4.2.5 stressed the importance of the matching between the task and the EMME approach; the use of EMME may only be helpful for tasks that are knowledge-rich and heavily depend on perceptual skills. In Study 3 the task to diagnose epileptic seizures in infants was used, which is both a knowledge-rich and a strongly perceptual task. This task is assumed to fit better to the EMME approach then the task to classify fish locomotion patterns. Although the videos of swimming fish included many potentially distracting elements, the task itself was rather simple: the relevant body parts were always salient, because they were moving (although not all salient areas were also relevant) and the interpretation of the relevant body parts was a dichotomous decision (body parts of fish may move in only two ways: undulating or oscillating). The effect of EMME is assumed to be stronger in the epileptic seizure task, which is not only more difficult from a perceptual perspective, but also has a more difficult underlying task procedure as the task analysis in Section 6.1 revealed. Hence, from this perspective EMME might be an appropriate instructional method to convey perceptual skills for diagnosing epileptic seizures in infants.

Model. As described in Section 4.2.3, the EMME expert model should be a domain as well as a teaching expert. Accordingly, in Study 3 the chosen expert was a physician of pediatric neurology with extensive experience in diagnosing epileptic seizures. Moreover, he has extensive experience in teaching pediatric neurology at a university level. Thus, the model knows from his own experience which aspects of this task provide difficulties for novices. Additionally, the expert was instructed to behave in a didactical manner. The instruction was in line with the instruction for the expert model in the fish locomotion classification task (see Section 4.2.3).

Presentation of Eye Movements. Study 2 compared a control group to two different types of presenting the eye movements of the expert model to the learners: a spotlight display, where non-attended areas were blurred, and a dot display, where attended areas were indicated by a yellow dot. Results of Study 2 indicated that the spotlight groups followed the expert model's eye movements more closely during learning and showed a more efficient visual search during testing than the control group. The dot display group, on the other hand, showed a better interpretation of the locomotion pattern than the control group. Taken together, although the use of EMME seems to have a positive effect on the acquisition of perceptual skills, none of the presentation types was superior in total. Thus, in Study 3 both presentation types were modified slightly in order to improve their effect on learning.

Although the *spotlight display* was successful to guide learners' visual attention and positively affected visual search, this was not reflected in an improved description of the locomotion pattern displayed in the test videos. As described above (see Section 5.3), this result might be due to the fact that the spotlight strongly occluded irrelevant information during training rendering it impossible to gain a holistic overview of the overall motion pattern. Thus, the spotlight display was designed in a less intrusive way in Study 3, so that it guides in a far more subtle way and at the same time preserves the impression of the context (cf. Nyström & Holmqvist, 2008). This was done in Study 3 by the so called "foveation" technique (Nyström, 2008). In that, the spatial contrast is lowered on unattended areas, while color and temporal contrast remain unaltered. So that, the overall motion is not blurred. As a consequence, a holistic overview over the entire scene – including motion – is possible.

In the *dot display* condition participants could interpret their observations better than the control group, however, their visual search was not improved compared to the control group. This might be due to the fact that the solid dot occluded the exact spot the expert model looked at. Thus, participants learning with the dot display were guided ”around” the relevant areas during learning instead being guided directly on the spot the expert model looked at. Hence, they might have had problems to correctly locate those areas quickly during testing. This would be even more severe in the task of describing epileptic seizures, because the relevant areas in this task are sometimes very small (e.g., an eye) and would be fully occluded by a dot. Thus, the dot display, was improved in Study 3 by displaying gazes not as a solid dot that enlarges and contracts, but instead as a transparent dot that does not change in size (i.e., a circle). Thus, in Study 3 a circle display was used instead of a dot display.

8.1.2.3 Example 2: Diagnosis of Epileptic Seizures in Infants

In Study 3 EMME were implemented to convey perceptual skills in medical education. Before the actual implementation, a task analysis had to confirm that EMME would suit the instructional goals in this task. As in the first example task, the task to diagnose epileptic seizures in infants was analyzed, first, in terms of a normative task analysis and second, based on empirical findings from expertise research. The normative task analysis revealed that this task is knowledge-rich and sufficiently depending on perceptual skills. In comparison to the fish locomotion classification task, the task to diagnose epileptic seizures in infants is even more difficult on a perceptual level (the dynamism is more complex) and on a procedural level (more steps need to be conducted to perform this task). The empirical task analysis revealed that novices are likely to lack the perceptual skills necessary for this task, which should lead to difficulties to perform it. Thus, it is assumed that novices require training of the perceptual skills necessary for this task. Furthermore, the empirical task analysis revealed that experts would outperform novices in this task. Thus, designing instruction based on experts’ task procedure seems appropriate. However, experts are also likely to use specific knwoeldge-based shortcuts in this task. These shortcuts would be not understandable for novices and thus not useful for instructional purposes. Hence, experts need to be instructed to avoid using shortcuts, before their task performance can serve as a model for novices. Altogether, EMME seem to fit the requirements of the task to diagnose epileptic seizures in infants very well. Since all empirical questions important for the instructional use of EMME in this task could be answered by means of the existing empirical literature, it was not necessary to perform an own study on expertise differences in the given task.

Study 3 investigated the improved implementation of EMME for this task. For the above stated reasons, the expert model was instructed to behave didactical as done in Study 2. Moreover, the presentation of the expert model's eye movements was altered. Following the results from Study 2 that indicated that the dot display might have guided the eye movements of the learners' slightly around the relevant areas, the dot was changed into a circle allowing the learners' to look exactly at the relevant spots. Moreover, the spotlight display was altered according to the findings from Study 2. Since this spotlight seemed to occlude too much context information, so that an overview over the moving body part in its context was not possible, the spotlight used in Study 3 was designed less intrusive, by blurring only contrast and not color or motion (Nyström & Holmqvist, 2008). Hypothesis 1 stated that due to this improvement of visibility of the relevant areas for the circle condition, both EMME groups should guide the learners' attention equally well. Hypothesis 2 assumed, that the learners' visual search should also be improved in both EMME groups. Hypothesis 3, finally, stated that since the improved spotlight display enabled a holistic overview over the motion, both EMME groups should outperform the control group with regard to the ability to interpret test videos. Results showed, however, that all three hypotheses were only true for the spotlight display, while the circle display group did not differ from the control group in any respect. This leads to the question on why the two presentation types of EMME resulted in such different effects. It might be crucial that the circle display adds new information to the learning material, while the spotlight display reduces the overall amount of information of the learning material. In line with this reasoning, one possible explanation for these results might be that adding information to the already demanding task might have overloaded learners' cognitive capacities in the circle group and hampered learning. On the contrary, in the spotlight group, reducing irrelevant information during learning might have freed learners' cognitive capacities to use them for cognitive activities promoting learning. Another possible explanation is that the circle display might have tempted the learners' to try to "look through" the small circle instead of attending to the entire body part that the circle pointed at. Trying to look through a very small circle, however, might have been a very difficult visual search task that again overwhelmed learners' capacities. Furthermore, trying to look through the circle instead of attending to the body part the circle pointed at might have also hampered getting a holistic overview over the entire body part. It has been shown for both studies that details in displaying the eye movements are of importance for instructional efficiency. The next section compares the results of both studies and provides possible explanations for the – at a first glance – diverse findings.

8.1.2.4 Comparison of Findings from Both Examples

This section compares both applications of EMME in this thesis. Table 8.1 provides an overview over potentially influencing differences in both studies. Table 8.2 provides an overview over the findings of both studies. Both applications of EMME in this thesis – teaching perceptual skills to classify fish locomotion patterns in Study 2 and teaching perceptual skills to diagnose epileptic seizures in infants in Study 3 – differ according to three crucial issues: requirements of the task, presentation types of the expert model's perceptual processes, and the studied population. All three differences may have caused the discussed differences in the results.

As described in the normative analyses (see Section 3.1.3 and Section 6.1.1.3), both tasks are characterized by their perceptual and procedural difficulties. Both tasks are perceptually difficult, since both require the inspection of realistic and dynamic material. The task to diagnose epileptic seizures in infants, however, is characterized by an even higher perceptual difficulty than the task to classify fish locomotion patterns. This is in particular the case because relevant body parts appear and disappear (due to the movement of the video camera)[1], and because several relevant body parts move at the same time, which have to be interpreted separately (e.g., face and leg). Furthermore, these motions are more time-critical

[1]Although this type of dynamism might have also occurred in the fish locomotion classification task, this was not the case for the video examples used in Study 2.

then the motion in the fish locomotion task, as the infants' motion is transient, while the fishs' locomotion is re-occurring. Hence, the task to diagnose epileptic seizures in infants is perceptually more difficult than the task to classify fish locomotion patterns. In addition to a higher perceptual difficulty, the task to diagnose epileptic seizures is also more difficult with regard to its procedure. As described in Section 3.1.3.1 to classify fish locomotion patterns, three steps need to be carried out. To state a first diagnosis of an epileptic seizure, six steps need to be carried out as described in Section 6.1.1.2. Thus, the task to diagnose epileptic seizures in infants is also more difficult from a procedural perspective than the task to classify fish locomotion patterns.

Furthermore, both applications of EMME differed in the way they presented the expert model's perceptual processes. On the one hand, both studies presented perceptual processes as additional information. The study investigating the task to classify fish locomotion patterns used a solid dot, whereas the study investigating the task to diagnose epileptic seizures used a transparent circle. The dot slightly occluded the exact spot the expert model looked at, but at the same time it indicated which body part was looked at. The circle highlighted the exact spot the expert model looked at, instead of occluding it. This might have tempted the learners' to try to "look through" the small circle instead of attending to the entire body part. This, however, might have been difficult from a visual search perspective and it might have hampered learners to get an holistic overview over the entire body part. On the other hand, both studies presented perceptual processes by reducing the amount information on non-attended areas. Study 2 – fish locomotion – investigated a strong blurring of non-attended areas in the videos. In that, contrast, color, and motion were reduced. Although this type of presentation guided the learners' attention very well, it might have prevented them from gaining a holistic overview over the motion. Study 3 investigated a less strong blurring of non-attended areas in the videos. In that, only contrast was reduced. Thus, motion was still visible, which might have enabled learners to gain a holistic overview over the motion.

A third crucial difference between both studies was the studied population. While participants in the fish locomotion study (Study 2) were psychology students with no specific prior interest in the topic, the epileptic seizure study (Study 3) investigated a real "target

Table 8.1:
Comparison of Both Studies that Investigated EMME in this Thesis

	Fish locomotion	**Epileptic seizures**
Task features		
Perceptual difficulty	Moderate: realistic and dynamic	Higher: realism and more difficult dynamism
Procedural difficulty	Moderate: 3 steps to perform the task	Higher: 6 steps to perform the task
Presentation of perceptual processes		
As additional information	Dot occluding relevant areas	Circle highlighting relevant areas
By reducing irrelevant information	Spotlight blurring contrast, color, motion	Spotlight blurring contrast only
Studied population		
Participants	Psychology students without specific interest in this task	Medical students, highly motivated to learn this task

population". Thus, in Study 2 participants were very likely less motivated to learn this topic, whereas in Study 3 participants were actual medical students for whom this task was of importance. These students have to learn this topic in their later training anyhow, thus, they were likely to be more motivated to learn this task and to actively engage in cognitive activities conductive to learning.

Table 8.2 provides an overview over the effects of EMME found in both studies of this thesis. Some findings of Study 2 were replicated in Study 3. Results of both studies show that the spotlight display successfully guides learners' visual attention during learning to the spots the expert model looked at. Furthermore, both studies could show that the participants learning with the spotlight display execute an efficient visual search on novel testing videos. At the same time, both studies showed that presenting perceptual processes as additional information does not improve visual search in comparison to a control group without attention guidance.

Other findings differ between the two studies. First, the presentation of perceptual processes as additional information has shown to guide attention in Study 2 – although this effect was not as strong as for the spotlight display –, whereas no significant effect was found for Study 3. The most likely reason for this difference in results is the difference of the type of ”marker” on the spots the expert model looked at. As described above, a solid dot might guide the learners around and not precisely on the exact spot the expert model looked at. Nevertheless, this type of marker guided the learners' attention towards the entire relevant body part. A small circle, however, might tempt the learners to look through it. Trying to look through a small circle would be a difficult task in itself and it might hamper the recognition of the entire relevant body part. This, however, is only an assumption and cannot be decided based on the given studies.

Second, the interpretation ability resulting from the dot and the circle display group differed between both studies. Again, the way of presenting the processes as additional information might have caused the differences. Since the dot was easier to follow – as indicated by the results of the attention guidance during learning – learners might have had enough cognitive capacities left to focus on what the expert model was saying and thus, might have been able to interpret novel videos more correctly on their own. Participants in the third study learning with the circle display might had difficulties to follow the circle, which might have consumed too much cognitive capacity so that there was not enough capacity left to attend to the expert model's verbalizations and to learn from them. Again, this is an assumption, which cannot be decided based on the given studies. Another possible explanation would be that the task to diagnose epileptic seizures in infants is more difficult than the task to classify fish locomotion and thus, participants in the third study (circle display) scored worse than in the second study (dot display). This explanation is very unlikely, however, because participants in the spotlight condition in Study 3 scored better than in the spotlight condition in Study 2.

Table 8.2:
Overview Over Findings on the EMME Effect Depending on Task and Presentation Type of Perceptual Processes

	Fish locomotion		Epileptic seizures	
Presentation of perceptual processes	**Dot**	**Spotlight**	**Circle**	**Spotlight**
During learning				
Attention guidance	+	+	0	+
During testing				
Visual search	0	+	0	+
Interpretation ability	+	0	0	+

Note. Gray cells are replications of effects in both studies.

Third, the interpretation ability of the spotlight group differed between both studies. As just elaborated the assumption that the task in itself caused the difference is unlikely, because the interpretation ability in both studies did not change for both display types in the same way. Thus, it is more likely that the change in the display type caused the difference. Since the spotlight display in the third study did only blur spatial contrast, it guided the learners to the not-blurred areas, but still allowed for an overview over the motion and color of the entire object. This might have facilitated the understanding of the motion as explained by the expert model and thus, enabled learning. The spotlight display in the second study, on the other hand, blurred not only contrast, but also color and motion. Although this obviously guided the learners' visual attention very well, it also did occlude the motion itself so strongly that the learners might have experienced difficulties to relate the perceptual input to the explanation of the expert model, which in turn might have hampered learning. This, however, is only a post-hoc explanation that cannot be finally decided based on the given studies.

In sum, this thesis showed that EMME may be helpful to better convey perceptual skills in the domains of classifying fish locomotion patterns and diagnosing epileptic seizures in infants. The effects of EMME, however, might strongly depend on the task under investigation, the studied population, and the used presentation of the expert model's perceptual processes. The studies in the given thesis suggest that the best effect – in terms of attention guidance during learning as well as visual search and interpretation performance during testing – might be achieved by an EMME that displays the perceptual processes of the expert model as a spotlight that blurs the contrast of the areas the expert model did not attend to, while the attended areas remain unaltered.

8.2 Evaluation of the Studies in this Thesis

All three studies of this thesis have strengths and drawbacks. These need to be taken into account when interpreting the results and drawing conclusions from them. This section, thus, discusses the strengths and weaknesses of those studies.

8.2.1 Strengths of the Studies Conducted in this Thesis

In all three studies, the tasks, the items, and their scoring were chosen based on thorough task analyses and were approved with the help of domain experts that had vast teaching experience. Hence, it can be assumed that the tasks that were investigated in this thesis are of relevance for educational practice in that they were similar to real-life tasks, and in that the performance of participants' on these tasks was scored in an appropriate way. These advantages were ensured by the following precautions.

First, all three studies used realistic videos. These videos were taken from large databases containing diverse videos. To avoid biases in terms of choosing inappropriate videos that might contain artifacts, the videos used for all three studies were chosen with the help of two domain experts, who approved and evaluated each of those as being a prototypical example for the to-be-learned motion. Moreover, the videos as well as the task to diagnose or classify them, as used in this thesis, were all highly realistic in terms of their similarity to a real-world task. Second, not only the videos, but also their analysis in terms of eye tracking was performed with the help of two domain experts per task. In that, two domain experts agreed upon which areas in the videos were defined as relevant AOIs. Thus, the decision whether or not participants performed an efficient visual search in terms of looking fast enough and sufficiently long on relevant body parts was made by experts of the domains and not by the researcher only. Third, Study 2 and Study 3 assessed participants' ability to interpret motion patterns displayed in the videos via a multiple-choice questionnaire. To make sure that these questionnaires are not biased by the hypotheses of the researcher, again domain experts determined and approved a priori to the experiments all questions and answer alternatives. Fourth, Study 1 recorded verbal data of the participants. Again, the coding schema for the verbal data was approved by two domain experts. Moreover, the actual coding was performed by two independent raters, who showed complete agreement.

8.2.2 Weaknesses of the Studies in this Thesis and their Possible Consequences

Besides these advantages, the studies in this thesis have also weaknesses. These weaknesses mainly affect the testing items used and the population tested as stressed in the following.

Besides the multiple choice questionnaire in Study 2 each dependent variable in each study was based on only two to four testing videos. This is due to the fact that each of those variables required very time consuming data preparation of either verbal or eye tracking data. The verbal data recorded in Study 1 required not only transcription of a protocol per participant and per video, but also the coding of it by two independent raters. Thus, only one video was presented per investigated fish locomotion pattern resulting in only four videos. Moreover, all videos used in Study 1 and the test videos used in Study 2 and Study 3 had to be prepared in a very time consuming way in order to create ”dynamic” AOIs (for a detailed description see Section 3.2.2.4). Thus, each study consisted of only three to four such videos. Finally, to capture the distance between the expert model's and the learners' gazes in Study 2 and Study 3 required a very time consuming data preparation via Excel per participant and per video. Thus, this dependent variable was based only on two to three videos. In sum, the low number of videos is due to the fact that the dependent variables calculated in this thesis are not part of any manufacturer based or open source data preparation software and required very time consuming manual preparations. This in turn leads to a low reliability (Nunnally, 1967). If these preparations could be automized in future, such recordings should include far more videos to be able to achieve a high reliability and thus more statistical power. Another problematic issue within these studies is that the used stimuli were realistic videos that – although carefully chosen with the help of domain experts – might still have included disturbing elements that are not representative of the according motion pattern. Again, a larger number of videos would help to reduce this danger. A final problem with the testing material occurs in Study 2. Since no standardized tests to estimate perceptual skills in the two domains of this thesis exist, self developed tests were used. As mentioned above, in Study 2 a ceiling effect might have occurred, because all participants scored very high. Thus, the used type of testing might have been too easy for this task. An alternative would be, instead of presenting the learners a multiple choice questionnaire that only requires the learners to recognize the answers, to ask the learners to freely describe the locomotion patterns. This might have led to more variance within the data, which in turn might have also have strengthen the results. In sum, the testing and learning materials might be improved in future studies.

Accordingly, the findings from this thesis might suffer at least from two issues concerning the investigated population. First, the small number of participants in Study 1 is problematic. This is, however, not uncommon in expertise research as experts are scarce. There is a relatively small number of specialists in marine zoology. Still, because of the small sample size, care has to be taken in interpreting the results from this study. Moreover, the statistical power of the results is very likely to be reduced by this fact. Second, Study 2 investigated not a realistic target group, but instead psychology students with no particular interest in the topic were investigated. Again, this is very common in psychological research, because such participants are easy to access. Nevertheless, the question is whether their potentially low motivation to learn this topic might have influenced the results. Furthermore, it is not clear, whether the results may be generalized to the real target group, namely biological students. This issue was ruled out in Study 3 by investigating only medical students to whom this topic was actually important.

8.3 Implications of the Findings from this Thesis for Research and Practice

8.3.1 Practical Implications

Despite the diversity of findings in this thesis, it has to be noted that EMME used in this thesis were an effective instructional approach, if implemented in a certain manner. Therefore, the question arises on how to transfer the findings of this thesis into practice. What we have clearly seen is that some tasks (here classifying fish locomotion and diagnosing epileptic seizures in infants) require perceptual skills, which are not in focus of the teaching so far. The same is probably true for several other tasks that require substantial perceptual skills as well, like sports judging, playing computer games, air traffic control, car driving, etc. When instructors are teaching novices to execute this kind of tasks, they should keep in mind to teach perceptual skills as well. EMME might be an appropriate way to do this for many tasks. Of course, it is not always possible to use eye movement recordings of experts, because eye trackers are not easily accessible. Still, if a teacher is aware of the fact that perceptual skills play a crucial role in the task s/he is teaching, s/he could use videos of the task, point as detailed as possible to the relevant areas, and explain how they should be interpreted. Furthermore, EMME could be used in online teaching where pointing is not possible (e.g., during a surgery) using an eye tracking device from which the novice receives the recordings. This proceeding might help a lot to learn a new perceptual task quickly. Again, these recordings could be later re-used for offline teaching situations.

8.3.2 Implications for Further Research

This thesis presented a first answer to the question on how to teach perceptual skills. Nevertheless, this thesis also raised several new questions. These new questions concern three levels: the task itself, the model, and the presentation type of the model's perceptual processes.

The two *tasks* used in the current thesis have many characteristics in common, like including the examination of a realistic and dynamic stimuli. Nevertheless, these tasks differ according to the difficulty of dynamism they involve and the difficulty of their procedure. Although I suggested above that the differences in the findings between those two studies are most likely to be due to the differences in presentation types used, it cannot be ruled out that the type of task influences the findings as well (in particular, since a ceiling effect occurred in Study 2). Thus, further research should investigate the effect of task types directly. This should be done by varying the levels of difficulty independently. For instance, the level of perceptual difficulty in dynamism might be crucial, because novices have a high risk to miss important information. The difficulty of realism, on the other hand, might not be the central issue. Rather the amount of information density could affect the need of attentional guidance. A high level of difficulty in the procedure in turn might require a stronger emphasis on conceptual pre-training. This kind of research would clarify the role of task characteristics on the need of perceptual skills that EMME might address.

An issue not varied in this thesis, but decided upon a priori, is the question on who should serve as a *model* and how this person should behave. Based on a literature review experts served as models in both tasks in this thesis. Moreover, these experts were instructed to behave didactical. Obviously, the effects of this instruction were not investigated empirically. Further research should investigate whether the instruction used really leads to a more didactical behavior of the experts. Another interesting research question is, whether an expert is really the best model. In the cognitive modeling approach, the model can also be a peer student with an equal or a slightly higher level of performance than the learner, in which case the demonstrated procedure may even contain errors (e.g., Braaksma et al., 2002;

Schunk & Hanson, 1985). Thus, further research should investigate the role of the model (in terms of level of prior knowledge and of instruction to behave more didactical) on the learning from EMME.

The differences between the two implementations of EMME in this thesis that is most likely to cause the difference in the findings is the *presentation type of the model's perceptual processes*. One assumption addressing the difference between the dot and the circle display is that participants seeing the dot looked at the element the dot ”pointed at”, while participants seeing the circle tried to ”look through” the circle, which should be more difficult and thus using up more cognitive capacity, which in turn would be missing for learning. Further research should investigate this in more detail. A detailed analysis of eye tracking data could distinguish whether participants look somewhere on the relevant element or whether they try to ”catch” the display with their gazes. Moreover, participants could be asked directly what they think the relevant part of the video is: the element or the small spot behind the display. Finally, measuring the amount of cognitive load caused by each display could give insight whether the initial assumption is true. The spotlight display could be investigated in more detail as well. Besides investigating the effects of blurring contrast, color, and motion separately, the degree of the blurring could be varied to find an optimal display type with least disturbance.

In sum, the current thesis showed that perceptual skills can be taught by the use of EMME. The different types of EMME used in this thesis, however, did not deliver consistent findings. The results indicate that the effect of EMME depends on the to-be-learned task, the used model, and the way in which the model's perceptual processes are presented to the learner. If further research can replicate these findings and clarify the open research questions, the EMME approach might be helpful in many teaching situations.

Bibliography

Acharya, J., Wyllie, E., Lüders, H., Kotagal, P., Lancman, M., & Coehlo, M. (1997). Seizure symptomatology in infants with localization-related epilepsy. *Neurology*, *48*, 189-196.

Ahmed, A. (2005). When is facial paralysis bell's palsy? Current diagnosis and treatment. *Clevelan Clinic Journal of Medicine*, *72*, 398-405.

Annett, J., & Duncan, K. D. (1967). Task analysis and training design. *Occupational Psychology*, *41*, 211-221.

Antes, J. R., & Kristjanson, A. F. (1991). Discriminating artists from nonartists by their eye-fixation patterns. *Perceptual and Motor Skills*, *73*, 893-894.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, *70*, 181-214.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, *95*, 774-783.

Ayres, P., & Paas, F. (2007). A cognitive load approach to the learning effectiveness of instructional animation. *Applied Cognitive Psychology*, *21*, 811-820.

Balslev, T., De Grave, W. S., Muijtjens, A. M. M., Eika, B., & Scherpbier, A. J. J. A. (2009). The development of shared cognition in paediatric residents analysing a patient video case versus a paper patient case. *Advances in Health Science Education*, *14*, 557-565.

Balslev, T., De Grave, W. S., Muijtjens, A. M. M., & Scherpbier, A. J. J. A. (2005). Comparison of text and video cases in a postgraduate problem-based learning format. *Medical Education*, *39*, 1086-1092.

Balslev, T., Jarodzka, H., Holmqvist, K., De Grave, W. S., Muijtjens, A., Eika, B., et al. (2011). *How do paediatricians diagnose? Visual attention and cognitive processes.* (Manuscript submitted for publication.)

Bandura, A. (1977). *Social learning theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Binet, A. (1893/1966). Mnemonic virtuosity: A study of chess players. *Genetic Psychology Monographs*, *74*, 127-162.

Blake, R. W. (2004). Fish functional design and swimming performance. *Journal of Fish Biology*, *65*, 1193-1222.

Blandin, Y., Lhuisset, L., & Proteau, L. (1999). Cognitive processes underlying observational learning of motor skills. *The Quarterly Journal of Experimental Psychology*, *52a*, 957-979.

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

Boekhout, P., Van Gog, T., Van De Wiel, M., Gerards-Last, D., & Geraets, J. (in press). Example-based learning: Effects of model expertise in relation to student expertise. *British Journal of Educational Psychology*, Retrieved from http://www.ingentaconnect.com/content/bpsoc/bjep/pre-prints/bjep871;jsessionid=4ruigm8b5msdh.alexandra.

Böhme, M., Dorr, M., Martinetz, T., & Barth, E. (2006). Gaze-contingent temporal filtering of video. In K.-J. Räihä & A. T. Duchowski (Eds.), *Proceedings of Eye Tracking Research & Applications* (p. 109-115). San Diego, California: ACM.

Bone, Q., & Moore, R. H. (2008). *Biology of fishes.* New York, NY: Taylor & Francis Group.

Boshuizen, H. P. A., & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates, and novices. *Cognitive Science*, *16*, 153-184.

Boucheix, J.-M., & Lowe, R. K. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction*, *20*, 155-166.

Braaksma, M. A. H., Rijlaarsdam, G., & Van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology*, *94*, 405-415.

Brooks, L. R., LeBlanc, V. R., & Norman, G. R. (2000). On the difficulty of noticing obvious features in patient appearance. *Psychological Science*, *11*, 112-117.

Bruner, J. (1967). *On knowing: Essays for the left hand.* Cambridge, Mass: Harvard University Press.

Bryman, A. (1984). The debate about quantitive and qualitative research: A question of method or epistemology? *British Journal of Sociology*, *35*, 75-92.

Burns, C., & Hajdukiewicz, J. R. (2004). *Ecological interface design*. Boca Raton, FL: CRC Press.

Burton, A. M., Shadbolt, N. R., Hedgecock, A. P., & Rugg, G. (1988). A formal evaluation of knowledge elicitation techniques for expert systems: Domain 1. In D. S. Moralee (Ed.), *Research and development in expert systems IV* (p. 136-145). Cambridge: Cambridge University Press.

Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction*, *20*, 155-166.

Catrambone, R. (1996). Generalizing solution procedures learned from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1020-1031.

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General*, *12*, 355-376.

Cavazos, J. E., & Spitz, M. (2009, November). *Seizures and epilepsy, overview and classification.* Paper retrieved from http://emedicine.medscape.com/article/1184846-overview.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*, 233-246.

Chandrasekaran, B. (1983). Towards a taxonomy of problem solving types. *AI Magazine*, *4*, 9-17.

Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 467-476.

Charness, N. (1991). Expertise in chess: The balance between knowledge and search. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (p. 39-63). New York: Cambridge University Press.

Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory and Cognition*, *29*, 1146-1152.

Chase, W. G., & Chi, M. T. H. (1981). Cognitive skill: Implications for spatial skill in large-scale environments. In J. Harvey (Ed.), *Cognition, social behavior, and the environment* (p. 111-136). Hillsdale, NJ: Erlbaum.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (p. 215-281). New York: Academic.

Chi, M. T. H. (2006a). Laboratory methods for assessing experts' and novices' knowledge. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The cambridge handbook of expertise and expert performance* (p. 167-184). Cambridge: Cambridge University Press.

Chi, M. T. H. (2006b). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, R. R. Hoffman, & P. J. Feltovich (Eds.), *The cambridge handbook of expertise and expert performance* (p. 21-30). Cambridge: University Press.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145-182.

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

Chow, R., & Vicente, K. J. (2002). *A field study of emergency ambulance dispatching: Implications for decision support.* Abstract retrieved from http://www.ingentaconnect.com/content/hfes/hfproc/2002/00000046/00000003/art00021. Santa Monica, CA: Human Factors and Ergonomics Society.

Clark, R. E., Feldon, D., Van Merriënboer, J. J. G. V., Yates, K., & Early, S. (2008). Cognitive task analysis. In J. M. Spector, J. J. G. Merrill, J. J. G. V. Van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (p. 577-594). Lawrence Erlbaum Associates.

Collins, A. F., Brown, J. S., & Newman, S. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. Resnick (Ed.), *Cognition and instruction: Issues and agendas* (p. 453-494). Mahwah, NJ: Erlbaum.

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies*, *41*, 801-849.

Cooper, G. A., & Sweller, J. (1987). The effects of schema acquisition and rule automationon mathematical problem-solving transfer. *Journal of Educational Psychology*, *79*, 347-362.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus driven attention in the brain. *Nature Review Neuroscience*, *3*, 201-215.

Couzijn, M. (1999). Learning to write by observation of writing and reading processes: Effects on learning and transfer. *Learning and Instruction*, *9*, 109-142.

Crandall, B., Klein, G. A., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. Cambridge, MA: MIT Press.

Crowley, R. S., Naus, G. J., Stewart, J., & Friedman, C. P. (2003). Development of visual diagnostic expertise in pathology - An information-processing study. *Journal of the American Medical Informatics Association*, *10*, 39-51.

Darabi, A. A., Nelson, D. W., & Palanki, S. (2007). Acquisition of troubleshooting skills in a computer simulation: Worked example vs. conventional problem solving instructional strategies. *Computers in Human Behavior*, *23*, 1809-1819.

De Groot, A. D. (1946/1978). *Thought and choice and chess.* The Hague, Netherlands: Mouton.

De Groot, A. D., & Gobet, F. (1996). *Perception and memory in chess.* Assen, The Netherlands: Van Gorcum.

De Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review*, *21*, 113-140.

De Leng, B. A., Dolmans, D. H. J. M., Van de Wiel, M., Muijtjens, A. M. M., & Van der Vleuten, C. P. M. (2007). How video cases should be used as authentic stimuli in problem-based medical education. *Medical Education*, *41*, 181-188.

Denzin, N. K. (1970). *The research act in sociology.* Chicago: Aldine.

Dequeker, J., & Jaspaert, R. (1998). Teaching problem-solving and clinical reasoning: 20 years experience with video-supported small-group learning. *Medical Education*, *32*, 384-389.

Diez, M., Boehm-Davis, D. A., Holt, R. W., Pinney, M. E., Hansberger, J. T., & Schoppek, W. (2001). *Tracking pilot interactions with flight management systems through eye movements.* Paper presented at the 11th International Symposium on Aviation Psychology. Columbus, Ohio. (Abstract retrieved from http://www.hf.faa.gov/docs/508/docs/gmugrant/papers/Diez01.pdf)

Dodge, R. (1903). Five types of eye movements in the horizontal meridian plane of the field of regard. *American Journal of Psychology*, *8*, 307-329.

Dodge, R., & Cline, T. (1901). The angle velocity of eye movements. *Psychological Review*, *8*, 145-157.

Dorr, M., Jarodzka, H., & Barth, E. (2010). Space-variant spatio-temporal filtering of video for gaze visualization and perceptual learning. In C. Morimoto & H. Instance (Eds.), *Proceedings of Eye Tracking Research & Applications ETRA '10* (p. 307-314). New York, NY: ACM.

Dorr, M., Vig, E., Gegenfurtner, K. R., Martinetz, T., & Barth, E. (2008). *Eye movement modelling and gaze guidance.* Paper presented at the Fourth International Workshop on Human-Computer Conversation. (Abstract retrieved from http://www.gazecom.eu/publications/dovigemaba08.pdf)

Dreifuss, F. E., & Nordli, D. (2000). Classification of epilepsy in childhood. In J. M. Pellock,

W. E. Dodson, & B. F. D. Bourgeois (Eds.), *Pediatric epilepdy: Diagnosis and therapy* (p. 69-80). New York: Demos Medical Publishing.

Drury, C. G., Paramore, B., Van Cott, H. P., Grey, S. M., & Corlett, E. N. (1987). Task analysis. In G. Salvendy (Ed.), *Handbook of human factors* (p. 371-401). New York: John Wiley & Sons.

Duchowski, A. T. (2003). *Eye tracking methodology: Theory and practice*. London: Springer-Verlag.

Duncker, K. (1945). On problem solving. *Psychological Monographs*, *58*, No. 270.

Dwyer, F. M. (1976). Adapting media attributes for effective learning. *Educational Technology*, *16*, 7-13.

Egger, J., Grossmann, G., & Auchterlonie, I. A. (2003). Benign sleep myoclonus in infancy mistaken for epilepsy. *British Medical Journal*, *326*, 975-976.

Emmerson, P., & Ross, H. (1985). Colour constancy with change of viewing distance under water. *Perception*, *14*, 349-358.

Ericsson, K. A. (1988). Concurrent verbal reports on reading and text comprehension. *Text*, *8*, 295-325.

Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The cambridge handbook of expertise and expert performance* (p. 223-241). Cambridge: Cambridge University Press.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaption to task constraints. *Annual Reviews in Psychology*, *47*, 273-305.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*, 215-251.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: Bradford books / MIT Press.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Ericsson, K. A., & Smith, J. (1991). Prospects and limits in the empirical study of expertise. In K. A. Ericsson & J. Smith (Eds.), *Towards a general theory of expertise: Prospects and limits* (p. 1-38). Cambridge, MA: Cambridge University Press.

Farrow, D., & Abernethy, B. (2002). Can anticipatory skills be learned through implicit video-based perceptual training? *Journal of Sports Sciences*, *20*, 471-485.

Farrow, D., Chivers, P., Hardingham, C., & Sacuse, S. (1998). The effect of video-based pereptual training on the tennis return of serve. *International Journal of Sport Psychology*, *29*, 231-242.

Feltovich, P. J., & Barrows, H. S. (1984). Issues of generality in medical problem solving. In H. G. Schmidt & M. L. D. Volder (Eds.), *Tutorials in problem-based learning: New directions in training for the health professions* (p. 128-142). Assen/Maastricht: Van Gorcum.

Feltovich, P. J., Johnson, P. E., Moller, J. H., & Swanson, D. B. (1984). LCS: The role and development of medical knowledge in diagnostic expertise. In W. J. Clancey & E. H. Shortliffe (Eds.), *Readings in medical artificial inteligence: The first decade* (p. 275-319). Reading, MA: Addison-Wesley Publishing Company.

Feusner, M., & Lukoff, B. (2008). Testing for statistically signifikant differences between groups of scan patterns. In S. N. Spencer (Ed.), *Proceeding of the 2008 Symposium on Eye Tracking Research & Applications* (p. 43-46). New York: ACM.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, *51*, 327-358.

French, K. E., Nevet, M. E., Spurgeon, J. H., Graham, K. C., Rink, J. E., & McPherson, S. L. (1996). Knowledge representation and problem solution in expert and novice youth baseball players. *Research Quarterly for Exercise and Sport*, *67*, 386-395.

Geiwitz, J., Klatsky, R. L., & McCloskey, B. P. (1988). *Knowledge acquisition techniques for expert systems: Conceptual and empirical comparison.* Fort Monmouth, NJ.

Gerjets, P., Imhof, B., Kühl, T., Pfeiffer, V. D. I., Scheiter, K., & Gemballa, S. (2010). Using static and dynamic visualizations to support the comprehension of complex dynamic phenomena in the natural sciences. In L. Verschaffel, E. De Corte, T. De Jong, & J. Elen (Eds.), *Use of external representations in reasoning and problem solving*

(p. 153-168). London: Routledge.

Gerjets, P., Scheiter, K., & Catrambone, R. (2004). Reducing cognitive load and fostering cognitive skill acquisition: Benefits of category-avoiding examples. In R. Alterman & D. Kirsch (Eds.), *Proceedings of the 25th Annual Cognitive Science Society* (p. 450-455). Mahwah, NJ: Erlbaum.

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development.* Oxford, UK: Oxford University Press.

Glaser, R., Lesgold, A., Lajoie, S., Eastman, R., Greenberg, L., Logan, D., et al. (1985). *Cognitive task analysis to enhance technical skills training and assessment.* (Tech. Rep.). Pittsburgh, PA: Learning Research and Development Center, University of Pittsburgh.

Gobet, F., & Charness, N. (2006). Expertise in chess. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The cambridge handbook of expertise and expert performance.* (p. 523-538). Cambridge: Cambridge University Press.

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, *31*, 1-40.

Gobet, F., & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science*, *24*, 651-682.

Gott, S. P., Parker-Hall, E., Pokorny, R. A., Dibble, E., & Glaser, R. (1993). A naturalistic study of transfer: Adaptive expertise in technical domains. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (p. 258-288). Norwood, NJ: Ablex.

Goulet, C., Bard, C., & Fleury, M. (1989). Expertise differences in preparing to return a tennis serve: A visual information processing approach. *Journal of Sport & Exercise Psychology*, *11*, 382-398.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, *14*, 462-466.

Green, A. J. F. (1998). *Using verbal protocols in language testing research: A handbook.* Cambridge, UK: Cambridge University Press.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*, 274-279.

Haider, H., & Frensch, P. A. (1999). Eye movevement during skill acquisition: More evidence for the information-reduction hypothesis. *Jornal of Experimental Psychology: Learning, Memory and Cognition*, *25*, 172-190.

Hamer, H., Wyllie, E., Lüders, H., Kotagal, P., & Acharya, J. (1999). Symptomatology of epileptic seizures in the first three years of life. *Epilepsia*, *40*, 837-844.

Hansen, J. K., & Balslev, T. (2009). Hand activities in infantile masturbation: A video analysis of 13 cases. *European Journal of Paediatric Neurology*, *13*, 508-510.

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

Helsen, W. F., & Starkes, J. L. (1999). A multidimensional approach to skilled perception and performance in sport. *Applied Cognitive Psychology*, *13*, 1-27.

Henderson, R. D., Smith, M. C., Podd, J., & Varela-Alvarez, H. (1995). A comparison of the four prominent user-based methods for evaluating the usability of computer software. *Ergonomics*, *39*, 2030-2044.

Hillstrom, A. P., & Chai, Y. C. (2006). Factors that guide or disrupt attentive visual processing. *Computers in Human Behavior*, *22*, 648-656.

Hinds, P. I. (1999). The curse of expertise: The effects of expertise and debiasing methods on predictions of novice performance. *Journal of Experimental Psychology: Applied*, *5*, 205-221.

Hinds, P. I., Patterson, M., & Pfeffer, J. (2001). Bothered by abstraction: The effect of expertise on knowledge transfer and subsequent novice performance. *Journal of Applied Psychology*, *86*, 1232-1243.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*, 99-107.

Hoar, W. S., Randall, D. J., & Conte, F. P. (1997). *Fish physiology*. San Diego, California:

Academic Press.

Höffler, T., & Leutner, D. (2007). Instructional animation versus static picture: A meta-analysis. *Learning and Instruction*, *17*, 722-738.

Hoffman, R. R. (1987). The problem of extracting the knowledge of experts from the perspective of experimental psychology. *AI Magazine*, *8*, 53-67.

Hoffman, R. R. (1995). Eliciting knowledge from experts: A methodological analysis. *Organizational Behavior and Human Decision Processes*, *62*, 129-158.

Hoffman, R. R., Coffey, J. W., & Ford, K. (2000). *A case study in the research paradigm of human-centered computing: Local expertise in weather forecasting* (Tech. Rep.). Pensacola, FL: National Technology Alliance.

Hoffman, R. R., Crandall, B., & Shadbolt, N. R. (1998). A case study in cognitive task analysis methodology: The critical decision method for the elicitation of expert knowledge. *Human Factors*, *40*, 254-276.

Hoffman, R. R., & Lintern, G. (2006). Eliciting and representing the knowledge of experts. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The cambridge handbook of expertise and expert performance* (p. 203-222). New York: Cambridge University Press.

Hoffman, R. R., & Woods, D. D. (2000). Studying cognitive systems in context: Preface to the special section. *Human Factors*, *42*, 1-7.

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.

Hoogveld, A. W. M., Paas, F., & Jochems, W. M. G. (2005). Training higher education teachers for instructional design of competency-based education: Product-oriented versus process-oriented worked examples. *Teaching and Teacher Education*, *21*, 287-297.

Hughes, J., & Parkes, S. (2003). Trends in the use of verbal protocol analysis in software engineering research. *Behaviour & Information Technology*, *22*, 127-140.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Imhof, B., Scheiter, K., & Gerjets, P. (2009). Realism in dynamic, static-sequential, and static-simultaneous visualizations during knowledge acquisition on locomotion patterns. In N. A. Taatgen & H. Van Rijn (Eds.), *Proceedings of the 31st annual conference of the cognitive science society* (p. 2962-2967). Austin, TX.

International League Against Epilepsy. (2010). Revised terminology and concepts for organization of the epilepsies: Report of the commission on classification and terminology. *Epilepsia*, *51*, 676-685.

Jankovic, J. (2008). Parkinson's disease: Clinical features and diagnosis. *Journal of Neurology, Neurosurgery, and Psychiatry*, *79*, 368-376.

Jentsch, F., Bowers, C., & Salas, E. (2001). What determines whether observers recognize targeted behaviors in modeling displays? *Human Factors*, *43*, 496-507.

Jonassen, D. H., Tessmer, M., & Hannum, W. H. (1999). *Task analysis methods for instructional design.* Mahwah, NJ: Lawrence Erlbaum Associates.

Jongman, R. W. (1967). *Het oog van de meester (In the eye of the master).* Unpublished doctoral dissertation, University of Amsterdam, Amsterdam.

Jucks, R., Schulte-Löbbert, P., & Bromme, R. (2007). Supporting experts' written knowledge communication through reflective prompts on the use of specialist concepts. *Journal of Psychology*, *215*, 237-247.

Just, M., & Carpenter, P. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*, 441–480.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*, 23-32.

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

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

Kamin, C., Deterding, R., Wilson, B., Armacost, M., & Breedon, T. (1999). The development of a collaborative distance learning program to facilitate pediatric problem-based learning. *Medical Education Online*, *4*, 1-12.

Kamin, C., O'Sullivan, P., Deterding, R., & Younger, M. (2003). A comparison of critical thinking in groups of third-year medical students in text, video, and virtual PBL case modalities. *Academic Medicine*, *78*, 204-211.

Karpf, D. (1973). *Thinking aloud in human discrimination learning.* Unpublished doctoral dissertation, State University of New York.

Kieras, D. (2004). Goms models for task analysis. In D. Diaper & N. A. Stanton (Eds.), *The handbook of task analysis for human-computer interaction* (p. 83-116). Mahwah, NJ: Lawrence Erlbaum Associates.

Kingdon, J. S. (1980). The role of visual signals and face patterns in african forest monkeys (guenons) of the genus cercopithecus. *The Transactions of the Zoological Society of London*, *35*, 425-475.

Kinney, J. A. S., Luria, S. M., & Weitzman, D. O. (1967). Visibility of colors underwater. *Journal of the Optical Society of America*, *57*, 802-809.

Kirwan, B., & Ainsworth, L. K. (1992). *A guide to task analysis.* London: Taylor & Francis.

Kitsantas, A., Zimmerman, B. J., & Cleary, T. (2000). The role of observation and emulation in the development of athletic self-regulation. *Journal of Educational Psychology*, *92*, 811-817.

Klahr, D. (1976). *Cognition and instruction*. Hillsdale, NJ: Erlbaum.

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambock (Eds.), *Decision making in action: Models and methods* (p. 138-147). Norwood, NJ: Ablex.

Krupinski, E. A. (1996). Visual scanning patterns of radiologists searching mammograms. *Academic Radiology*, *3*, 137 - 144.

Krupinski, E. A. (2005). Visual search of mammographic images: Influence of lesion subtlety. *Academic Radiology*, *12*, 965 - 969.

Krupinski, E. A., Tillack, A. A., Richter, L., Henderson, J. T., Bhattacharyya, A. K., Scott, K. M., et al. (2006). Eye-movement study and human performance using telepathology virtual slides: Implications for medical education and differences with experience. *Human Pathology*, *37*, 1543 - 1556.

Kuipers, B., & Kassirer, J. P. (1984). Causal reasoning in medicine: Analysis of a protocol. *Cognitive Science*, *8*, 363-385.

Kundel, H., Nodine, C., Krupinski, E., & Mello-Thoms, C. (2008). Using gaze-tracking data and mixture distribution analysis to support a holisitc model for the detection of cancers on mammograms. *Academic Radiology*, *15*, 881-886.

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday live*. Cambridge: Cambridge University Press.

Law, J., & Lynch, M. (1990). Lists, field guides, and the descriptive organization of seeing: Birdwatching as an exemplary observational activity. In M. Lynch & S. Woolgar (Eds.), *Representation in scientific practice.* (p. 267-299). Cambridge: MIT Press.

Lesgold, A. (1984). Acquiring expertise. In J. R. Anderson & S. M. Kosslin (Eds.), *Tutorials in learning and memory* (p. 31-60). San Francisco, CA: Freeman.

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing x-ray pictures. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise* (p. 311-342). Hillsdale, NJ: Erlbaum.

Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physice – Doklady*, *10*, 707-710.

Lieske, E., & Myers, R. (1994). *Korallenfische der Welt: 2044 Arten in Farbe.* Hamburg: Jahr Top Spezial.

Lindsey, C. (1978). Form, function, and locomotory habits in fish. *Fish Physiology*, *7*, 1-88.

Lintern, G., Miller, D., & Baker, K. (2002). Work centered design of a USAF mission planning system. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (p. 531-535). Santa Monica, CA: Human Factors and Ergonomics Society.

Loddenkemper, T., Kellinghaus, C., Wyllie, E., Najm, I. M., Gupta, A., & Rosenow, F. (2005). A proposal for a five-dimensional patient-oriented epilepsy classification. *Epileptic Disorders*, *7*, 308-316.

Lorenz, K., & Tinbergen, N. (1939). Taxis und Instinkthandlung der Eirollbewegung der Graugans. *Zeitschrift für Tierpsychologie*, *2*, 1-29.

Low, R., & Sweller, J. (2005). The modality principle in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (p. 147-158). New York: Cambridge University Press.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, *14*, 225-244.

Lowe, R. K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction*, *13*, 157-176.

Lowe, R. K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, *14*, 257-274.

Lowe, R. K. (2008). Learning from animation: Where to look, when to look. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation: Research implications for design* (p. 49-68). New York: Cambridge University Press.

Luczak, H. (1997). Task analysis. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (p. 340-416). New York: John Wiley & Sons.

Lüders, H., Acharya, J., Baumgartner, C., Benbadis, S., Bleasel, A., Burgess, R., et al. (1998). Semiological seizure classification. *Epilepsia*, *39*, 1006-1013.

Manning, D. J., Gale, A., & Krupinski, E. A. (2005). Perception research in medical imaging. *The British Journal of Radiology*, *78*, 683-685.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology*, *93*, 377-389.

Mayer, R. E. (2005a). Principles for managing essential processing in multimedia learning: Segmenting, pretrainung, and modality principles. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (p. 169-182). New York: Cambridge University Press.

Mayer, R. E. (2005b). Principles of multimedia learning based on social cues: Personalization, voice, and image principles. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (p. 201-212). New York: Cambridge University Press.

McLaren, B. M., Lim, S., & Koedinger, K. R. (2008). When and how often should worked examples be given to students? New results and a summary of the current state of research. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (p. 2176-2181). Austin, TX: Cognitive Science Society.

Medin, D. L., Lynch, E. B., Coley, J. D., & Atran, S. (1997). Categorization and reasoning among tree experts: Do all roads lead to rome? *Cognitive Psychology*, *32*, 49-96.

Medin, D. L., Ross, N., Atran, S., Cox, D., Coley, J., Proffitt, J. B., et al. (2006). Folkbiology of freshwater fish. *Cognition*, *99*, 237-73.

Meister, D. (1999). *The history of human factors and ergonomics*. Mahwah, NJ: Lawrence Erlbaum Associates.

Meyer, B. J. F. (1975). *The organization of prose and its effects on memory.* New York: Elsevier.

Meyer, K., Rasch, T., & Schnotz, W. (2009). Effects of animation's speed of presentation on perceptual processing and learning. *Learning and Instruction*, *20*, 136-145.

Militello, L. G., & Hutton, R. J. B. (1998). Applied cognitive task analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. *Ergonomics*, *41*, 1618-1641.

Miller, R. B. (1953). *A method for man-machine task analysis* (Tech. Rep.). Dayton, OH: Wright Air Development Center.

Miller, R. B. (1962). Task description and analysis. In R. M. Gagne (Ed.), *Psychological principles in systems development* (p. 187-228). New York: Holt, Rinehart & Winston.

Moreno, F. J., Reina, R., Luis, V., & Sabido, R. (2002). Visual search strategies in experienced and inexperienced gymnastic coaches. *Perceptual and Motor Skills*, *95*, 901-902.

Myles-Worsley, M., Johnston, W. A., & Simons, M. A. (1988). The influence of expertise on X-ray image processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *14*, 553-557.

Nachtigall, W. (2002). *Bionik: Grundlagen und Beispiele für Ingenieure und Naturwissenschaftler.* Berlin: Springer.

Nachtigall, W. (2006). *Ökophysik: Plaudereien über das Leben auf dem Land, im Wasser* *und in der Luft.* Berlin: Springer.

Naikar, N., & Sanderson, P. M. (2001). Evaluating system design proposals with work domain analysis. *Human Factors*, *43*, 529-542.

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

Nievelstein, F., Van Gog, T., Van Dijck, G., & Boshuizen, H. P. A. (2010). *The worked-example and expertise revearsal effect in less structured tasks: Learning to reason about legal cases.* (Manuscript submitted for publication.)

Nodine, C. F., Kundel, H. L., Lauver, S. C., & Toto, L. C. (1996). Nature of expertise in searching mammograms for breast masses. *Academic Radiology*, *3*, 1000 - 1006.

Nordli, D. (2002). Infantile seizures and epilepsy syndromes. *Epilepsia*, *43*, 11-16.

Nordli, D., Bazil, C. W., Scheuer, M. L., & Pedley, T. A. (1997). Recognition and classification of seizures in infants. *Epilepsia*, *38*, 553-560.

Nordli, D., Kuroda, M. M., & Hirsch, L. J. (2001). The ontogeny of partial seizures in infants and young children. *Epilepsia*, *42*, 986-990.

Norman, G. R., Coblentz, C. L., Brooks, L. R., & Babcook, C. J. (1992). Expertise in visual diagnosis: A review of the literature. *Academic Medicine*, *67*, 78-83.

Nückles, M., Winter, A., Wittwer, J., Herbert, M., & Hübner, S. (2006). How do experts adapt their explanations to a layperson's knowledge in asynchronous communication? An experimental study. *User Modeling and User-Adapted Interaction*, *16*, 87-127.

Nunnally, J. C. (1967). *Psychometric theory.* New York: McGraw-Hill Book Company.

Nyström, M. (2008). *Off-line foveated compression and scene perception: An eye tracking approach.* Unpublished doctoral dissertation, Lund University.

Nyström, M., & Holmqvist, K. (2007). Deriving and evaluating eye-tracking controlled volumes of interest for variable-resolution video compression. *Journal of Electronic Imaging*, *16*, 013006.

Nyström, M., & Holmqvist, K. (2008). Semantic override of low-level features in image viewing – both initially and overall. *Journal of Eye Movement Research*, *2*, 1-11.

Ohlsson, S., & Rees, E. (1991). The function of conceptual understanding in the learning of arithmetic procedures. *Cognition and Instruction*, *8*, 103-179.

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

Paas, F., & Van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive load appraoch. *Journal of Educational Psychology*, *86*, 122-133.

Parr, W. V., Heatherbell, D., & White, K. G. (2002). Demystifying wine expertise: Olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chemical Senses*, *27*, 747-755.

Pfeiffer, V. D. I., Gemballa, S., Jarodzka, H., Scheiter, K., & Gerjets, P. (2009). Situated learning in the mobile age: Mobile devices on a field trip to the sea. *ALT-J, Research in Learning and Technology*, *17*, 187-199.

Pfeiffer, V. D. I., Scheiter, K., Kühl, T., & Gemballa, S. (2010). *Learning how to identify species in a situated learning scenario: Using dynamic-static visualizations to prepare students for their visit to the aquarium.* (Manuscript under preparation)

Pinheiro, V. D., & Simon, H. A. (1992). An operational model of motor skill diagnosis. *Journal of Teaching in Physical Education*, *11*, 288-302.

Pirolli, P. L., & Anderson, J. R. (1985). The role of learning from examples in the acquisition of recursive programming skills. *Canadian Journal of Psychology*, *39*, 240-272.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology.*, *32*, 3-25.

Prensky, A. L. (2000). An approach to the child with paroxysmal phenomena with emphasis on nonepileptic disorders. In J. M. Pellock, W. E. Dodson, & B. F. D. Bourgeois (Eds.), *Pediatric epilepdy: Diagnosis and therapy* (p. 97-116). New York: Demos Medical Publishing.

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

Raab, M., & Johnson, J. G. (2007). Expertise-based differences in search and option-generation strategies. *Journal of Experimental Psychology: Applied*, *13*, 158-170.

Rao, R. P. N., Zielinsky, G. J., Hayhoe, M. M., & Ballard, D. H. (2002). Eye movements in iconic visual search. *Vision Research*, *42*, 1447-1463.

Rasmussen, J. (1986). *Information processing and human-machine interaction. An approach to cognitive engineerin*. Amsterdam: Elsevier Science Publishers.

Reimann, P., & Rapp, A. (2008). Expertiseerwerb. In A. Renkl (Ed.), *Lehrbuch Pädagogische Psychologie* (p. 155-203). Bern: Huber.

Reingold, E., & Charness, N. (2005). Perception in chess: Evidence from eye movements. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (p. 325-354). Oxford: Oxford University Press.

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

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

Renkl, A. (2002). Worked-out examples: Instructional explanations support learning by self-explanations. *Learning and Instruction*, *12*, 529-556.

Renkl, A. (2010). *Towards an instructionally oriented theory of example-based learning.* (Manuscript submitted for publication.)

Resnick, L. B. (1987). Learning in school and out. *Educational Researcher*, *16*, 13-20.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, *29*, 1045-1060.

Ripoll, H. (1991). The understanding-acting process in sport - the relationship between the semantic and the sensorimotor visual function. *International Journal of Sport Psychology*, *22*, 221-243.

Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem-solving in computer-mediated settings. *Journal of the Learning Sciences*, *14*, 201-241.

Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *International Journal of Computer-Supported Collaborative Learning*, *4*, 69-92.

Sabers, D. S., Cushing, K. S., & Berliner, D. C. (1991). Differences among teachers in a task characterized by simultaneity, multidimensionality, and immediacy. *American Educational Research Journal*, *28*, 63-88.

Salvucci, D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eyetracking protocols. In A. T. Duchowski (Ed.), *In Proceedings of the Eye Tracking Research and Applications (ETRA) Symposium 2000* (p. 71-78). New York: ACM Press.

Schaafstal, A. M. (1991). *Diagnostic skill in process operation: A comparison between experts and novices.* Unpublished doctoral dissertation, University of Groningen, The Netherlands.

Schaafstal, A. M. (1993). Knowledge and strategies in diagnostic skill. *Ergonomics*, *36*, 1305-1316.

Schaafstal, A. M., & Schraagen, J. M. (1992). *A method for cognitive task analysis* (Tech. Rep. No. Report IZF 1992 B-5). Soesterberg, NL: TNO Institute for Perception.

Schaafstal, A. M., & Schraagen, J. M. (2000). Training of troubleshooting: A structured, task analytical approach. In J. M. Schraagen, S. F. Chipman, & V. L. Shalin (Eds.), *Cognitive task analysis* (p. 57-70). Mahwah, NJ: Lawrence Erlbaum Associates.

Schaafstal, A. M., Schraagen, J. M., & Van Berlo, M. (2000). Cognitive task analysis and innovation of training: The case of structured troubleshooting. *Human Factors*, *42*, 75-86.

Scheiter, K., Gerjets, P., Huk, T., Imhof, B., & Kammerer, Y. (2009). The effects of realism in learning with dynamic visualizations. *Learning and Instruction*, *19*, 481-494.

Scheiter, K., & Van Gog, T. (2009). Using eye tracking in applied research to study and stimulate the processing of information from multi-representational sources [Special Issue]. *Applied Cognitive Psychology*, *23*.

Schmidt, D., & Schachter, S. C. (2000). *Epilepsy: Problem solving in clinical practice*. London: Martin Dunitz.

Schmidt, H. G., & Boshuizen, H. P. A. (1992). Encapsulation of biomedical knowledge. In

D. A. Evans & V. L. Patel (Eds.), *Advanced models of cognition for medical training and practice* (p. 265-282). New York, NY: Springer Verlag.

Schmidt, H. G., & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational Psychology Review*, *5*, 205-221.

Schnotz, W., & Lowe, R. K. (2008). A unified view of learning from animated and static graphics. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation: Research and design implications* (p. 304-356). New York: Cambridge University Press.

Schraagen, J. M. (1993). What information do river pilots use? In *Proceedings of the international conference on marine simulation and ship manoeuvrability MARSIM '93* (p. 509-517). St. John's, Newfoundland: Fisheries and Marine Institute of Memorial University.

Schraagen, J. M. (2006). Task analysis. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The cambridge handbook of expertise and expert performance* (p. 185-201). New York: Cambridge University Press.

Schraagen, J. M., Chipman, S. F., & Shalin, V. L. (2000). *Cognitive task analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schraagen, J. M., & Schaafstal, A. M. (1996). Training of systematic diagnosis: A case study in electronics troubleshooting. *Le Travail Humain*, *59*, 5-21.

Schunk, D. H. (1981). Peer models and children's behavioral change. *Journal of Educational Psychology*, *73*, 93-105.

Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficiacy and achievement. *Journal of Educational Psychology*, *77*, 313-322.

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.

Seagull, F. J., & Yan, X. (2001). *Using eye-tracking video data to augment knowledge elicitation in cognitive task analysis.* Paper presented at the Human Factors and Ergonomics Society 45th Annual Meeting. Minneapolis, MN: Human Factors and Ergonomics Society. (Abstract retrieved from http://citeerx.ist.psu.edu/viewdoc/download?doi=10.1.1.113.3849&rep=rep1&type=pdf)

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (p. 325-348). Hillsdale, NJ: Erlbaum.

Simon, H. A. (1983). Why should machines learn? In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (p. 25-38). Palo Alto, CA: Tioga.

Simon, H. A., & Kaplan, C. A. (1989). Foundations of cognitive science. In M. I. Posner (Ed.), *Foundations of cognitive science* (p. 1-47). Cambridge, MA: MIT Press.

Simon, S. J., & Werner, J. M. (1996). Computer training through behavior modeling, self-paced, and instructional approaches: A field experiment. *Journal of Applied Psychology*, *81*, 648-659.

Simpson, S. A., & Gilhooly, K. J. (1997). Diagnostic thinking processes: Evidence from a constructive interaction study of electrocardiogram (ECG) interpretation. *Applied Cognitive Psychology*, *11*, 543-554.

Smeets, J., & Hooge, I. (2003). Nature of variability in saccades. *Journal of Neurophysiology*, *90*, 12–20.

Smith, P. L., & Ragan, T. J. (1999). *Instructional design* (2nd ed.). Upper Saddle River, NJ: Merril/Prentice Hall.

Soloway, E. M., & Johnson, W. L. (1984). Remembrence of blunders past: A retrospective on the development of PRUOST. In *Proceedings of the sixth cognitive science society conference.* (p. 57). Colorado: Boulder.

Speelman, M. (1998). Implicite expertise: Do we expect too much from our experts? In K. Kirsner, M. Speelman, M. Maybery, M. O'Brien-Malone, M. Anderson, & C. MacLeod (Eds.), *Implicit and explicit mental processes* (p. 135-147). Mahwah, NJ: Erlbaum.

Stelmach, L. B., Campsall, J. M., & Herdman, C. M. (1997). Attentional and ocular movements. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 823-844.

Ste-Marie, D. M., & Lee, T. D. (1991). Prior processing effects on gymnastic judging. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *17*, 126-136.

Struve, D. (2008). Process-oriented worked examples for training older adults to use interactive systems. *Gerontechnology*, *7*, 216.

Sudman, S., Bradbrun, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology.* San Francisco, CA: Jossey-Bass.

Sweller, J. (1988). Element interactivity and intrinsic, extraneous, and germane cogntive load. *Educational Psychology Review*, *22*, 123138.

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

Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (p. 159-167). Cambridge: University Press.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*, 59-89.

Sweller, J., & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology*, *4*, 434-458.

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

Taylor, F. W. (1998/1911). *The principles of scientific management*. Mineola, NY / New York and London: Dover Pulications / Harper & Brothers.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, *57*, 247-262.

Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics*, *46*, 629-646.

Van Gog, T. (2006). *Uncovering the problem-solving process to design effective worked examples*. Unpublished doctoral dissertation, Open Universiteit Nederland.

Van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model‘s eye movements. *Computers in Human Behavior*, *25*, 785-791.

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

Van Gog, T., Paas, F., & Van Merriënboer, J. J. G. (2005). Uncovering expertise-related differences in troubleshooting performance: Combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology*, *19*, 205-221.

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

Van Gog, T., Paas, F., & Van Merriënboer, J. J. G. (2008). Effects of studing sequences of process-oriented and product-oriented worked examples on troubleshooting transfer efficiency. *Learning and Instruction*, *18*, 211-222.

Van Gog, T., Paas, F., Van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied*, *11*, 237-244.

Van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cogntive research perspectives. *Educational Psychology Review*, *22*, 155-174.

Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning [Special Issue]. *Learning and Instruction*, *10*.

Van Lehn, K. (1989). Problem solving and cognitive skill acquisition. In M. Posner (Ed.), *Foundations of cognitive science.* (p. 527-579). Mahwah, NJ: Erlbaum.

Van Lehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology*, *47*, 513-539.

Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training.* Englewood Cliffs, NJ: Educational Technology Publications.

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

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

Velichkovsky, B. M. (1995). Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics and Cognition*, *3*, 199-224.

Vogt, S., & Magnussen, S. (2007). Expertise in pictorial perception: Eye movement patterns and visual memory in artists and laymen. *Perception*, *36*, 91-100.

Wade, N. J., & Tatler, B. (2005). *The moving tablet of the eye: The origins of modern eye movement research*. Oxford: Oxford University Press.

Ward, P., Hodges, N. J., Williams, A., & Starkes, J. L. (2004). Deliberate practice and expert performance: Defining the path to excellence. In A. Williams & N. J. Hodges (Eds.), *Skill acquisition in sport: Research, theory, and practice.* (p. 231-258). London, UK: Routledge.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*, 158-177.

Watson, J. B. (1920). Is thinking merely the action of language mechanisms? *British Journal of Psychology*, *11*, 87-104.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive measures in the social sciences.* Chicago: Rand McNally.

Webb, P. W. (1984). Der Fischkörper: Form und Bewegung. *Spektrum der Wissenschaft*, *9*, 84-97.

Williams, A., & Ward, P. (2003). Perceptual expertise: Development in sport. In J. L. Starkes & K. A. Ericsson (Eds.), *Expert performance in sports* (p. 219-250). Champaign, IL: Human Kinetics.

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, *83*, 73-87.

Wittwer, J., Nückles, M., & Renkl, A. (2010). Using a diagnosis-based approach to individualize instructional explanations in computer-mediated communication. *Educational Psychological Review*, *22*, 9-23.

Wolf, F. M., Miller, J. G., Borzynski, M. E., Schlesinger, A., Rosen, D. S., & al. et. (1994). Effects of expertise and case difficulty on the interpretation of pediatric radiographs. *Academic Medicine*, *69*, 31-34.

World Health Organisation. (2006). *International classification of diseases (ICD-10).* Genf: World Health Organisation (WHO).

Wouters, P., Paas, F., & Van Merriënboer, J. J. G. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load. *Review of Educational Research*, *78*, 645-675.

Wouters, P., Paas, F., & Van Merriënboer, J. J. G. (2009). Observational learning from animated models: Effects of modality and reflection on transfer. *Contemporary Educational Psychology*, *34*, 1-8.

Zusammenfassung

Forschung zum Erwerb kognitiver Fertigkeiten hat bisher weitgehend vernachlässigt, dass viele Aufgaben im wahren Leben die Verarbeitung und Interpretation komplexer, visueller Reize voraussetzen. Daher erfordern viele dieser Aufgaben nicht nur kognitive, sondern auch perzpetuelle Fertigkeiten. Beispiele für solche Aufgaben, die im Rahmen der vorliegenden Dissertation untersucht wurden, sind die Klassifikation vom Schwimmbewegungsmustern bei Rifffischen sowie die Diagnose epileptischer Anfälle bei Kleinkindern. Somit war das Ziel der vorliegenden Dissertation zweifältig: Zum einen, sollte die Rolle perzeptueller Fertigkeiten für die Expertise innerhalb solcher Aufgaben untersucht werden. Zum anderen, sollte eine effektive Instruktionsmaßnahme zur Entwicklung jener perzpetuellen Fertigkeiten entwickelt werden.

Um erfolgreich eine Instruktionsmaßnahme zu entwickeln, wurden diejenigen Aufgaben, die innerhalb der vorliegenden Dissertation von Interesse sind, analysiert. Dadurch sollte ein besseres Verständnis, dafür wie diese Aufgaben bearbeitet werden und welche Schwierigkeiten dabei entstehen können, geschaffen werden. Kapitel 2 stellte verschiedene Methoden zur Aufgabenanalyse vor und wie diese genuzt werden können, um Instruktionsmaßnahmen zu entwickeln. Basierend auf diesem Literaturüberblick, wurde die Methode zur Entwicklung eines Aufgabenmodels von Schaafstal und Schraagen (1992) ausgewählt und im Rahmen dieser Dissertation angewendet. Diese Aufgabenanalysemethode besteht aus der Beschreibung der schrittweisen Prozedur der Aufgabenbearbeitung, der Identifikation des Wissens und der Fertigkeiten, die für diese Aufgabe von Bedeutung sind, der Voraussage von Schwierigkeiten, die bei der Bearbeitung dieser Aufgabe auftreten können, sowie

Schlussfolgerungen darüber wie ein Training dieser Aufgabe gestaltet werden sollte.

Kapitel 3 präsentiert eine zweistufige Aufgabenanalyse. Im ersten Teil ergab eine normative Analyse, wie die Aufgabe der Klassifikation der Schwimmbewegungen von Rifffischen bewältigt werden *sollte*. Dieser Teil der Aufgabenanalyse wurde mit Hilfe von Domänenexperten und Lehrbüchern zur marinen Zoologie durchgeführt. In einem zweiten Schritt wurde eine empirische Aufgabenanalyse durchgeführt deren Ziel es war zu ergründen, wie diese Aufgabe *tatsächlich* durchgeführt wird (Kapite 3.2). Dadurch sollten die Ergebnisse der normativen Analyse bestätigt und ergänzt werden. Dies wurde umgesetzt, indem Experten und Novizen auf Prozess- und Performanzebene, während der Klassifikation von Schwimmbewegungen von Rifffischen, verglichen wurden. Dabei waren die interessierenden Prozesse sowohl kognitiver, als auch perzpetueller Natur. Die Hauptbefunde ergaben, dass Experten im Vergleich zu Novizen diese Aufgabe schneller und akkurater bewältigten, indem sie schneller die relevanten Körperteile des Fisches beachteten und erfahrungsbasierte "Abkürzungen" nutzten, während Novizen durch irrelevante, aber auffällige Merkmale abgelenkt wurden. Insgesamt bestätigten und ergänzten diese Ergebnisse, die Schlussfolgerungen aus der normativen Aufgabenanalyse.

In Kapitel 4 wurden die aus der normativen und empirischen Aufgabenanalyse gewonnen Einsichten genutzt, um eine neue Instruktionsmaßnahme für die Vermittlung perzeptueller Fertigkeiten für die Aufgabe der Klassifizierung der Schwimmbewegungen von Rifffischen zu entwickeln. Diese Methode ist abgeleitet aus der Forschung zum Lernen aus ausgearbeiteten Lösungsbeispielen (Sweller et al., 1998) und kognitiver Modellierung (Collins et al., 1989). Diese Methode besteht aus zwei Komponenten. Die erste Komponente sind verbale Erklärungen eines Domänenexperten (d.h., Modellierung kognitiver Prozesse), die sie oder er während der Aufgabenbearbeitung äußert. Die zweite Komponente ist eine Videoaufzeichnung von einem dynamischen Reiz (z.B. ein schwimmender Fisch oder ein sich bewegendes Kind) mit einer Überlagerung der Blickbewegungen des Experten auf diesem Video. Diese Blickbewegungen sollen zum einen, dem Lernen Informationen darüber bieten, wie die Aufgabe auf einer perzeptuellen Ebene bearbeitet werden soll und zum anderen, die visuelle Aufmerksamkeit des Lerners lenken. Zusammengenommen wurden die verbalen Erklärungen des Experten und die Überlagerung ihrer oder seiner Blickbewegungen auf dem Video Blickbewegungs-Modellierung-Beispiele genannt (englisch *eye movement modeling examples*, EMME; Van Gog et al., 2009).

In der zweiten Studie im Rahmen dieser Dissertation (Kapitel 5) wurde die neue Instruktionsmaßnahme EMME angewendet. Dabei wurden zwei verschiedene Arten der Darstellung der Blickbewegungen des Experten dahingehend mit einer Kontrollgruppe, die keine Blickbewegungen des Experten sah, verglichen, wie gut sie die visuelle Aufmerksamkeit der Lerner lenkten und verschiedene Arten des Lernens unterstüzten. Die Hauptbefunde legen nahe, dass EMME die Vermittlung perzeptueller Fertigkeiten fördert, welche wiederum den Lerner zu einer effizienten visuellen Suche relevanter Körperteile und Interpretation ihrer Bewegungsmuster befähigen. Interessanterweise förderten die beiden Darstellungsarten der Expertenblickbewegungen unterschiedliche Aspekte des Lernens: Während eine schweinwerferartige Darstellung der Blickbewegungen die visuelle Suche förderte, verbesserte eine Darstellung der Blickbewegungen als Punkt die Interpretation der Schwimmbewegungen. Da keine der beiden Darstellungsarten sowohl die visuelle Suche als auch die Interpretation erfolgreich vermitteln konnte, war eine der Schlussfolgerungen aus Studie 2, dass EMME weiter verbessert werden musste.

Um die Generalisierbarbeit von EMME auf andere Domänen zu überprüfen, wurde diese Instruktionsmaßnahme auf eine zweite Aufgabe angewendet: die Diagnose epileptischer Anfälle bei Kleinkindern. Wiederum wurde eine Aufgabenanalyse durchgeführt, die die Anwendbarkeit von EMME auf diese Aufgabe rechtfertigte (Kapitel 6). Basierend auf dieser Aufgabenanalyse wurde eine adaptierte Version von EMME in zwei verbesserten Darstellungsarten implementiert um perzeptuelle Fertigkeiten in der Medizindiagnose zu vermitteln (Kapitel 7). Die Ergebnisse von Studie 3 bestätigen die Befunde aus Studie 2 weitestgehend, indem EMME die Lerner dazu befähigt die relevanten Körperteile des Kleinkindes visuell zu suchen und deren Bewegungen korrekt zu interpretieren. Insbesondere fördert eine im Vergleich zu Studie 2 optimierte Scheinwerferdarstellung perzpetuelle Fertigkeiten, sowohl auf Prozess- als auch auf Performanzebene, im Vergleich zu einer anderen EMME Darstellungsart und einer Kontrollgruppe.

Abschließend fasst Kapitel 8 die Ergebnisse zusammen und diskutiert diese im Hinblick auf künftige Forschung. Insgesamt hat diese Dissertation die Wichtigkeit perzeptueller Fertigkeiten für Aufgaben die auf komplexen, visuellen Reizen basieren, gezeigt. Darüberhinaus hat die vorliegende Dissertation gezeigt, dass perzeptuelle Fertigkeiten durch EMME verbessert werden können, wobei eine effizientere visuelle Suche relevanter Informationen und deren korrektere Interpretation erfolgt. Bei der Implementation dieser Instruktionsmaßnahme muß jedoch sowohl die Art der zu lernenden Aufgabe, als auch die Art der Darstellung von EMME beachtet werden.

Appendices

Appendix Study 1

Defining ”Dynamic” AAOIs

The computer tool that enables the transformation of AOIs to ”dynamic” AAOIs was programmed in Java (java.sun.com). The input files are ClearView 2.7.0 output TXT files. The tool first calculates the dwell time for each AOI across all 100 ms segments per video. It then aggregates the AOIs into AAOIs (relevant for locomotion, species, both, or irrelevant) across all four videos. In addition, the tool generates a sequence of the AAOIs for each video, which serves as input for the sequence analysis tool. All output of the tool is given in TXT files.

Sequence Analysis

The computer tool that was used for the sequence analysis was programmed in Ruby (www.ruby-lang.org). The input data files are in TXT file types, which were converted in CSV file types. The output data files are in TXT file types. The program calculates the edit distance and the similarity percentage between two strings according to Levenshtein. The program can be assessed at the URL code.google.com/p/eye-tracker-tools.

Appendix Study 2

Data Study 2

Table 3:
Means (and SD) for Eye Movement Similarity During Example Study (in pxl), for Eye Movement Measures (in ms), and for Performance During Multiple-Choice Test (in perc)

		Spotlight display group	Dot display group	Control group
Learning	Eye movement similarity	93.89 (8.36)	121.61 (7.31)	136.07 (20.99)
Testing	Time until looked at relevant AOIs	1205.00 (391.75)	1530.36 (509.55)	1662.87 (697.05)
	Time spent on relevant AOIs	751.82 (300.57)	701.00 (332.07)	505.19 (353.68)
	Interpretation score	71.28 (11.04)	78.55 (11.92)	70.82 (10.45)

Appendix Study 3

Data Study 3

Table 4:
Means (and SD) for Eye Movement Similarity During Example Study (in pxl), for Eye Movement Measures (in ms), and for Performance During Multiple-Choice Test (in perc)

		Spotlight display group	Circle display group	Control group
Learning	Eye movement similarity	210.79 (37.06)	238.24 (44.02)	237.43 (27.46)
Testing	Time until looked at relevant AOIs	189.52 (86.60)	274.80 (123.70)	289.88 (169.44)
	Time spent on relevant AOIs	702.36 (129.56)	607.43 (116.22)	605.58 (142.16)
	Interpretation score	60.33 (12.88)	53.33 (7.17)	50.67 (10.01)