Event cognition at the workplace: perceiving, understanding, and practicing assembly tasks

Vom Fachbereich Sozialwissenschaften
der Technischen Universität Kaiserslautern
zur Verleihung des akademischen Grades
Doktor der Philosophie (Dr. phil.)
genehmigte

Dissertation

vorgelegt von

Dipl.-Psych. Katharina Sebastian
aus Rybnik (Polen)

Acknowledgments

This work was conducted during my employment as a researcher at the Innovative Factory Systems group at the Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH and the Technologie-Initiative SmartFactoryKL e.V. both led by Prof. Dr. Dr. h.c. Detlef Zühlke. His insistent, sincere promotion of interdisciplinary research built the foundation for the present thesis, and I am thankful for this opportunity.

I would like to thank Prof. Dr. Tandra Ghose for her dedicated supervision and countless feedbacks during the last four years. Especially, I appreciate her persistent confidence, companionship, and encouragement.

A big piece of good fortune was meeting Prof. Dr. Markus Huff at a conference in 2012 constituting the beginning of a rewarding collaboration. I wish to express my deep gratitude to him for numerous helpful discussions, words of advice, and suggestions. Beyond that, he initiated my visit to the Washington University in St. Louis. There, I benefited a lot from the stimulating exchange with Prof. Dr. Jeffrey Zacks and his support concerning event segmentation data analysis.

Hans-Joachim Berberich, his team, and the workers at the Westpfalz-Werkstätten im Gemeinschaftswerk für Menschen mit Behinderungen GmbH made the event segmentation study for Experiment 3 possible. I appreciate our collaboration very much.

Furthermore, I thank the EU project consortiums of COGNITO and VISTRA for developing the computer-based systems relevant to the present thesis. In particular, I want to thank Dr. Nils Petersen who was responsible for the software development of the IBES tool and Dr. Dominic Gorecky who successfully pulls together interdisciplinary teams.

Sincere thanks go to Barbara Seitz, Jedrzej Pelka, Katrin Schilling, Konrad Lauenroth, and Ricarda Moses who helped me with data collection.

Working at the DFKI and the SmartFactory was a rewarding experience also due to my colleagues: Dr. Moritz Kümmerling, Dr. Dominic Gorecky, Dr. Marc Seißler, Dr. Marius Orfgen, Moritz Ohmer, Fabian Quint, Dennis Kolberg, Christian Stahl, Stefanie Fischer, Ricardo Campos, Stephan Weyer, Dr. Matthias Loskyll, Ines Dahmann, Dr. Tobias Gerber, and Yannik Schelske. Together with the "Pflaumenbaum" they contributed to having a memorable time in Kaiserslautern.

Finally, I am very grateful for the consistent backup of my family and friends, and especially for the support of my anchor Tom.

Abstract

The Event Segmentation Theory (Kurby & Zacks, 2008; Zacks, Speer, Swallow, Braver, & Reynolds, 2007) explains the perceptual organization of an ongoing activity into meaningful events. The classical event segmentation task (Newtson, 1973) involves watching an online video and indicating with key presses the event boundaries, i.e., when one event ends and the next one begins. The resulting hierarchical organization of object-based coarse events and action-based fine events gives insight into various cognitive processes. I used the Event Segmentation Theory to develop assistance and training systems for assembly workers in industrial settings at various levels - experts, new hires, and intellectually disabled people.

Therefore, the first scientific question I asked was whether online and offline event segmentation result in the same event boundaries. This is important because assembly work requires not only watching activities online but processing the information offline, e.g., while performing the assembly task. By developing a special software tool that enables assessment of offline event boundaries, I established that online perception and offline elaboration lead to similar event boundaries. This study supports prior work suggesting that instructions should be structured around event boundaries.

Secondly, I investigated the importance of fine versus coarse event boundaries when learning the sequence of steps in virtual training, both for novices and experts in car door assembly. I found memory, tested by ability to predict the next frame, to be enhanced for object-based coarse events from the nearest fine event boundary. However, virtual training did not improve memory for action-based fine events from the nearest coarse event boundary. I conjecture that trainees primarily acquire the sequence of object-based coarse events in an initial training. Based on differences found in memory performance between experts and novices, I conclude that memory for action-based fine events is dependent on expertise.

Thirdly, I used the Event Segmentation Theory to investigate whether the simple and repetitive assembly tasks offered at workshops for intellectually disabled persons utilize their full cognitive potential. I analyzed event segmentation performance of 32 intellectually disabled persons compared to 30 controls using a variety of event segmentation measures. I found specific deficits in event boundary detection and hierarchical organization of events for the intellectually disabled group. However, results suggest that hierarchical organization is task-dependent. Because the event segmentation task accounted for differences in general cognitive ability, I propose the event segmentation task as diagnostic method for the need for support in executing assembly tasks.

Based on these three studies, I argue that the Event Segmentation Theory offers a framework for assessment and assistance of important attentional, perceptual, and memory processes related to assembly tasks. I demonstrate how practical applications can make use

of this framework for the development of new computer-based assistance and training systems that are tailored to the users' need for support and improve their quality of life.

Abbreviations

A'	sensitivity, non-parametric calculation of the Signal Detection Theory measure
ANOVA	analysis of variance
AR	Augmented Reality
B"	response bias, non-parametric calculation of the Signal Detection Theory measure
C	coarse
e	exponential function
eCALC	exponential function calculation
F	fine
HA	hierarchical alignment
HE	hierarchical enclosure
IBES	instructions based on event segmentation
ID	intellectual disability
ms	milliseconds
PE	position errors
s	seconds
SA	segmentation agreement
SD	standard deviation
SE	sequence errors
SRT	simple reaction time
VR	Virtual Reality
WMT-2	Wiener Matrizen Test-2
*	significant on the 5% level
**	significant on the 1% level

Content

1 Introduction

Assembly workers rely on instructional support in order to execute their complex manual tasks successfully, e.g., they have to be acquainted with the correct assembly sequence and detect potential errors. Paper instructions are the most prevalent source of information in order to support workers' cognitive processing during task execution. However, the trend is towards replacing them with technologies based on advanced human-machine interaction. In contrast to conventional means, computer-based assistance and training systems are promising because they provide intuitively comprehensive and interactive information and feedback. Furthermore, they reduce the development time for authoring new instructions because they quickly and (semi-)autonomously adapt to frequent changes in assembly tasks with the help of advanced algorithms. In this thesis, I argue that these systems have to be tailored to the human cognitive processes to generate appropriate support and prevent negative consequences.

In order to ensure the adaptivity of systems to human users, I aim at scientifically investigating the underlying cognitive processes involved in executing assembly work. Specifically, the Event Segmentation Theory offers a framework that deals with the processing of dynamic information like the sequential manual tasks in assembly work (Kurby & Zacks, 2008; Zacks et al., 2007). The theory explains the perceptual organization of an ongoing activity into meaningful events. The corresponding classical event segmentation task (Newtson, 1973) involves watching an online video and indicating the end of events with key presses. The task's output is the hierarchical organization of object-based coarse events and action-based fine events which gives insights into various cognitive processes (Radvansky & Zacks, 2014).

Until now, the cognitive processes in the context of assembly work in automotive industry are not well understood. How do daily working experiences lead to certain cognitive structures and long-term memory representations and how do they, in turn, influence learning of new assembly tasks? I will present an instruction creation paradigm and a research tool with which users semi-automatically create instructions based on a video. Assessing underlying cognitive structures with the help of this new software tool is theoretically and practically relevant; for instance, design suggestions for computer-based assistance and training systems can be derived.

Besides specific expertise in a domain like automotive industry, the general cognitive ability of workers is important, too, for instance, in the context of special assembly workshops for intellectually disabled people. There is a lack of a comprehensive assessment on their cognitive processes and abilities. Therefore, their daily work environment contains mostly simple and repetitive tasks. What is their actual cognitive potential to successfully execute structured manual tasks and are they able to perform more complex assembly tasks? In this thesis, I will use the Event Segmentation Theory to investigate the mentioned open scientific

questions concerning assembly workers at various levels - experts, new hires, and intellectually disabled people. Once I filled the knowledge gaps with respect to cognitive processing and need for support, I will discuss tailored assistance and training means.

This thesis has the following outline. Chapter 2 describes human-machine interaction at assembly workplaces and states the motivation for using Event Segmentation Theory from an applied point of view. In Chapter 3, I will describe the theoretical background of event cognition research and introduce the theoretical and methodological aims of this thesis. Next, the research questions will be elaborated in the empirical part in Chapters 4, 5, and 6. The resulting findings are summarized with respect to their contribution to the theoretical background of event cognition research as well as to their practical implications in Chapter 7.

2 Human-machine interaction at assembly workplaces

2.1 Assembly work

Manual assembly plays an important role in the industrial production of large, medium, and small sized companies that manufacture complex products like cars, machine parts, and electronic devices (Richardson, Jones, Torrance, & Baguley, 2006).

In this chapter, I will introduce assembly work in two selected domains, namely, automotive manufacturing and workshops for adapted work. I will present challenges with respect to assistance and training appropriate for workers and I will describe two open questions from an applied point of view.

2.1.1 Automotive manufacturing

Employees of automotive companies have mixed levels of expertise. Some manual workers are highly experienced in assembly operations; some are inexperienced first-time employees or seasonal work force. At production lines, assembly processes are predefined and rigorous. For instance, workers have to assemble a car door at a special assembly station in a strict sequence of assembly steps under specified time constraints with the help of parts and tools, i.e., car door windows, proper screws, and screwdrivers.

With increasing demand for individualization, the variety of car models and variants that are produced within the same factory is increasing. Workers have to know different assembly procedures for dissimilar cars, e.g., cars with power window versus conventional power lifter. Additionally, the product life cycle of a particular car model gets shorter requiring workers to frequently update their knowledge of assembly procedures.

Figure 1: Manual assembly workplaces in automotive manufacturing: Production line for assembly of car engines (left) and work station for assembly at the car shell (right). Source: Adam Opel AG.

It becomes clear that appropriate employee training is crucial to enable workers to cope with these varying and continuously changing manual assembly tasks. In order to prepare for the introduction of a new car model, workers go through repeated practice sessions for acquiring

the required procedures for the upcoming task with the help of pre-series hardware prototypes and paper instructions (Hermawati et al., 2015). This so-called hardware-based training is beneficial because workers can get acquainted with the real assembly procedures without much time pressure while interacting with qualified trainers in a safe environment.

Disadvantages of hardware-based training include the restricted extent of practicing different variants because producing appropriate pre-series hardware is costly, the limited number of training repetitions due to effortful assembly, disassembly and re-assembly, a late onset of training because of late availability of hardware, and partly incomplete or not up-to-date training because car characteristics may have been changed even shortly before start of production (Gorecky, Mura, von Falkenhausen, Apold, & Arlt, 2013; Hermawati et al., 2015). One can overcome these limits by using computer-based assistance and training (Malmsköld, Örtengren, Carlson, & Svensson, 2007b; Moskaliuk, Bertram, & Cress, 2013) that will be described in the next Section 2.2.

To sum up, the presented automotive domain constitutes a very dynamic production environment because the assembly tasks keep on changing continuously and the work force is quite heterogeneous consisting of both experts and novices (Gorecky et al., 2013; Hermawati et al., 2015). The complex working environment is amplified by upcoming developments like demographic change, an internationalized work force, and rapidly evolving new technologies which together challenge workers' flexibility of continuously coping with new tasks. Traditional hardware-based training goes along with a number of shortcomings.

A part of assembly tasks for the automotive industry is completed in workshops for adapted work which will be described in the next section.

2.1.2 Workshops for adapted work

Workshops for adapted work (or, formerly, sheltered workshops) provide work places for people with different types and degrees of physical and intellectual disabilities. It is estimated that 2 to 3 million (Europe) and 135.000 persons (USA) with a variety of physical, mental, or psychological deficits are occupied in workshops (European Association of Service Providers for Persons with Disabilities [EASPD], 2012a; Migliore, 2010) executing activities like assembly tasks and others (Dulaney, 1998; Migliore, 2010; Visier, 1998).

Workers with intellectual disability can have varying deficits in reasoning, problem solving, planning, abstract thinking, judgment, academic learning, or learning from experience (American Psychiatric Association [APA], 2013). However, there are no strict inclusion criteria for working in workshops and no unified definition of disability (EASPD, 2012a). Most workers have an intellectual disability with onset during developmental period (APA, 2013), they never worked at the regular job market, and entered workshops for adapted work directly after special school.

Figure 2: Manual assembly workplaces at the workshops for adapted work: Production line (left) and single assembly workplace (right). Source: Westpfalz-Werkstätten im Gemeinschaftswerk Pfalz.

Workshops can be a good option for disabled persons compared to open labor market because workshops aim at fostering workers' personal growth through manageable, interesting, and qualifying work activities as well as providing a fitting task to the worker (Migliore, 2010; Tomporowski & Hayden, 1990). In practice, workers can choose from a wide range of activities and receive an individually adapted work place as far as necessary and possible. For instance, assembly work places may be extended by additional installations like gripper arms so that individual motor deficits can be addressed. However, the tasks themselves are mostly highly repetitive (Dulaney, 1998); simple routines prevent mental overload, but they also lead to monotony and missed chance for personal growth. On the other hand, instructing more complex activities requires appropriate support incorporating the individual potential and needs of each worker. However, resources for an elaborated cognitive assistance through supervisors helping each worker individually are very restricted (EASPD, 2012b). These limits can be addressed with the help of computer-based assistance and training, for instance, interactive instructions (Korn, Schmidt, & Hörz, 2013b), which will be described in more detail in the next Section 2.2.

To sum up, the presented domain of workshops for adapted work represents a safe production environment for a highly heterogeneous group of intellectually disabled people with varying cognitive deficits. Due to lack of resources, workshops provide simple and repetitive tasks with risk of monotony, boredom, and missed opportunity for personal growth on the part of the workers.

2.2   Assistance and training means

In the following sections, I will demonstrate the potential of computer-based assistance and training systems to address the above introduced challenges. I will be comparing them to conventional ways of instructional support, i.e., paper-based assembly instructions.

2.2.1 Assembly instructions

In order to support successful execution of assembly tasks, the most predominant method is instructions which are "messages that guide people to perform procedural tasks by

describing the steps or rules required for completing the task" (after Eiriksdottir & Catrambone, 2011, p. 750). Instructions in working context contain textual and graphical information and are communicated via paper manuals or by interacting with trainers and supervisors.

Instructional design was influenced by cognitive psychology (e.g., Gagne & Dick, 1983; Shneiderman, 1989); for instance, the Dual Coding Theory (Paivio, 1986) and the Cognitive Theory of Multimedia Learning (Mayer, Hegarty, Mayer, & Campbell, 2005) suggested that pictures together with textual descriptions are superior in communicating procedural information compared to pictures or descriptions alone (labeled as *multimedia effect*). Appropriate structure of instructions is important, irrespective of the medium being used for giving instructions. Event cognition research (refer to Chapter 3) informed instructional design about how to optimally organize manuals (Zacks & Tversky, 2003) and where to place pauses in instructional videos (Adamczyk & Bailey, 2004; Spanjers, van Gog, & van Merriënboer, 2010; van Gog, Paas, & Sweller, 2010).

However, usage of paper-based assembly instructions entails shortcomings; they are inaccessible for non-native speakers, persons with a reading difficulty, or intellectually disabled persons, especially, if graphical information is not self-explaining. Furthermore, paper-based assembly instructions are not interactive and do not provide performance feedback, for instance, when an error occurs. One-to-one supervision by trainers is a suitable alternative but highly limited by available resources. The next Section 2.2.2 introduces computer-based methods for assistance and training addressing these limitations.

2.2.2 Computer-based systems

Promising ways of assistance and training may be through application of new technologies (Wehmeyer, Smith, & Palmer, 2004). Intelligent, context-aware systems autonomously provide workers with online virtual instructions directly into or close to the work space, e.g., on a monitor mounted at the workplace (Gorecky, Campos, & Meixner, 2012), through optical projection into the work space (Korn et al., 2013b) or by displaying information on a head-mounted display (Bleser et al., 2015). Dependent on the current task, the current assembly step, and the correctness of execution, workers receive appropriate guidance and, if necessary, correction. This context-aware enrichment of the real world with virtual information at the appropriate time and place is called Augmented Reality (Azuma, 1997); in contrast, the notion Virtual Reality is used when virtual information is displayed without such a connection to the current, real-world context, e.g., a simulation of the assembly task on a monitor (Brough et al., 2007). Two exemplary systems will be illustrated in the following.

A virtual training system supplementing hardware-based training

Workers may practice new assembly tasks with the help of a virtual training depicted in Figure 3 and Figure 4 (Gorecky et al., 2013). Users of this system receive a realistic simulation of the car model on a monitor and complete the assembly procedure sequentially

by moving car objects to the correct assembly position using their hand motions. Similar approaches have been reported elsewhere (Brough et al., 2007) and are available for purchase (e.g., from LivingSolids, Magdeburg, Germany; Cortona3D, Dublin, Ireland).

Figure 3: Virtual simulation of a car assembly for training purposes. Source: DFKI GmbH.

Traditional hardware-based training (see Section 2.1.1) can be enhanced by such a virtual training approach (Gorecky et al., 2013) providing the following benefits.

- Virtual training promotes acquisition of declarative knowledge with the help of realistic simulations (Ericsson, 2008; Malmsköld et al., 2007). In contrast to hardware-based training, it is not restricted to selected car variants but can include all models and variants that need to be learnt (Gorecky et al., 2013).

- Within virtual training visual, haptic, or auditory performance feedback enhances learning (Ericsson, 2004).

- It involves the opportunity for repeated practice (Ericsson, Krampe, & Tesch-Romer, 1993). In contrast to hardware-based training, there is no effortful disassembly thereby reducing the cost of hardware material wear-off (Gorecky et al., 2013).

- It may incorporate different difficulty levels ensuring adaptation to learning progress and continuous, effortful experiences (Charness, Kelley, Bosman, & Mottram, 2001; Sonnentag & Kleine, 2000).
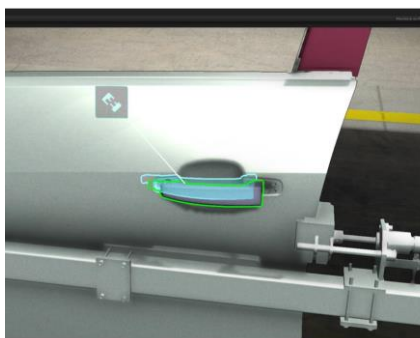
Figure 4: Graphical user interface of virtual training system: The correct placement of the door handle is indicated by (green) borders. Source: DFKI GmbH.

An Augmented Reality-based assistance system

Another way of computer-based support is virtual step-by-step instructions. Workers receive virtual information on how to execute the assembly while working on an assembly workplace. Gorecky, Campos, and Meixner (2012) presented such an Augmented Reality-based workstation (see Figure 5). The system previously *learned* the assembly workflow and now provides appropriate step-by-step guidance as well as monitors the assembly steps carried out by the worker to enable direct feedback to the current work. Thus, dependent on the workers' actions, the correct next assembly instruction or an error notification can be given.
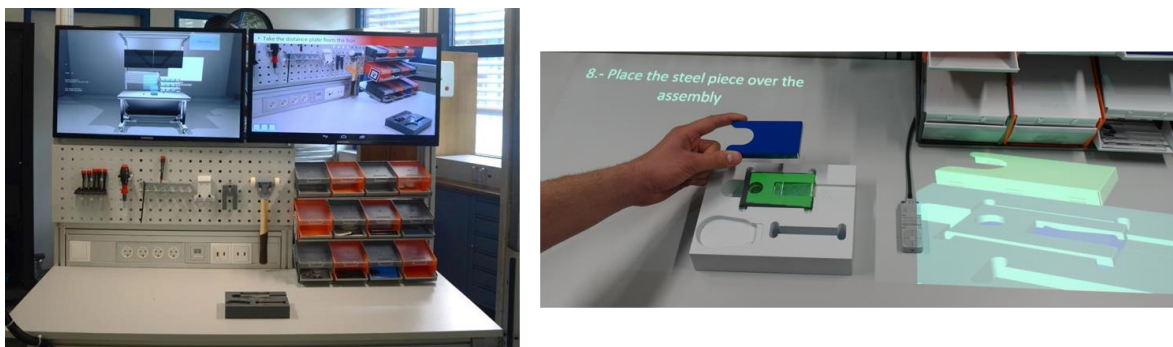
Figure 5: Manual assembly work station: Step-by-step instructions by virtual animations on a monitor (left) or projections into the workspace (right). Source: DFKI GmbH.

This computer-based assistance system and similar approaches (Bleser et al., 2015; Goto, Uematsu, & Saito, 2010; Henderson & Feiner, 2011; Korn et al., 2013b; Nilsson & Johansson, 2006; Petersen & Stricker, 2012; Stork & Schubö, 2010; Webel et al., 2011) offer the following benefits (see also Biocca, Tang, Owen, and Xiao (2006); Henderson and Feiner (2011)):

- Augmented Reality-based assistance involves in-situ, online instructions supporting attention allocation to the right task aspects (Stork & Schubö, 2010) and is usable for different people, e.g., intellectually disabled persons (Korn, Schmidt, & Hörz, 2013a).

- Augmented Reality-based assistance incorporates textual description together with dynamic, graphical animations beneficial for understanding and learning (Ainsworth & VanLabeke, 2004; Mayer, 2005).

- Within an Augmented Reality-based assistance different visual, haptic or auditory feedback can be immediately given on correct assembly performance (Ericsson, 2004; Webel et al., 2011).

- It covers different presentation modes, e.g., it can incorporate more or less detailed instructions (Eiriksdottir & Catrambone, 2011).

2.2.3 Trend towards adaptivity of computer-based systems

Despite the benefits of computer-based assistance and training systems, they mostly provide the same assistance and training for all users disregarding their individual characteristics. In contrast, *adaptive* assistance and training systems (Evenson, Rheinfrank, & Dubberly, 2010;

Jipp, Wagner, & Badreddin, 2008) are able to sense the individual user's potential and need for assistance. Adaptive systems adjust their interactions to individual *traits* like level of expertise, manual skills, motivation, learning strategies, and cognitive abilities. Furthermore, they address relevant *states* like increased familiarity with a task (Vicente & Rasmussen, 1992) and mental overload (Parasuraman, 1990). Whereas the above presented systems are able to adapt to the assembly task in question, true adaptive systems address the unique particularities of each individual user, too.

There are good reasons for the development of such adaptive systems. Kalyuga, Ayres, Chandler, and Sweller (2003) demonstrated that instructional support which helps inexperienced users may not be beneficial for participants with high expertise, or even deteriorate their performance. This negative consequence of missing adaptation has been noted as expertise reversal effect. In addition, motivation can be negatively affected if experienced workers feel not acknowledged or even patronized receiving too detailed directions. Furthermore, lack of adaptivity leads to exclusion from usage for some groups (Wobbrock, Kane, Gajos, Harada, & Froehlich, 2011), e.g., some intellectually disabled persons find these systems too difficult to use (Korn et al., 2013a). Finally, inadequate assistance in companies leads to a waste of financial and personal resources.

Adaptive systems show potential to overcome these negative effects. However, in order to be able to provide tailored assistance (Evenson et al., 2010) two questions have to be elaborated.

1. **What are relevant work-related user characteristics and how to integrate their assessment into a technical system? (User assessment)**

A number of psychometric assessments applies to performance in assembly work, i.e., test of perceptual-motor (e.g., Fleishman, 1972) and cognitive abilities (e.g., Formann, Waldherr, & Piswanger, 2011; Shepard & Metzler, 1971). Despite the right choice for a relevant test battery, additional aspects have to be regarded with respect to the integration in a computer-based system. The test has to be available as a stand-alone computer-based version so that a user can autonomously perform it. Furthermore, an automatic test evaluation should be provided. Consequently, vocational testing methods that rely on real world interactions cannot be used in this context, e.g., tower of Hanoi (Zook, Davalos, DeLosh, & Davis, 2004) or manipulating wires (Lienert, 1976). In contrast, computer-based intelligence tests could be used. However, Jipp et al. (2008) noted that integrating intelligence tests into computer-based systems is not suitable because they need detailed explanation, they incorporate artificial, ecologically invalid tasks, and their acceptance by users is low.

Because of the limitations of standard tests with respect to their integration into computer-based systems, appropriate assessment methods have to be defined. In this thesis, I will introduce and discuss a new diagnostic method as possible solution. Given the work-related user characteristics were assessed, the subsequent question is:

2. How should tailored instructions be designed? (Adaptation guidelines)

Informed by the results on a diagnostic assessment, support can be designed accordingly with the help of a variety of design means. For instance, Vicente and Rasmussen (1992) defined guidelines for interface design depending on user's familiarity with a situation. For assembly instructions, Eiriksdottir and Catrambone (2011) suggested adaptation by fading information dependent on expertise level. Similarly, a task can be presented in segments or at full length (Moreno, 2007). Other options are to manipulate the output modality, e.g., presentation of visual or auditory information, depending on the user's preference (Vignais et al., 2013). Furthermore, communication of task structure can be differently presented, either in a hierarchically structured format versus an unstructured step-by-step guidance depending on the personal goal (Zacks & Tversky, 2003). This short, incomplete list of examples indicates the abundance of design variance.

Because there is such a high range of possible adaptations that can be made in an adaptive computer-based assistance and training system, adaptation guidelines have to be elaborated. They inform system developers about the appropriate design decision depending on relevant user characteristics. In this thesis, I will discuss a selection of adaptation guidelines based on my empirical findings (Section 7.2).

2.3 Motivation for this thesis

In this thesis, I propose a theoretical framework to address the previously introduced questions concerning user assessment and adaptation guidelines, i.e., the Event Segmentation Theory (Kurby & Zacks, 2008; Zacks et al., 2007). Whereas the theory's applicability for assistance and training has been demonstrated for paper-based manuals, animations, and instructional videos (Adamczyk & Bailey, 2004; Spanjers et al., 2010; van Gog et al., 2010; Zacks & Tversky, 2003), I will further exploit it in the context of adaptive computer-based assistance and training systems for assembly work. Thereby, the theory will be used as a comprehensive framework for analyzing cognitive processes which are related to perceiving, understanding, learning, and executing assembly tasks. I will demonstrate its following benefits:

- It prompts a diagnostic assessment method to investigate assembly work-related cognitive ability. Thereby, the diagnostic method is easy to integrate in computer-based systems and applicable to a wide range of users.

- Empirical findings based on the Event Segmentation Theory offer suggestions for adaptation guidelines.

- The Event Segmentation Theory is the basis for further investigation of an increasingly dynamic human-machine interaction.

3 Event cognition literature

This chapter explains event cognition by introducing the perceptual organization of dynamic activities into events (Section 3.1). Further, I will review that the principle of segmenting dynamic content into events is not only relevant for perception, but also for actual performance (Section 3.1.5) and memory (Section 3.2). In Section 3.3, I will identify three gaps in the event cognition literature. They deal with understanding, practicing, and perceiving events, respectively.

3.1 Event perception

When observing a dynamic sequential activity, observers automatically divide it into meaningful, hierarchically structured events. Between two events, people perceive event boundaries (Zacks, Tversky, & Iyer, 2001). Despite the immediate comprehensiveness of this account, there are many cognitive processes involved in event segmentation. Their complex integration is described within the Event Segmentation Theory (Kurby & Zacks, 2008; Zacks et al., 2007).

3.1.1 Event Segmentation Theory

When watching a dynamic activity, a stream of information enters the human sensory channels. For instance, the observer perceives information like physical shapes, person's movements, and environmental features. This perception is guided by a working memory model of the event that contains a robust representation of the current event (e.g., "Closing the cover" in Figure 6a) influenced by prior knowledge, e.g., facts about human movement and own experiences (e.g., familiarity with assembling furniture, see Figure 6a). Based on the current observations, the person anticipates future actions. Such predictions are adaptive because they enable anticipatory behavior. An error detection mechanism monitors potential deviations between predicted future inputs and actual outcomes. As long as predictions are accurate, the current event model is valid. (Kurby & Zacks, 2008; Zacks et al., 2007)

However, prediction errors may increase due to meaningful changes in activities, i.e., time, location, character, intention, and causation (Zwaan, Langston, & Graesser, 1995). In assembly, a whole new object may appear (see Figure 6b). The more indices change at the same time, the more difficult it becomes to integrate this information into the current model (Huff, Meitz, & Papenmeier, 2014). The consequence is that the current working memory model of the event has to be reset and the observer has to establish a new event model (Figure 6b). This updating process incorporates incoming sensory information as well as existing knowledge from long-term memory (i.e., schemas; see Section 3.1.2) in order to build a stable working memory model of the new event. This stable representation means

that occlusions and interruptions do not result in changes of the model. The described transition between old and new event is perceived as an event boundary. Event boundaries are distinct points in time that go along with a higher attention level and go along with an elaborated long-term memory encoding (see Section 3.2.1). Therefore, they are important strategic points in the course of an activity. (Kurby & Zacks, 2008; Zacks et al., 2007)
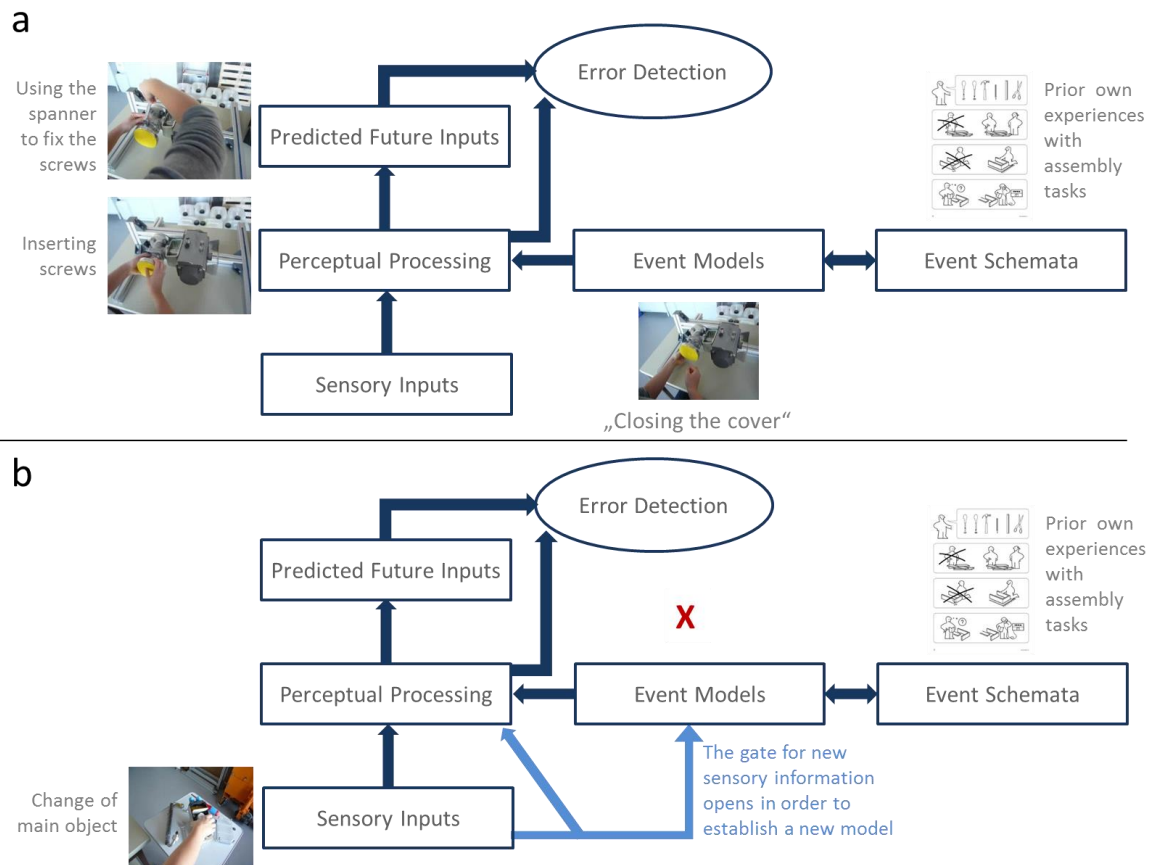
Figure 6: Event Segmentation Theory illustration (adapted from Kurby and Zacks (2008), Zacks et al. (2007)): (a) If predictions correspond with the outcomes in the video, no prediction errors are detected and the current event model "Closing the cover" can remain valid. (b) Updating the event model became necessary because of increased prediction error. For this purpose, sensory information may directly enter working memory in order to establish a new event model.

Furthermore, observers perceive dynamic activities at different grains simultaneously, so that at a specific point in time, several working memory models of events at different time scales can be active. That is, both a coarse-grained representation like assembling a table leg and a fine-grained representation like screwing the second screw of table leg may be active at the same time. The perception of fine event boundaries is due to bottom up processes, e.g., lower-level changes in movement (Zacks, Kumar, Abrams, & Mehta, 2009) and brief increases in prediction errors (Radvansky & Zacks, 2014; Zacks, Speer, Swallow, Braver, & Reynolds, 2007). Coarse event boundaries are perceived due to higher-level conceptual changes (Zacks et al., 2009), represent larger changes in goals (Zacks, Tversky, & Iyer, 2001), involve more physical change (Hard, Recchia, & Tversky, 2011), and go along with more sustained increases in prediction errors (Radvansky & Zacks, 2014; Zacks et al., 2007).
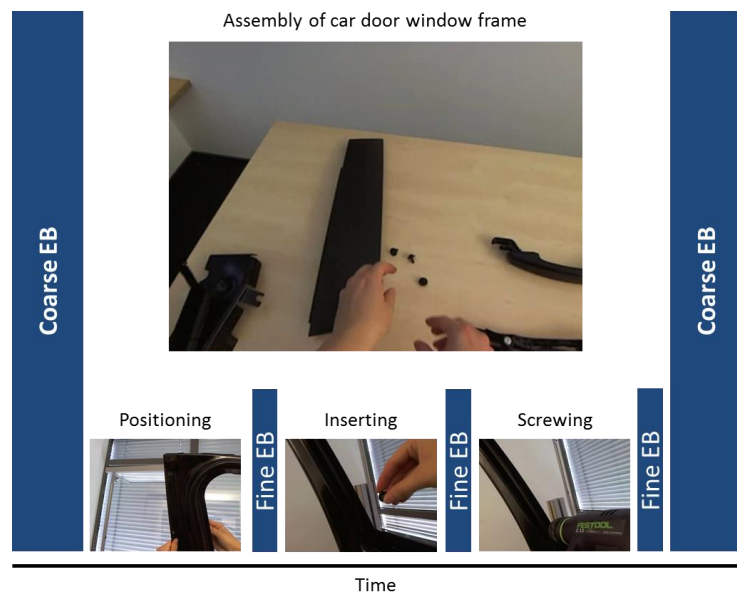
Figure 7: Hierarchical structure of an assembly task: Two coarse event boundaries (EB) frame the beginning and end of the superordinate event of assembling a specific car object. Between these coarse EB, there are several fine EB that frame several subordinate actions with this object.

"Assembly is a paradigm case of a complex event" (Zacks & Tversky, 2003, p. 89). Figure 7 illustrates the relationship between working memory models of both coarse and fine events for an exemplary assembly task. A coarse event is represented by attaching a major part and fine events are depicted by orienting the part and attaching it with the help of screws (Daniel & Tversky, 2012). The sequence of an assembly task may be either strict or allow for variations (Zacks & Tversky, 2003). Yet, the domains presented within this thesis, i.e., automotive manufacturing and workshops for adapted work, require workers to carry them out in a strict sequential order.

The relation between fine and coarse events has been described by the concepts of hierarchical alignment and enclosure (Hard, Lozano, & Tversky, 2006; Zacks & Tversky, 2001). According to these concepts, it is assumed that several fine events group together and precede a common coarse event boundary as outlined in Figure 7. Hierarchical alignment and enclosure will be described in more detail (Section 3.1.4).

3.1.2 Cognitive functions linked to event segmentation

The last section demonstrated that, in the context of event segmentation, a variety of cognitive processes has to be integrated in order to make sense of an activity, i.e., attention, visual perception, working memory, and long-term memory. Figure 8 summarizes them. In this section, I review the importance of these separate, cognitive functions for performance in assembly work.
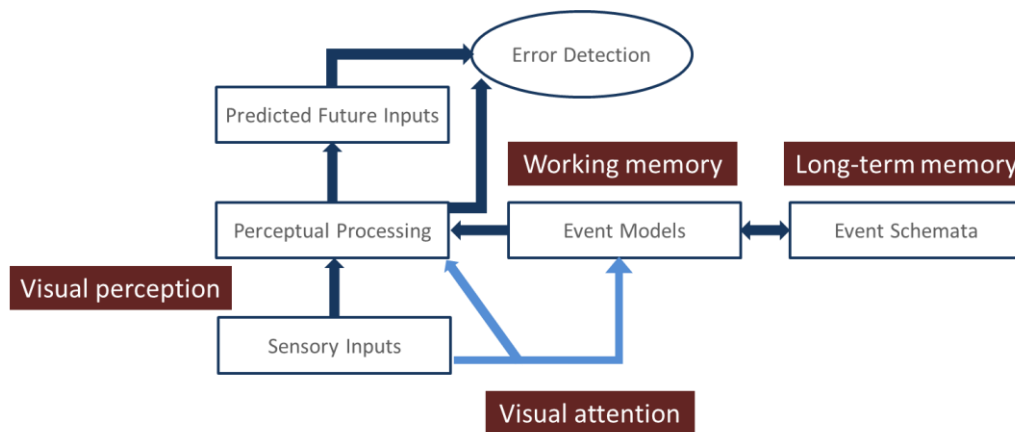
Figure 8: Cognitive functions related to processes during event segmentation when watching dynamic activities.

Visual attention

Visual attention is crucial for event segmentation, i.e., when an old event model was reset (Huff, Papenmeier, & Zacks, 2012). In this case, new incoming sensory information has to be brought into focus in order to establish a working memory model of the new event.

Visual attention is, furthermore, needed when executing assembly work (Stork & Schubö, 2010), for instance, when the assembly requires a new screw or when the worker has to allocate attention to the right box of screws. Thus, attentional processes are important both for successful assembly execution and in event segmentation.

Visual perception

When segmenting an ongoing activity, it is especially important for an observer to detect changes in movement in terms of positions, velocities, and accelerations of visual objects (Radvansky & Zacks, 2014). Motion changes yield to increases in prediction error (Zacks et al., 2007) and imply perception of event boundaries (Zacks et al., 2009; Zacks, 2004).

In assembly work, the importance of visually perceiving changes within a task has been acknowledged by the Situation Awareness Theory (Endsley, 2013), for instance, when workers perform a function test after assembly and interpret whether a component's movement is correct. Furthermore, perception is important because the execution of manual actions involves a constant interaction between both perception and action (Stork & Schubö, 2010).

Working memory

The central role of the working memory for event segmentation has been described in detail above (Section 3.1.1). Cognitive frameworks for assembly work stress the working memory's importance for successful task execution as well (Endsley, 2013; Richardson & Ball, 2009).

Long-term memory

Besides the described bottom-up processes like detection of movements, top-down processes from long-term memory are also important in event segmentation. They guide the perception of event boundaries by making prior experiences and higher-level concepts of typical situations available. These generalized long-term memory representations are known as scripts (Schank & Abelson, 1977) or schemas (Brewer, 1981). They support encoding and retrieval of specific event models (also termed as situation models), i.e., a situation or episode that a person actually perceived or experienced.

Long-term memory scripts for assembly tasks contain knowledge about the general sequential structure and the typically involved actions and objects. Situation models in the context of assembly are supposed to contain the exact assembly sequence, all involved objects, the local context, instruments, conditions, and consecutive context (Malmsköld, Örtengren, Carlson, & Svensson, 2007a).

From the previous paragraphs, it became clear that cognitive processes required for event segmentation are also needed for successful assembly task execution. The close relationship between both perception and actual execution will be further demonstrated in Section 3.1.5.

3.1.3 Assessing online event representations

So far, we saw how observers automatically divide dynamic activities into events during *online* perception of these activities, i.e., in the course of watching them. In order to assess the event boundaries in the classical event segmentation task (Newtson, 1973), subjects watch a video depicting a dynamic activity. At the same time, they indicate the end of one and the beginning of the next event by pressing a key button, respectively (Figure 9). The event segmentation task is executed two times; subjects segment both fine- and coarse-grained meaningful units (Newtson, 1973; Zacks et al., 2001). In Section 3.2.2, I will contrast this online event segmentation with the state of the art in *offline* event segmentation.
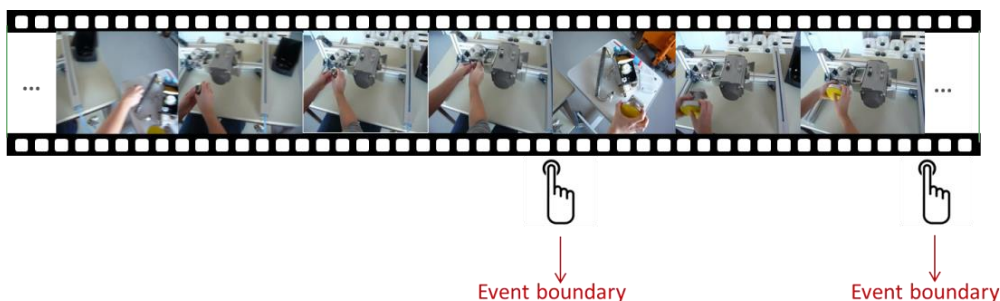
Figure 9: Classical event segmentation task: People watch a video and, thereby, divide the shown activity into meaningful events by pressing a key button. The task's output suggests perception of event boundaries.

In the original online event perception studies by Newtson (1973), videos depicted everyday activities like answering a telephone or setting a table. Other authors used a range of additional stimuli, for instance, simple moving dots (e.g., Maguire, Brumberg, Ennis, &

Shipley, 2011; Zacks, 2004), different assembly tasks (e.g., Zacks & Tversky, 2003), actions within a virtual environment (e.g., Radvansky & Copeland, 2006), or long visual narratives in terms of movies and sitcoms (e.g., Huff et al., 2014; Zacks, Speer, & Reynolds, 2009). Besides watching videos, online event representations can be assessed in verbal stories, either when reading (e.g., Zwaan et al., 1995) or listening (e.g., Whitney et al., 2009). Thus, the principles of event segmentation apply to varying materials where a stream of information is meaningfully structured.

A characteristic output of an online event segmentation task, i.e., key presses across time, is shown for an exemplary subject in Figure 10. The illustration depicts both fine- (Figure 10a) and coarse-grained segmentation (Figure 10b). In order to quantify online event segmentation, these key presses can be further analyzed. I will review different measures of the event segmentation behavior in the following section.

3.1.4 Measures of event segmentation behavior

Segmentation data have been analyzed by a variety of different measures treating time as either discrete or continuous variable. Treating time discretely involves the binning of data into intervals where each bin is interpreted as perceived event boundary if it contains a key press (Zacks et al., 2001). Prior work has predominantly used these 1-s bins. The advantage is that the statistical analyses are easily understood (Zacks et al., 2001). In contrast, treating time continuously retains all information (Royston, Altman, & Sauerbrei, 2006) without requiring arbitrary choice of a bin size (Zacks et al., 2001). Each person's key press is an estimate of a perceived event boundary distributed as a Kernel density function around the key press (Papenmeier, 2014). The Kernel density function is shifted to the left in order to account for the delay between event boundary perception and key press which has been set to approximately 1 s (Huff, Papenmeier, & Zacks, 2012). Thus, the underlying theoretical assumption of the continuous analysis is a probabilistic relationship between empirical key press and actual event boundary perception (compare Figure 10e and Figure 10f). In the discrete analysis, the distinction between key press and event boundary is nonexistent (compare Figure 10a and Figure 10b).

Summing up all individual segmentation plots, in both discrete (Figure 10c and Figure 10d) and continuous analyses (Figure 10g and Figure 10h), provides segmentation plots with characteristic variations including peaks that indicate chronological correspondences across individuals. In the discrete analysis, the segmentation plot displays a group histogram of identified event boundaries with binned time at the x-axis. Similarly, the continuous analysis depicts segmentation magnitudes (y-axis) across time (x-axis).

In the following sections, different discrete and continuous measures will be explained with the help of Figure 10. I will present the respective calculation, report exemplary values from existing literature, if available, and inform about the measure's usage in this thesis.
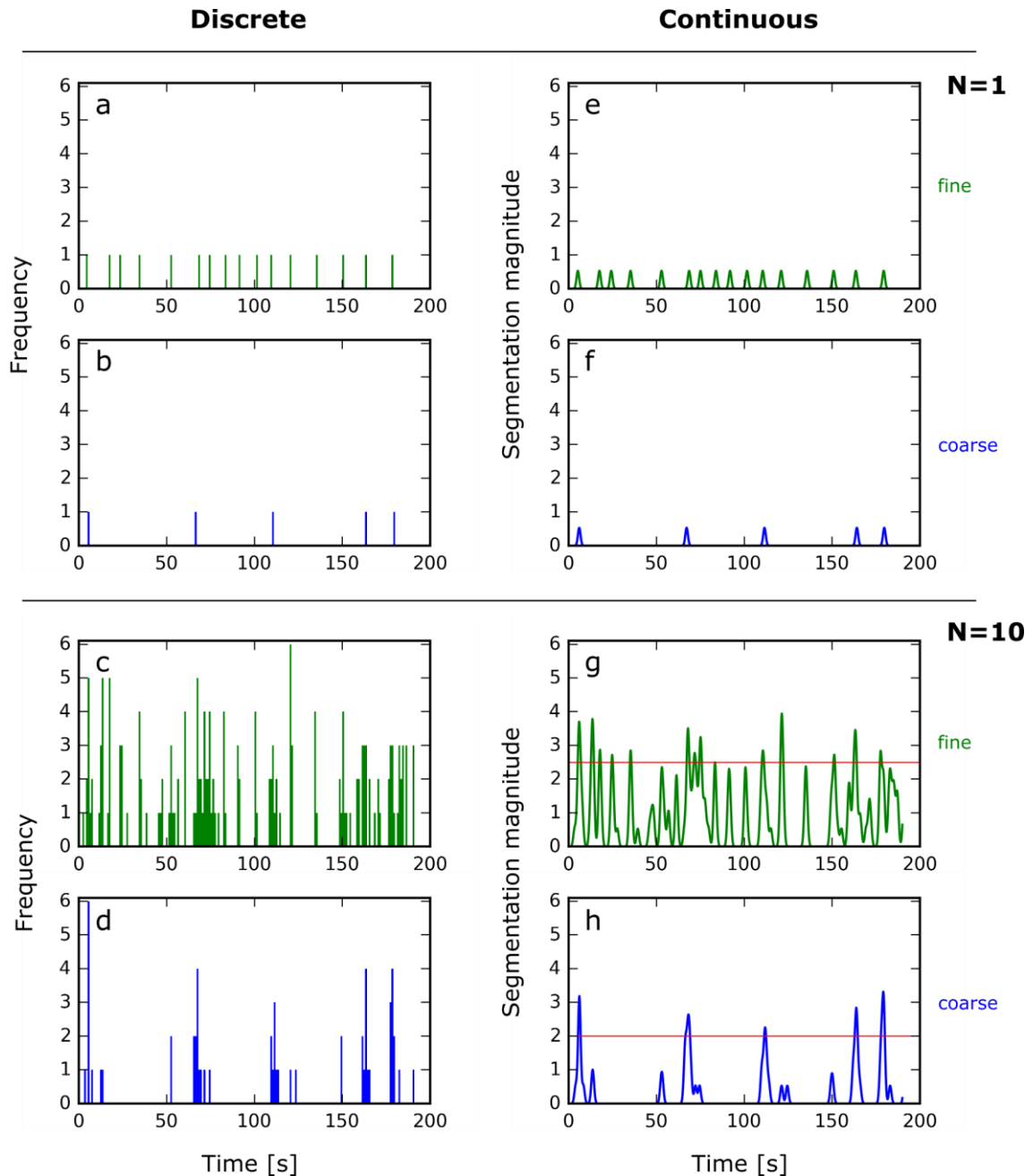
Figure 10: Output dependent on treating time as discrete (left) versus continuous (right) variable: Segmentation plots for one exemplary subject (upper plots) and across subjects (lower plots). Horizontal (red) lines in panels g and h illustrate critical cutoffs for determining significant event boundaries in the continuous analysis, respectively.

Number of events

For both fine and coarse grains, the sum of key presses is counted for each individual. As can be seen from Figure 10a versus Figure 10b, this typically results in more fine than coarse key presses. The respective mean event lengths, i.e., the time between two key presses, is the quotient of overall duration divided by number of key presses plus 1. Accordingly, the fine segmentation results in shorter events compared to the coarse segmentation.

The relation between numbers of identified event boundaries in fine to coarse condition has been further analyzed by calculating the ratio between them (Zacks et al., 2001). The value has been reported to be around 3 (Zacks et al., 2001) meaning that subjects defined three times more fine than coarse events (inspect the example in Figure 10 where the sample participant's ratio is 16 fine / 5 coarse events = 3.2). Additional relations between fine and coarse event boundaries have been analyzed by measures of hierarchical structure (see below for hierarchical alignment and enclosure). I applied the basic measures mentioned here, i.e., number of events, to describe event segmentation data in Experiments 1 and 3.

Significant event boundaries

As already noted, segmentation plots have characteristic peaks of chronological correspondence. Despite this graphical inspection, the continuous analysis offers to determine which of the characteristic peaks are significantly higher compared to peaks that occur by chance (Papenmeier, 2014). For this purpose, a critical segmentation magnitude is determined by simulating key presses under the null hypothesis that they were randomly distributed across individuals. Imagine that, even under the null hypothesis of randomly distributed key presses, peaks with a certain segmentation magnitude will occur by chance. They are called local maxima. Iterating this simulation, for instance, 1000 times, leads to 1000 local maxima in segmentation magnitude that get ordered according to their size. Depending on the choice of a confidence probability, e.g., 95% or 99%, a certain local maximum is kept as critical segmentation magnitude. This critical segmentation magnitude is unlikely to occur (e.g., probability of 5% or 1%) under the assumption of random key presses (Papenmeier, 2014). Consequently, segmentation magnitudes above the derived segmentation magnitude cut-off are defined as significant event boundaries. (Papenmeier, 2014) Exemplary critical cutoffs in segmentation magnitude are added as horizontal (red) lines in Figure 10g and Figure 10h.

With the help of this method, I determined significant event boundaries for groups of participants in Experiments 1, 2, and 3, respectively.

Segmentation agreement

Oftentimes, it is of interest to determine how similar the segmentation behavior is between single individuals and a comparison group, between different groups, or between single individuals. The extent to which an individual segmented in agreement with a comparison group can be calculated by a point-biserial correlation (e.g., Zacks, Speer, Vettel, & Jacoby, 2006). Concretely, for each time bin, it has to be checked whether there was a key press or not (i.e., a dichotomous variable) and the respective relative frequency of the comparison group has to be determined. To illustrate, data as displayed in Figure 10a are correlated with data from Figure 10c (note that relative frequencies are used). Furthermore, the point-biserial correlation has to be scaled in order to control for individual differences in the number of key presses (Kurby & Zacks, 2011).

This value was reported to be $r = .30$ when comparing individuals with dementia to a healthy comparison group (Bailey, Kurby, Giovannetti, & Zacks, 2013). In contrast, in the same study, the segmentation agreement of healthy individuals with the same comparison group resulted in a mean segmentation agreement of $r = .40$.

In addition to this segmentation agreement between individual and group, segmentation agreement between groups can be determined by correlating both groups' relative frequency histograms using the Pearson correlation, as reported in Zacks, Swallow, Vettel, and McAvoy (2006). In their experiment, while watching and segmenting a video of two simple objects, one group was told that these objects were moving intentionally; the other group thought they were moving randomly. In order to determine if their segmentation behavior was similar, the correlation of both groups' respective relative frequency data was calculated and resulted in values of $r_{fine} = .76$ and $r_{coarse} = .56$, respectively. To illustrate, imagine correlating two segmentation plots that are alike Figure 10c.

Furthermore, by using pair-wise correlations, i.e., Cohen's kappa, it has been tested if *pairs of individuals* chose event boundary locations that were more similar to each other than expected by chance (Zacks et al., 2006). In order to test for significance, bootstrap confidence intervals were constructed for each mean correlation. If the interval did not include 0, agreement for chosen pairs was significant. This calculation has been performed for pairs of individuals within the same and across different groups, respectively, in order to determine the within-group homogeneity as well as the agreement between pairs of different groups. Zacks et al. (2006) reported a mean pairwise k_{fine} of .19 when comparing two individuals from different groups, i.e., intentional and random group (see above). Respective coarse event segmentation resulted in mean pairwise $k_{coarse} = .08$ between groups.

Based on the calculation of these different kappas, the differences between agreement within the same group and across different groups can be analyzed. This test reveals if persons from the same group are more similar to each other than persons from different groups, and vice versa. Again, bootstrap confidence intervals for these differences inform about significance of the difference. For instance, if the difference in agreement between pairs of individuals from the same group versus different groups is not significantly different from 0, this means that the persons within the same group are not more similar to each other than to the persons of the other group.

Finally, in addition to these correlational analyses based on binned data, a continuous analysis method to determine segmentation agreement has been proposed as well. Based on two segmentation plots produced by two different groups (for instance, imagine two segmentation plots similar to Figure 10g), Papenmeier and Sering (2014) suggested subtracting these segmentation plots from each other with respect to time in order to define the overlap of both groups. In order to further determine critical cutoffs for the *differences* in segmentation magnitude, they provided a similar simulation method as described above (see the section on significant event boundaries). That is, for both groups, random key presses can be repeatedly simulated and resulting differences in segmentation magnitude are

calculated and ordered. This time, two cutoffs, i.e., a local minimum and a local maximum value, respectively, are derived.

In order to determine different types of segmentation agreement, I used point-biserial correlations between individual and group in Experiment 3, pair-wise kappas in Experiment 3, correlations between groups in Experiments 1 and 3, and the difference method based on the segmentation magnitude of two groups in Experiments 1 and 3.

Hierarchical alignment

Hierarchical alignment is a measure to reveal the quality of hierarchical structure of an individual's segmentation. As indicated by Figure 11, it refers to the temporal closeness between perceived fine and coarse event boundaries (Zacks et al., 2001). High hierarchical alignment means that a coarse event boundary constantly goes along with the end of a respective fine event. Figure 11 illustrates good alignment indicated by five pairs of temporally close fine and coarse event boundaries. Zacks et al. (2001) defined both discrete and continuous analysis methods. Yet, the continuous alternative has been predominantly applied in the subsequent literature and will be described in the following.

First, the distances between each person's coarse event boundaries to the nearest fine event boundaries are determined and averaged leading to the *observed* average distance. Figure 11 displays five coarse event boundaries that have a nearest fine event boundary, respectively. The temporal distances between these pairs of time points are calculated ("d_1" to "d_5") and averaged. Next, an *expected* average distance under the assumption that coarse and fine event boundaries were independent is calculated (refer to Zacks et al. (2001) for a detailed discussion). The individual's alignment score is the difference between the expected and the observed average distance. The higher the value, the more hierarchically aligned a person segmented.
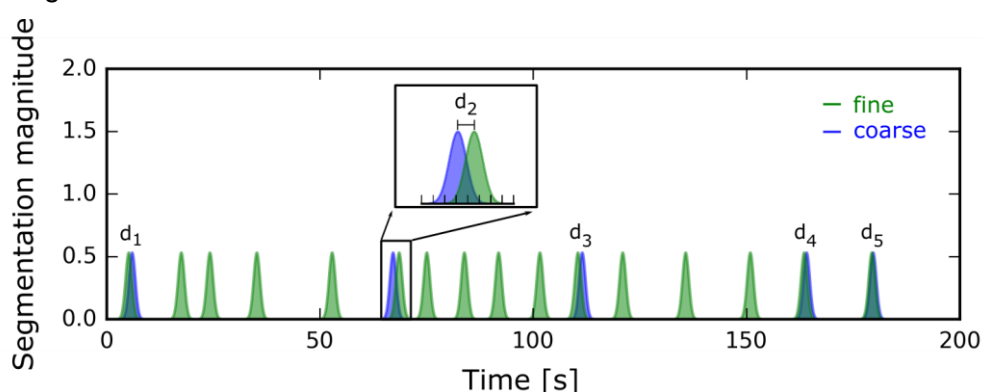
Figure 11: Temporal closeness between coarse and fine event boundaries: Based on distances between the coarse and the nearest fine event boundaries (d_1 to d_5) the observed average distance is calculated.

In previous work, mean observed versus expected distances lied around 1.7 versus 4.8 (Swallow, Zacks, & Abrams, 2009) and 2.8 versus 4.7 s (Zacks et al., 2001), respectively. In this thesis, the hierarchical alignment was calculated for segmentation data in Experiments 2 and 3, respectively.

Hierarchical enclosure

Another measure to reveal the ability to hierarchically structure is the hierarchical enclosure measure. It refers to the extent to which an individual segmented according to a "chunking pattern", i.e., fine event boundaries have to *precede* its corresponding coarse event boundary (Hard, Lozano, & Tversky, 2006; Zacks et al., 2001). For calculating the hierarchical enclosure score, the nearest fine event boundary for each coarse event boundary is determined like in the hierarchical alignment computation. Afterwards, there is a check if the nearest fine event boundary is temporally before or after its coarse event boundary. The hierarchical enclosure score is the proportion of the nearest fine event boundaries preceding the coarse event boundary in relation to all nearest fine event boundaries. Consequently, the enclosure value ranges from 0 to 1. In the example in Figure 11, four of five nearest fine event boundaries precede its respective coarse event boundary. This leads to an enclosure value of .80 and indicates a clear chunking pattern of the sample participant.

In the literature, enclosure scores have been reported to be between .40 and .67 (Hard et al., 2006). In this thesis, I investigated hierarchical enclosure in Experiments 2 and 3, respectively.

3.1.5 Event segmentation behavior and performance

In Section 3.1.2, I suggested an overlap between event perception and assembly performance, because the Event Segmentation Theory and literature on successful assembly execution share the required cognitive processes. In a recent study (Bailey et al., 2013), the close relationship between perception and performance was demonstrated empirically. Bailey et al. (2013) instructed participants suffering from Alzheimer's disease to perform both an event segmentation task and a naturalistic action task, i.e., packing a lunch box. Quality of event perception was defined as degree of agreement of an individual patient's segmentation with a healthy control group's segmentation concerning identification of event boundaries. The level of segmentation agreement measured by point-biserial correlation (see Section 3.1.4) predicted the Alzheimer patients' task performance.

Reciprocally, repeated execution, familiarity, and expertise also affect segmentation behavior. Graziano, Moore, and Collins (1988) showed that, compared to novices, experts segmented familiar material more coarsely (see also Schwan & Garsoffky, 2008). Furthermore, Zacks et al. (2001) found evidence for a positive relationship between task familiarity and amount of individual hierarchical alignment which indicated that if familiarity increases hierarchical structuring improves.

Finally, segmentation agreement is associated with memory performance (Kurby & Zacks, 2011; Sargent et al., 2013; Zacks, Speer, Vettel, & Jacoby, 2006). In their study (Sargent et al., 2013), the extent to which participants segmented everyday activities in agreement with the whole sample predicted how well participants recalled the shown activities afterwards.

Similarly, persons with an intellectual disability showed poorer memory performance for events dependent on their previous event segmentation ability (Zalla, Labruyère, & Georgieff, 2013).

These findings demonstrate that the event segmentation task measures the cognitive ability to structure activities into goals and sub-goals. This ability, again, is suggested to be the basis for further important functions, most importantly, action performance (Bailey et al., 2013). Carried over to assembly work, the event segmentation task could offer a diagnostic measure that supports assessment of cognitive potential of assembly workers to execute assembly tasks. I will investigate this question on the group of intellectually disabled workers from workshops for adapted work (Section 3.3 and Experiment 3).

3.2 Long-term memory for events

So far, the focus was on the Event Segmentation Theory as framework for perception of dynamic events. I presented how existing knowledge from long-term memory influences event boundary perception. Here, I focus on how new procedural knowledge finds its way into long-term memory.

3.2.1 Event boundaries as memory anchors

After persons watched or read about an activity one time, memory is generally better for event boundaries than for non-event boundaries (Lassiter & Slaw, 1991; Newtson & Engquist, 1976; Schwan & Garsoffky, 2004; Swallow et al., 2009; Zacks, Speer, et al., 2006). Because event boundaries go along with increased attention (Huff, Papenmeier, & Zacks, 2012), they are more likely to be encoded into long-term memory (Radvansky & Zacks, 2014). Furthermore, deletions, delays, or disturbances at the points of event boundaries are more detrimental for memory as compared to time points within event boundaries (Boltz, 1992; Schwan & Garsoffky, 2004).

Memory for fine events is more fragile than for coarse events in written and pictorial narratives (Bransford, Barclay, & Franks, 1972; Gernsbacher, 1985; Johnson-Laird & Stevenson, 1970; Treisman & Tuxworth, 1974). Memory performance is better for coarse information but it also takes more effort to recall it compared to fine information (Franklin & Bower, 1988 after Zacks, 2001). For instance, participants who were asked to memorize a previously read text, answered more slowly when they integrated coarse compared to fine events suggesting better processing for coarse information. Furthermore, fine events may be more similar and less distinct compared to coarse events (Hard et al., 2011; Radvansky & Zacks, 2014; Zacks et al., 2001; Zacks et al., 2009) which further results in differences between fine and coarse events concerning their memory representations.

3.2.2 Acquisition of new events

The Event Horizon Model (Radvansky & Zacks, 2014) describes the transition from current perception to long-term memory as well as retrieval from long-term memory. Online presentation of an event creates a long-term memory representation, called experience model (Radvansky & Zacks, 2014) or situation model (Zwaan et al., 1995). Thereby, several events of an activity are linked together by their causal relationships. Different events can be similar to each other and similarity influences ease of long-term memory retrieval. There is a facilitated retrieval of an event element, when this element is represented in multiple event models (Radvansky & Zacks, 2014). However, "when several event models are similar, accessing any specific event model is more difficult" (Radvansky & Zacks, 2014, p. 29). Thus, similarity of events affects retrieval performance from long-term memory.

The Event Horizon Model does not make predictions about retrieval after repeated presentations of events but focuses on one-time presentation. However, the key to acquiring new skills is indeed exposing learners to multiple challenging experiences with tasks summarized under the Deliberate Practice Framework (Ericsson et al., 1993). For instance, effortful, repeated practice is crucial for learning in chess (Charness, Tuffiash, Krampe, Reingold, & Vasyukova, 2005), music (Krampe & Ericsson, 1996), and in the workplace (Charness, Kelley, Bosman, & Mottram, 2001; Sonnentag & Kleine, 2000). However, it is not well understood yet how repeated presentation alters memory for events that may be more or less similar to each other.

The interrelation of both aspects, i.e., material differing in similarity and repeated presentations, has been focus of recent research. Reagh and Yassa (2014) compared conceptually different versus similar material and found that repetition enhanced discrimination only for conceptually different pictures. In their basic research study, participants were more likely to correctly detect a target picture when they saw it three times compared to only once. However, pictures that were similar to the targets but were not presented in the study phase (distractors) were more likely to be falsely identified as a target after three repetitions than after one repetition. This provides empirical evidence for a deteriorating effect of repetition depending on similarity of stimulus material.

Even if it is clear that fine and coarse events differ conceptually and lead to different memory representations, we do not know how memory performance for fine and coarse events will develop after repeated presentation. I will address this gap in this thesis (see Experiment 2).

3.2.3 Assessing offline event representations

Until now, it has been indicated that event segmentation plays an important role not only in online perception of activities but it is also important when processing information *offline* (Zacks et al., 2001; Zacks & Tversky, 2001), e.g., memorizing events (Black & Bower, 1979), planning an action (Hommel, 2004), communicating about events (Tversky, Zacks, Morrison, & Hard, 2011), and instruction making (Daniel & Tversky, 2012; Tversky, Zacks, Lee, &

Heiser, 2000). The methods used for assessment of offline events were free recall (Zacks et al., 2001; Zalla et al., 2013), recognition tests (Radvansky & Copeland, 2006; Swallow et al., 2009; Zalla et al., 2013), and picture sorting tasks (Zacks, Speer, et al., 2006) amongst others.

In order to investigate free recall data, Zacks et al. (2001) applied detailed language analysis and reported qualitative relations. For instance, they found that descriptions of coarse events predominantly entail objects whereas descriptions of fine events focus on verbs. They compared the offline with the online event descriptions and revealed a meaningful overlap. Other support for similarity between online perception and offline elaboration, i.e., memory performance (Kurby & Zacks, 2011; Sargent et al., 2013; Zacks, Speer, Vettel, & Jacoby, 2006), provides Section 3.1.5. However, no attempt has been made, so far, to quantitatively compare the exact time points of event boundaries between online and offline event segmentation. I will address this gap in this thesis (see the following section and Experiment 1).

3.3   Aims of this thesis

3.3.1 Research questions

Organizing dynamic activity into events is a cognitive activity that encompasses, uses, and influences all crucial processes of cognition, from online perception (Section 3.1) to offline elaboration (Section 3.2). Quantification methods have been used to investigate online event segmentation and the different measures have been summarized in Section 3.1.4. However, there has been no rigorous quantitative analysis of offline event segmentation which has been mostly analyzed qualitatively (Section 3.2.3). Since, in working context, workers process assembly tasks predominantly offline (Section 2.1), the classical, online event segmentation task might not result in ecologically valid event boundaries. Nevertheless, event boundaries from online event segmentation have been widely used for creating instructions (Section 2.2 and 2.2.1). I aimed at confirming quantitatively whether the online event boundaries correspond to the offline event boundaries. Therefore, Experiment 1 deals with the following question:

R1. Are event boundaries during offline event segmentation similar to event boundaries during online event segmentation? (Chapter 4)

Based on earlier studies demonstrating that online and offline processing have similar numbers of event boundaries and correspondent verbal descriptions (Section 3.2.3), I expect the exact time points of event boundaries to be the same. Empirical validation of this statement is necessary to check whether the online event segmentation task offers a valid way to investigate the structure of activities regardless of whether they are processed online or offline.

The event segmentation task yields fine and coarse event boundaries, respectively (Section 3.1.1), that are encoded in memory differently (Section 3.2.1). So far, memory for dynamic events has been tested only after one presentation (Section 3.2.1 and Section 3.2.2). Yet, repeated presentation of stimulus material is crucial for learning and repetitions change the basic memory processes (Section 3.2.2). An important but unanswered question is whether repetition benefits memory for coarse events more than fine events or vice-versa, or whether there is no difference between the two. In addition, expertise and familiarity influences event cognition (Section 3.1.5), so that memory processes after training could differ between experts and novices. Therefore, Experiment 2 aimed at answering the following question:

R2. How does memory for events develop when repeatedly practicing the sequence of events, both in novices and experts? (Chapter 5)

Finally, the event segmentation task offers an assessment method of event segmentation ability (Section 3.1.3) which is understandable by a wide range of persons including Alzheimer patients and intellectually disabled participants (Section 3.1.5). Furthermore, it was used to predict action performance (Section 3.1.5). In the light of the abundance of repetitive and monotonous assembly tasks in workshops for adapted work (Section 2.1.2), Experiment 3 will test, with the help of the event segmentation task, if cognitive potential of intellectually disabled persons allows them to perform more complex tasks. Therefore, my last research question is:

R3. Do the simple and repetitive assembly tasks offered at workshops for adapted work utilize the full cognitive potential of intellectually disabled persons? (Chapter 6)

3.3.2 Methodological aims

Besides these theoretical aims, I formulated the following methodological goals. First, there is no assessment tool available for providing event boundaries during offline event segmentation. Since instruction creation represents a way to investigate offline event representations (Section 3.2.3) and instruction creation is important for assembly work (Section 2.2.1), I pursued the following goal:

M1. Developing a tool for assessing offline event segmentation by using an instruction creation paradigm (Chapter 4)

Another open question is whether event segmentation measures including the ability to hierarchically structure events (Section 3.1.4) holds for groups with varying intellectual abilities, if so under what conditions. Therefore, I aimed at:

M2. Evaluating and refining existing event segmentation measures with respect to their suitability for intellectually disabled persons (Chapter 6)

4 Experiment 1: Offline event segmentation of assembly tasks

Since assembly workers process their tasks mostly offline rather than during online perception, this experiment investigates event segmentation during offline understanding of assembly tasks.

4.1 Introduction

The conceptual distinction between online and offline event segmentation has been noted by Zacks et al. (2001). Online event segmentation takes place fast, automatically, and in the course of perceiving an activity. In contrast, offline event segmentation takes place during elaboration of a task. This process is slower, involves no time constraints, and there is an explicit aim like planning future actions, understanding narratives, remembering past events, or creating an instruction manual.

In both online and offline processing of a task, people structure activities with respect to important, strategic points in time, i.e., event boundaries. Event boundaries during online event segmentation have been assessed with the help of the event segmentation task (Newtson, 1973). In contrast, there is no such method for assessing event boundaries during offline event segmentation. Rather, these event boundaries were derived indirectly from free recall data (Zacks et al., 2001; Zalla et al., 2013), recognition tests (Radvansky & Copeland, 2006; Swallow et al., 2009; Zalla et al., 2013), and picture sorting tasks (Zacks, Speer, et al., 2006) amongst others.

The event boundaries conveniently assessed by the classical event segmentation task have been used in the applied field in order to provide guidelines for designing instructions, e.g., how to sequentially structure them (Zacks & Tversky, 2003) or where to put pauses (Adamczyk & Bailey, 2004; Spanjers et al., 2010; van Gog et al., 2010). Hence, users received manuals structured according to event boundaries based on online event segmentation but they used them for offline elaboration of the task, i.e., during usage of instructions. This is appropriate if online and offline event segmentation lead to similar event boundaries; otherwise, there could be interferences.

Yet, there is evidence for a meaningful overlap between online and offline event segmentation. Zacks et al. (2001) compared them by collecting both online and offline event descriptions. Indeed, language analysis revealed that numbers of respective events were similar. They argued that the same cognitive structures guide both online perception and offline conception, i.e., scripts (Brewer, 1981; Schank & Abelson, 1977). Further support comes from research showing that performance in the event segmentation task is the basis for further functions like memory (Kurby & Zacks, 2011; Sargent et al., 2013; Zacks, Speer, et al., 2006) and action performance (Bailey et al., 2013). However, the quantitative overlap

of event boundaries was not tested yet. Finding similar locations of event boundaries would eventually prove that the online event segmentation task provides a valid way to investigate the structure of activities regardless of whether they are processed online or offline.

Until now, an important obstacle to a comparison between both types of event boundaries, i.e., online and offline, was the lack of an appropriate tool. More specifically, there is no method to assess exact locations of event boundaries during offline event segmentation. In the present experiment, I will introduce a tool that solves this problem, namely, the IBES tool (**I**nstructions **b**ased on **e**vent **s**egmentation). With the help of the IBES tool, participants use static frames of a video in order to design instructions. I use this instruction creation paradigm in order to detect event boundaries during offline event segmentation. I will compare this tool's output to event boundaries labeled during the classical online event segmentation task. The question is whether the participants' mental representation of the task assessed during instruction creation is similar to the automatic event perception processes involved in online watching of the video.

In this experiment, one group of participants performed online event segmentations for two assembly tasks using the classical event segmentation task. Another group of participants executed offline event segmentations for the same assembly tasks by the help of the IBES tool. I compared both groups' event boundaries.

4.2 Methods

4.2.1 IBES tool

Together with Nils Petersen, I developed the IBES tool in order to assess event boundaries during offline event segmentation. The IBES tool is released as freeware and is available at http://www.ict-cognito.org/demo. It is based on the approach for automatic task segmentation and instructions generation, described in Petersen and Stricker (2012). Nils Petersen was responsible for the software development of the IBES tool whose detailed technical specifications can be found in Mura, Petersen, Huff, and Ghose (2013). In the following, I provide an overview of the tool's characteristics.

Overall, this computer-based tool uses an instruction creation paradigm in which participants are asked to make instructions based on static frames of a task's video. In the first, most important step, participants have to define an appropriate structure for the task in question. Second, they choose illustrative static frames from the video in order to add them to their instruction manual. Third, they add textual descriptions. Fourth, the manual can be printed. Moreover, the IBES tool provides an output file that has a significant impact on psychological research. The output file called "results.csv" not only documents the instruction creation process but also records time stamps for the starting and ending frames of each assembly step. These time stamps can be further used for assessing event boundaries during offline

event segmentation. In the following, I will describe the four-step-workflow within the IBES tool in more detail.

On the start screen, the users are asked to segment the sequence of video frames into instructional steps (see Figure 12g). More specifically, the users choose segments from the stream of frames by mouse clicks. The chosen pictures and the corresponding frame numbers are shown in a small, transparent window (Figure 12c) and, additionally, amplified in a bigger window below the stream (Figure 12d). If the users hold down the left mouse key while moving over the stream of pictures, the big window shows a movie clip consisting of the marked pictures.
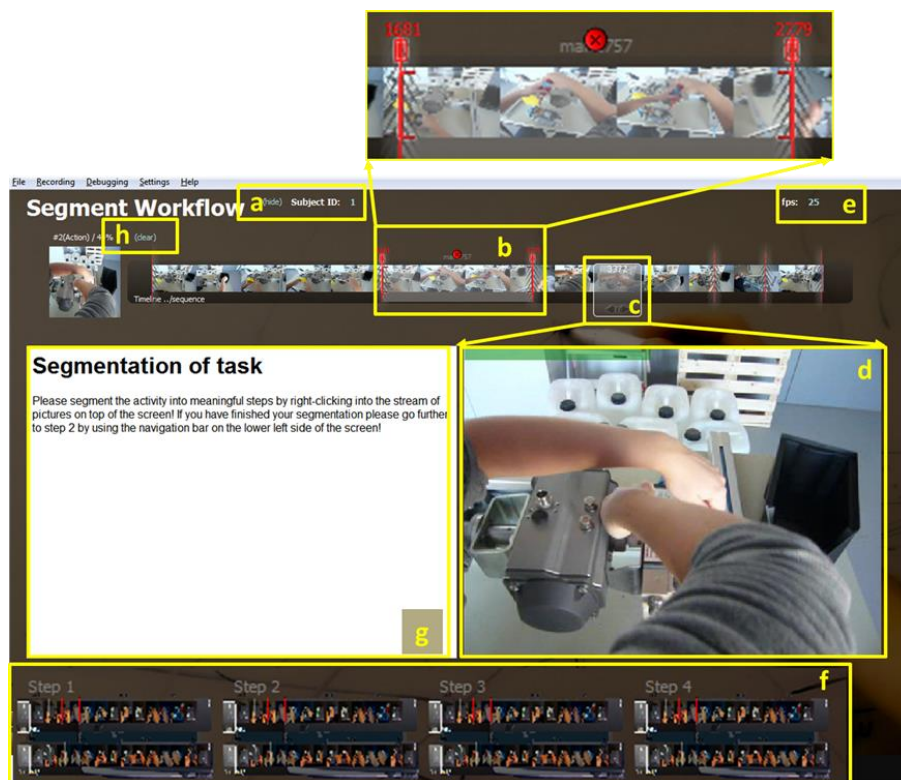
Figure 12: Screenshot of start screen where offline event segmentation takes place: (a) The subject identification may be entered and is hidden. (b) Start and end frames of an event are indicated by two (red) marks which can be used to drag the boundaries (additionally amplified on top of this figure). The chosen event can be deleted by clicking the (red) circle in the middle. (c) The white window highlights the current picture (d). The default frame rate value of 25 fps may be changed (e). The navigation bar (f) enables moving back and forth between the four steps. (g) A field for instructions or information for the user can be edited. (h) The complete segmentation may be deleted by clicking on "clear".

Specifically, for segmenting the stream of pictures, the (white) transparent window in Figure 12c has to be placed at the starting point of a new step followed by a right mouse click. A default time window (see Figure 12b) with two (red) marks appears when the mouse is moved a little above the filmstrip. Then the users adjust the end point by dragging the right (red) boundary to the appropriate end frame. The two (red) boundary-marks represent the start and end frames of an instructional step that should go into the instruction manual. The subsequent instructional step for the manual can start with the very next frame after the end frame of the preceding segment. However, if the immediate next frames are not meaningful, the start point can be moved forward until the next important step begins. The users may

delete a step by clicking on the cross displayed above the selection window (see Figure 12b). By pressing "Clear" on the upper left side of the screen (see Figure 12h) they may delete the entire segmentation.

After the segmentation is complete, the users have to navigate to the second step using the navigation bar (Figure 12f). As shown in Figure 13, each of the event segments chosen in step 1 of the IBES tool workflow appears in a separate row on a new screen (in Figure 13 nine steps are displayed for clarity). By default, each of the event segments is displayed as a sequence of eleven images. The users' task is to choose the essential and most representative pictures that have to be incorporated into the instruction manual by clicking on them. The users usually choose at least one picture from every event segment. Users may cancel their selection by clicking on the picture again.
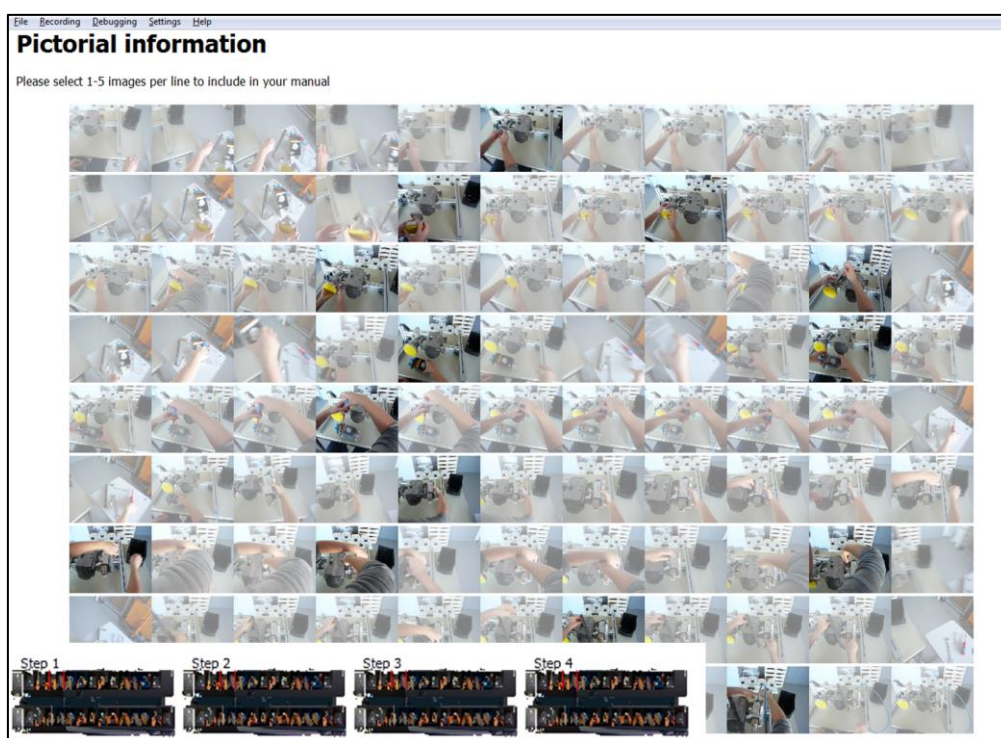
Figure 13: Screenshot for step 2 within the IBES tool: Users choose appropriate pictures representing each event. Pictures that have been chosen for the manual are shown more clearly than the rest.

In the third step within the IBES tool, subjects see their preliminary manual consisting of all instructional steps row by row along with their associated pictures (Figure 14). In this phase of the instruction design they can add textual descriptions for each step into the corresponding text box.

In step 4 within the IBES tool, the completed manuals are displayed and may be printed out. They are either ready for immediate use or users may manually add overlays, like arrows, boxes, circles, and so on.
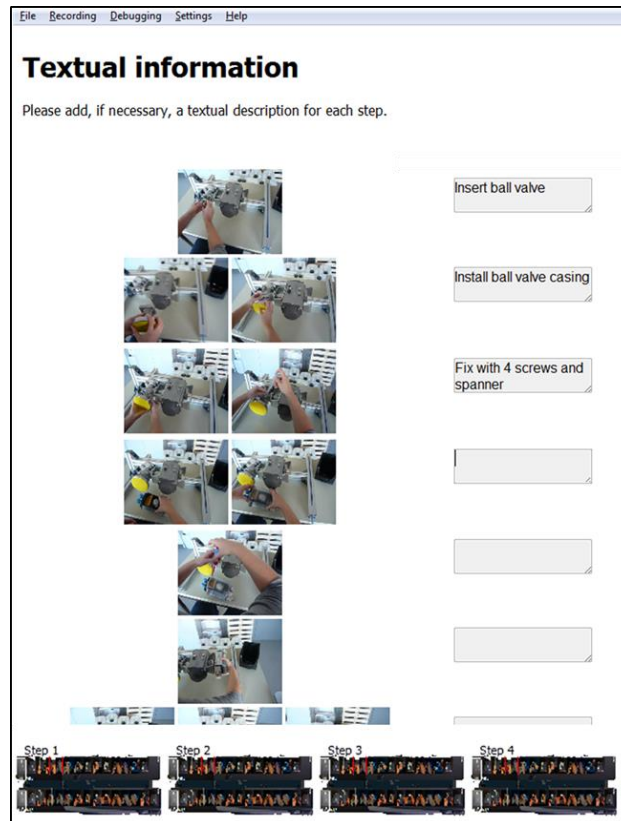
Figure 14: Screenshot for step 3 within the IBES tool: Users may edit textual description for each event.

4.2.2 Material

In both online and offline event segmentations, I used two industrial tasks in which the actor performs some manual operations. One task involved changing a notebook RAM and the other task involved assembling a pump system (further screenshots for the pump task may be found in Figure 6). The videos of the tasks were recorded from a first-person perspective (Figure 15). The notebook task took 1 minute and 12 seconds and the pump task took 3 minutes and 16 seconds.

Figure 15: Screenshots of videos depicting an installation of a new notebook RAM (a) and an assembly of a pump system (b).

4.2.3 Participants

In the offline event segmentation, 20 participants (average age of M = 25.1 years, SD = 1.9) including 11 male and 9 female students from the University of Kaiserslautern created manuals for both tasks with the help of the IBES tool.

For the online event segmentation, I recruited 22 new participants from the same university; twelve subjects segmented the video of the notebook task (6 female and 6 male with an average age of M = 25.5 years (SD = 1.9)) and ten subjects segmented the video of the pump task (4 female and 6 male with an average age of M = 24.8 years (SD = 2.5)).

4.2.4 Procedure

The participants in the offline event segmentation initially saw a video of the notebook task in order to become familiar with it. Then, they were introduced to the functionality of the IBES tool. They had to divide the whole task into steps that they thought will be "useful for giving instructions" by defining the start and end points of each instructional step, respectively. No time limit was given and participants had the opportunity to modify their choice of steps during segmentation. Afterwards, they sequentially assigned descriptive pictures and wrote textual explanations according to the assembly sequence that they chose within the tool. Participants executed the same procedure a second time when they created instructions for the pump task.

During the online event segmentation, participants saw the video in question three times; the first time without any instruction in order to get familiar with it, and the second and third time to segment it into fine and coarse events while watching the videos. The order of fine and coarse segmentation was counterbalanced across participants. While watching the video they tapped a button whenever they thought one meaningful event ended and another meaningful event had begun.

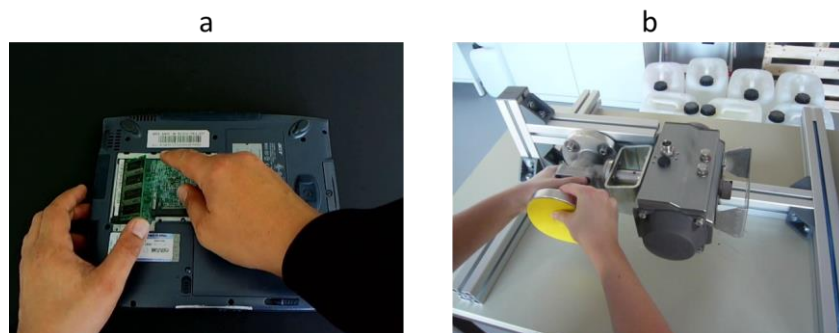To summarize, in the offline event segmentation, the identification of events was without any time constraints, with the explicit aim to create instructions, and without specification of grain. In the online event segmentation, the participants' task was to segment the video according to their subjective perception of fine and coarse event segments, respectively.

4.2.5 Data analysis and statistical methods

In contrast to the statistical analyses with binned data already reported in Mura et al. (2013), the present chapter adds the analyses without binning the data. That is, I take advantage of treating time as a continuous variable. The continuous analysis provides not only a graphical inspection of characteristic peaks assumed to be event boundaries. It offers means to determine significance of a segmentation magnitude (Papenmeier, 2014). I estimated a person's perception or definition of an event boundary as a Kernel density distributed function around the person's key press or end point (Newtson, 1973), respectively. Then, I

summed up all participants' individual distributions for determining the online and offline segmentation plots, respectively. Simulation methods were applied to check for significance of the resulting peaks, on a 95% confidence level.

Furthermore, I made use of the segmentation difference method (Papenmeier, 2014) in order to contrast event boundaries from online and offline event segmentation. I checked for significance of the resulting difference values on a 99% confidence level.

I used R (R Development Core Team, 2008) for all statistical analyses and additional R package segmag (Papenmeier, 2014) for plotting segmentation behavior, determining significant event boundaries, and subtracting groups' segmentation data.

In sum, the following sample of event segmentation measures (Section 3.1.4) applied for this Experiment 1. I analyzed the **number of events** for fine and coarse online event segmentation as well as offline event segmentation. I tested for **significant event boundaries** in all segmentation data. I investigated the **segmentation agreement** between online and offline event segmentation groups by, first, correlating both groups' histograms (binning data) and, second, computing the differences in their segmentation magnitudes (without binning data).

4.3 Results

First, I analyzed the number of events in online and offline event segmentation, respectively. Second and third, I defined the exact locations of event boundaries and compared them between online and offline event segmentation. Fourth, I analyzed the data qualitatively.

4.3.1 Number of events

During online event segmentation, as expected, participants perceived more events boundaries in the fine segmentation conditions compared to the coarse ones. More specifically, participants defined a mean number of 12.0 fine and 4.6 coarse event boundaries in the notebook task and 18.1 fine and 6.1 coarse event boundaries in the pump task, respectively (see Table 1).

In the offline event segmentation, participants segmented only once. The mean number of event boundaries[1] was 6.6 in the notebook task and 11.0 event boundaries in the pump task, respectively (Table 1). That is, for both tasks, the mean number of offline event boundaries lied in between the mean number of event boundaries perceived during the online fine and coarse event segmentation. This implies that the spontaneous structure chosen for the instruction manual creation is a compromise between coarse and fine granularities.

---

[1] I counted the end points identified by each participant (Newtson, 1973).

Table 1: Number of events during both online and offline event segmentation

	M	*Median*	*SD*	*Min*	*Max*
Notebook task					
offline event segmentation ($N_1 = 20$)					
events	6.6	6.5	2.6	2	13
online event segmentation ($N_2 = 12$)					
coarse events	4.6	5.0	1.3	3	6
fine events	12.0	11.5	4.5	5	19
Pump task					
offline event segmentation ($N_1 = 20$)					
events	11.0	11.5	3.2	5	16
online event segmentation ($N_3 = 10$)					
coarse events	6.1	6.0	1.8	4	10
fine events	18.1	18.5	6.7	10	26

However, since the IBES tool also incorporated graphical content, it is possible that the participants initially defined rather coarse offline events in order to subdivide them later into more fine-grained events by choosing more pictures. Therefore, I checked if participants chose offline events that were actually complete. I correlated the number of events chosen within the IBES tool and the number of chosen pictures. If the participants selected fewer events during offline event segmentation and more frames during picture selection in order to represent sub-events within the event, then I would expect a negative correlation. The average number of pictures per event was 2.3 with $SD = 0.9$ (in a range between 1.1 and 5.0 pictures) and no significant correlation (Pearson's r) was found between number of events and pictures chosen per event ($r = -.28$, $p = .08$). Absence of a negative correlation indicates that the information content within an offline event was complete. There are no relevant sub-events present within each event.

4.3.2 Significant event boundaries

Most importantly, I was interested in the exact time points of event boundaries during both online and offline event segmentation. The analyses on event boundary locations are summarized in Figure 16 (notebook task) and Figure 17 (pump task). For each of the two videos, there were three sets of segmentation data: one data set belonged to the offline event segmentation (upper plots in Figure 16 and Figure 17, respectively) and two data sets accompanied the online event segmentation, i.e., fine and coarse event segmentation (two middle plots in Figure 16 and Figure 17, respectively). All three sets of segmentation plots show characteristic peaks representing chronological correspondences in event boundary perception across participants (Figure 16 and Figure 17). I computed the respective critical segmentation magnitude (displayed as a horizontal (red) line, respectively). Values above this cutoff are considered as significant event boundaries (highlighted as vertical (green) lines).

First, I looked at the event boundaries from the offline and the online event segmentation in the notebook task (Figure 16). During the offline event segmentation, participants defined 7 significant event boundaries. During the online event segmentation, not all characteristic peaks exceeded the critical cutoff. There were 6 and 2 significant event boundaries in fine and coarse condition, respectively. Each peak in the coarse condition was related to a peak in the fine condition. This suggests that participants perceived a hierarchically structured stream of information (Kurby & Zacks, 2008).
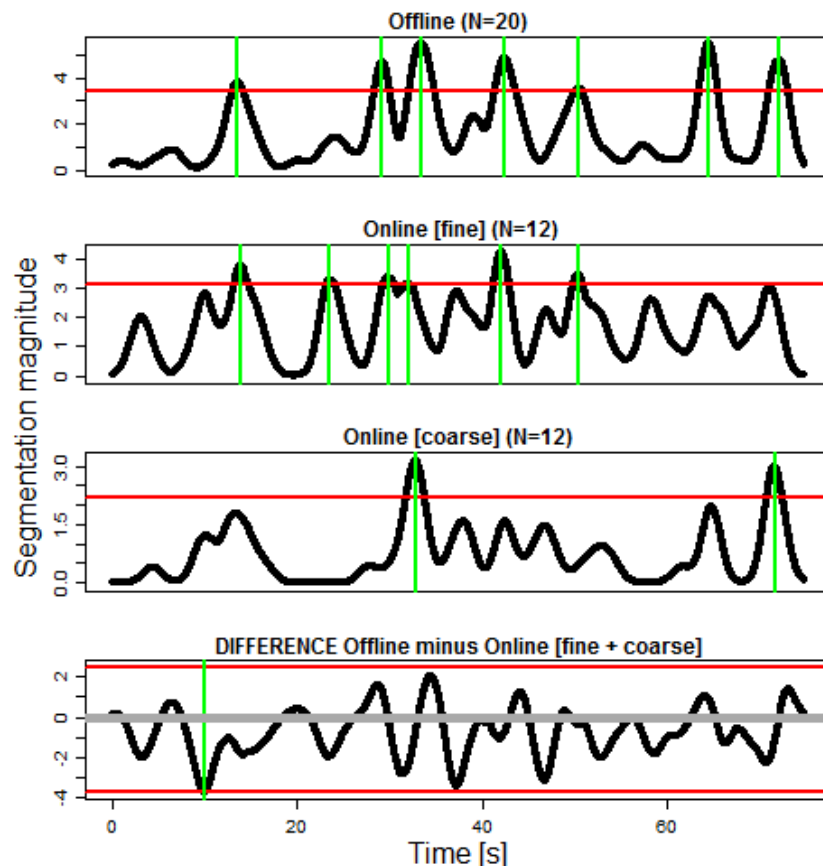
Figure 16: Segmentation plots for the notebook task: The upper plot shows the segmentation behavior during offline event segmentation; the two middle plots depict segmentation behavior during online event segmentation in fine and coarse condition, respectively. The lower plot displays the segmentation magnitude's difference values when subtracting the two middle plots from the upper plot. Significant event boundaries (confidence level of 95%) and differences (confidence level of 99%) are displayed as vertical (green) lines.

When inspecting and comparing the offline and online event segmentation peaks, it becomes apparent that for each peak in the offline event segmentation, there was a corresponding peak in the fine or the coarse online event segmentation. The relation became even clearer when inspecting adversely: Each coarse online event boundary which represents higher-level changes had a corresponding offline event boundary. Yet, most but not all fine event boundaries which represent lower-level changes had a corresponding offline event boundary. In sum, the graphical inspection indicates that there are chronological correspondences between offline and online event segmentation. Furthermore, definition of offline event boundaries seems to incorporate both higher- and lower-level changes.

Similarly, I checked the event boundaries for the pump task (Figure 17). During offline event segmentation, participants defined 12 significant event boundaries. During online event segmentation, I found 8 and 5 significant event boundaries in fine and coarse condition, respectively. As can be seen, there were more than 8 peaks in the fine segmentation; yet, the segmentation magnitude did not exceed the critical cutoff. This is likely due to the small sample size of $N = 10$. Again, each peak in the coarse condition related to a peak in the fine condition (Kurby & Zacks, 2008).

As in the notebook task, inspection of characteristic peaks in the pump task showed a correspondence between the offline and the online event segmentation. That is, each peak in the offline event segmentation matches with a peak in the online event segmentation, either in fine, coarse, or both conditions. In more detail, for each coarse online event boundary, I found a corresponding offline event boundary. Yet, not all of the peaks in the fine condition represented an offline event boundary. Again, offline event boundaries represent both higher- and lower-level changes in activities.
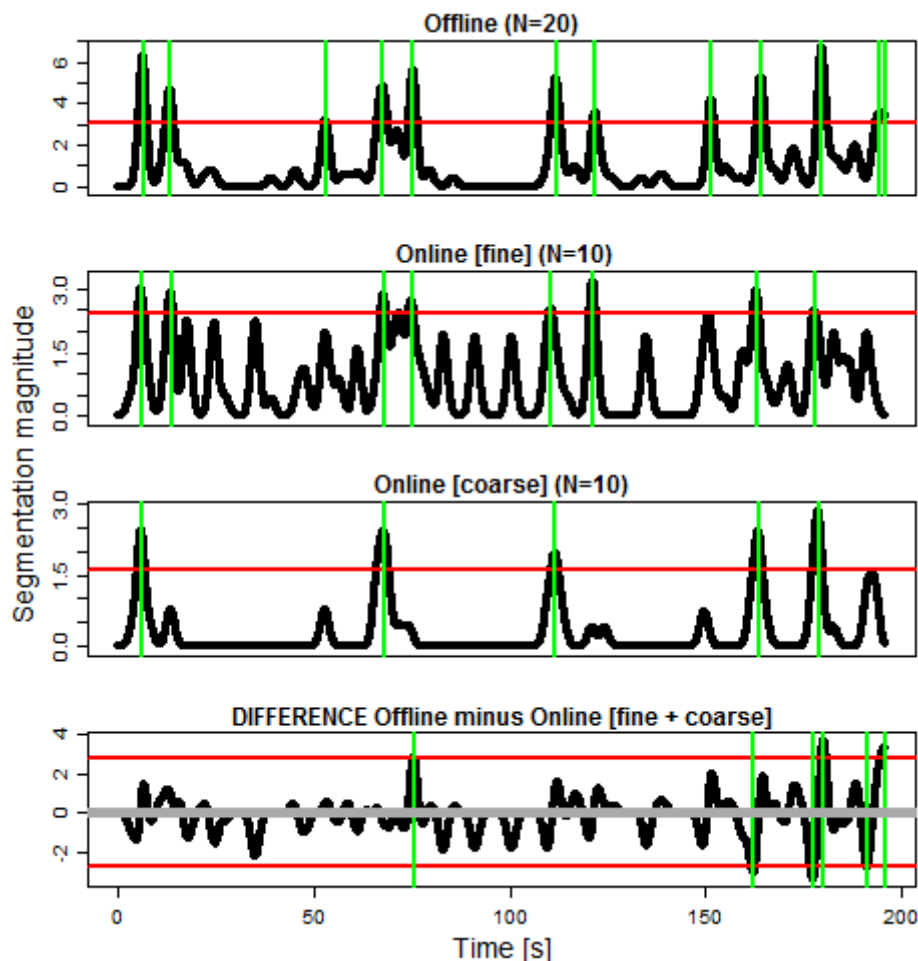
Figure 17: Segmentation plots for the pump task: The upper plot shows the segmentation behavior during offline event segmentation; the two middle plots depict segmentation behavior during online event segmentation in fine and coarse condition, respectively. The lower plot displays the segmentation magnitude's difference values when subtracting the two middle plots from the upper plot. Significant event boundaries (confidence level of 95%) and differences (confidence level of 99%) are displayed as vertical (green) lines.

4.3.3 Segmentation agreement

In order to quantitatively evaluate these observed overlaps between the offline and the online event segmentation, I performed two analyses, i.e., the correlation of group histograms and the difference method.

First, I calculated the overlap between the online and the offline event segmentation groups by correlating their group histograms based on the binned data. For the notebook task, the offline event segmentation behavior substantially correlated with the online event segmentation behavior both in fine ($r = .59$, $p < .01$) and coarse ($r = .66$, $p < .01$) condition. The same was true for the pump task. Significant correlations between offline and online event segmentation were found for the fine ($r = .38$, $p < .01$) and the coarse ($r = .48$, $p < .01$) condition.

Second, I compared the offline and the online event segmentation by subtracting the online event segmentation data from the offline event segmentation data. Note that, in order to compute this difference, I initially collapsed fine and coarse online event segmentation data into one by adding them. The resulting differences in segmentation magnitude between the offline and the online event segmentation are displayed in the lower plots in Figure 16 and Figure 17, respectively. If participants' event boundaries were the same between the offline and the online event segmentation, the difference in segmentation magnitude would correspond to a line around 0 (highlighted by a thick, horizontal (grey) line). If there were event boundaries detected in the offline event segmentation but not in the online event segmentation, the difference value would be positive. Conversely, event boundaries that were detected in the online but not in the offline event segmentation would be displayed as negative values. By simulation techniques, I determined critical cutoffs for negative and positive differences in segmentation magnitude and highlighted these cutoffs as horizontal (red) lines. I chose a confidence probability of 99%.

In the notebook task (Figure 16), collapsing the fine and the coarse event segmentation data sets resulted in a mutual data set based on $N = 24$. The offline event segmentation data set was based on $N = 20$. I found that differences between offline and online event segmentation were mostly 0 (lower plot of Figure 16). The significantly negative difference indicates perception of online but not offline event boundary. Potentially, participants perceived an event boundary due to lower-level changes (supported by the peak in the fine condition in the middle plot of Figure 16). However, this lower-level change was not processed as an important offline event boundary. Overall, this provides support for the hypothesized, quantitative overlap between the offline and the online event boundaries.

In the pump task (Figure 17), the collapsing of the fine and the coarse event segmentation data sets resulted in an overall online data set of $N = 20$ which equaled the group size in the offline event segmentation. Again, I found that subtracting the online from the offline event segmentation data resulted in difference values which did not exceed the respective cutoffs

most of the time (lower plot in Figure 17). This provides additional support for a close relation between offline and online event boundaries.

However, there were infrequent significant differences. In the pump task, the first significant deviation was the positive indicating that the boundary is present in offline but not in online event segmentation (see lower plot of Figure 17). By graphical inspection, we can see, though, that there is a corresponding online fine event boundary for this offline event boundary. Because many participants defined this event boundary during offline event segmentation, the sum of the fine and the coarse segmentation magnitudes was insufficient to make the difference zero despite of an existing overlap between the online and the offline event segmentation. The same is true for the next deviation. This time, the difference is significantly negative suggesting that participants defined it during the online but not during the offline event segmentation. However, through graphical inspection, we can see that it was defined in offline and in *both* fine and coarse online conditions. The fact that participants agreed that strong during fine and coarse event perception led to a high overall magnitude when adding fine and coarse segmentation data. This is responsible for the negative difference.

Towards the end of the video, I found a significant negative difference followed by a significant positive difference indicating that an offline event boundary was defined a little later than the online event boundary. This could be due to the fact that assembly operations were faster around this time point. In offline event segmentation, the processing was less flustered for the participants compared to the participants in the online event segmentation condition who were more under pressure to press the key due to the fact that the video will end soon. Finally, the last deviation indicates that an event boundary was defined during offline but not during online event segmentation. At this time point, the video abruptly stopped so that participants who perceived the last event did not have enough time to press the key before the video stopped.

In sum, the correlational analyses as well as the methods based on continuous treatment of time provided evidence for temporal similarity between online and offline event boundaries.

4.3.4 Qualitative analysis

In a final step, two independent raters analyzed the identified offline events across all participants including the textual descriptions. The aim was to confirm the events found in Figure 16 and Figure 17. Furthermore, it provided a more detailed insight into the offline event segmentation in the context of the instruction creation process of participants.

First, the two raters identified a consensus version of an instruction manual for each task. An offline event became the part of a consensus version of the task in case both raters agreed that it was defined by at least half of all participants (N ≥ 10). This resulted in 7 events for the notebook and 11 events for the pump task, respectively (see Table 2). These numbers go along with the numbers of significant offline event boundaries, i.e., 7 in the notebook and 12

in the pump task (Section 4.3.2). Furthermore, the numbers of offline events again indicate that the structure of the manual is a combination of fine and coarse events (Section 4.3.1).

Table 2: Consensus version for the notebook and pump task, respectively: Events that more than half of all participants (N >= 10) identified and described textually within the IBES tool.

Notebook task	Pump task
1. Turn the notebook upside down	1. Put ball valve into base
2. Unscrew both screws of the cover	2. Put casing onto base
3. Remove the cover	3. Fix with four screws
4. Insert the RAM	4. Tighten the screws with spanner
5. Put the cover on again	5. Put positioner covering on positioner
6. Screw both screws of the cover	6. Screw four screws
7. Turn the notebook back	7. Put the positioner onto the actuator
	8. Fix it with 2 nuts
	9. Tighten the nuts with spanner
	10. Connect actuator and positioner by pipe
	11. Connect the tube with the positioner

Second, as already shown in Section 4.3.1, the numbers of events identified across participants in the offline event segmentation differed across participants. They ranged from 2 to 13 in the notebook and 5 to 16 in the pump task (Table 1). This suggests that participants varied both towards more detailed and broader segmentations during instruction creation compared to the consensus versions listed in Table 2. Therefore, both raters analyzed all deviations from consensus and agreed on 7 deviations for the notebook and 9 deviations for the pump task. In the following two paragraphs, I will describe a few examples of deviations in either direction compared to the consensus shown in Table 2. They illustrate that the individuals defined different granularities.

On the one hand, there were participants who created more coarse assembly steps. In the manuals for the notebook task, five participants summarized "Putting the cover on again" (see step 5 of the notebook task in Table 2) and "Screw both screws of the cover" (step 6) into one single step of "Closing the cover". Similarly, steps 2 and 3 were summarized into "Open the cover". A reduced notebook manual incorporating these consolidations would consist of five assembly steps. Consolidations for the pump task would result in a coarser pump manual of six steps according to the inspections of the raters. If participants segmented in a coarser way compared to the consensus version of the pump task in Table 2, then they might merge steps 2 to 4, 5 and 6, and 7 to 9 into one step, respectively. For instance, they did not segment "Put the positioner onto the actuator", "Fix it with two nuts", and "Tighten the nuts with spanner" (steps 7 to 9 from Table 2) into separate steps but perceived all three of them as one common step "Assemble the positioner onto the actuator". The number of instructional steps in the coarser instruction manuals equals the mean number of coarse event boundaries (5 in the notebook and 6 in the pump task, respectively).

On the other hand, some manuals created by participants had a more detailed structure than indicated in Table 2. A number of participants added steps like "Initial state" and "Final state"

to their manual. However, even the most detailed instruction manuals included 13 and 16 steps but did not reach the levels of fine granularity of the fine event segmentations (see 19 and 26 fine events in Table 1). For example, no subject understood laying down a tool as a separate step whereas during fine event segmentation some participants did.

Taken together, the initial graphical comparison showed an overlap between offline and online event boundaries. Second, the difference values between offline and online data confirmed their chronological correspondence. Finally, qualitative results confirmed and extended comprehension of the offline event segmentation within the instruction creation process.

4.4   Discussion

In this experiment, I investigated the question if online event perception results in the same event boundaries as offline elaboration. My aim was to show a quantitative overlap. Indeed, the event boundaries during the offline event segmentation were at similar temporal locations as the event boundaries during the online event segmentation. This demonstrates that boundaries in offline elaboration are boundaries in event perception and, vice versa. Thus, this finding is in line with the claim that organizing dynamic activity into events is at the bottom of different processes from cognition (Radvansky & Zacks, 2014).

Despite the overlap of respective event boundaries, I found that not all fine event boundaries constituted important strategic points during offline elaboration. Fine event boundaries depict lower-level movement changes and were not always incorporated as strategic points in the offline event segmentation, especially, if they illustrated repetitive actions, e.g., screwing screw 1, screwing screw 2, and so on. These fine event boundaries were summarized in the offline event segmentation into one. Instead, coarse event boundaries had corresponding offline event boundaries throughout. It means that similar higher-level conceptual changes guided both online and offline event segmentation. This finding is in line with prior work claiming that, regardless of whether perceiving online or elaborating more deeply, persons make usage of the same scripts (Radvansky & Zacks, 2014; Zacks et al., 2001). Especially, they are guided by the same situation models when it comes to higher-level changes. Nevertheless, they also use lower-level changes to define offline event boundaries. To sum up, higher-level, object-based as well as lower-level, action-based information is used for offline event segmentation.

As reported for the online event segmentation (Zacks et al., 2001), I found substantial interindividual differences in the offline event segmentation, too. My qualitative analysis confirmed that individuals differed in number of event boundaries they defined during instruction creation. Again, this finding shows that higher- and lower-level information can be used for definition of offline event boundaries. Furthermore, this result shows that the IBES tool represents a tool that can be further used for investigating the individual differences in the offline event segmentation process.

For instance, it may be useful for further examining different instruction creation processes of participants with differing expertise levels, i.e., novices versus experts, in order to understand their situation models potentially differing in granularity and content. It may also support further research on theory of mind. For instance, Killingsworth, Saylor, and Levin (2005) were interested in finding if their participants would create different instructions when asked to create instructions either for humans or for computers. Hence, they showed that participants defined more segments for computers likely because they attributed limited reasoning capabilities to them. An important contribution of this newly created IBES tool for suchlike research is that it supports log files with which investigators can further analyze the offline segmentation or the instructional design process.

Former research in instructional design already elaborated that the structure of instructions should be based on event boundaries as important strategic points (e.g., Zacks & Tversky, 2003). However, it was an empirical question if perceivers of an activity actually share the same event boundaries as creators of instruction manuals. My experiment showed that the perceivers of a task and the creators of instructions for this task rely on the same strategic points. Thus, I provided an additional justification for connecting event cognition research with its application in instructional design.

Another asset is that the IBES tool is the first software tool that makes it possible to create instructions based on event segmentation semi-automatically compared to current approaches. Thus, manuals based on event boundaries which are important for understanding and memory can be created easily. This is advantageous both for research and practical application. In research, easy creation of manuals could promote evaluations of different types of instructions, for instance, with varying structure (fine- versus coarse-grained) or differing contents (graphical versus textual). For a more detailed discussion on the further usage of the IBES tool refer to Section 7.1.

The difference method resulted in the same finding as the correlation method, i.e., there was a meaningful quantitative overlap between online and offline event segmentation. However, the segmentation magnitude is dependent on overall key presses which, in turn, are influenced by the sample size. As differences are calculated based on segmentation magnitudes, group sizes should be approximately equal. Furthermore, even if both groups show a significant event boundary at a given time point, the method is still sensitive to differences. Therefore, a high confidence level of 99% should be preferred in order to avoid such significant differences based on mathematical interferences. Nevertheless, the more sample sizes deviate from each other, the less valid the method gets. Having said that, the difference method provides an additional analysis compared to graphical inspection alone and it is comprehensive in order to compare segmentation behavior between groups.

In sum, I concluded from this experiment that the easy to perform classical online event segmentation task captures the event structure of assembly tasks which, in working context, are processed mostly offline. Consequently, the event segmentation task can be used for my further studies because it provides ecologically valid event boundaries.

5 Experiment 2: Practicing assembly tasks

As motivated in Chapter 4, assembly workers process their tasks offline rather than during online perception. In addition, they process a specific task not only once but practice it repeatedly in order to prepare for their work at the production line. Therefore, the goal of Experiment 2 was to test the effects of repeated presentation on processes from event cognition.

I used the classical online event segmentation task in the following experiment to determine fine and coarse event boundaries. Based on them, I investigated if workers acquire fine versus coarse events differently in the context of training.

5.1   Introduction

It was shown that working conditions in manufacturing (Section 2.1.1) are characterized by a regular change of assembly sequences due to new car models. Introducing new assembly sequences requires workers to regularly practice them on specially built hardware prototypes (Hermawati et al., 2015). Additionally, recently, workers have been able to practice new assembly tasks in a virtual environment (Gorecky et al., 2013; Malmsköld et al., 2007b). According to the Deliberate Practice Framework (Ericsson et al., 1993) all such opportunities for repeatedly practicing the same or similar assembly sequences are the key to learning and expertise development. According to the Event Segmentation Theory (Kurby & Zacks, 2008; Zacks et al., 2007), however, sequences consist of two conceptually different strategic points, i.e., coarse event boundaries depicting higher-level conceptual changes and fine event boundaries depicting lower-level, less salient changes representing "ongoing activity" (Radvansky & Zacks, 2014; Swallow et al., 2009). Consequently, working memory models depict content differing in hierarchical level.

Acquisition of fine and coarsely segmented information after repeated presentation has not been in the focus of research until now. Rather, studies exploring memory for dynamic events presented the stimulus material only once (Lassiter & Slaw, 1991; Newtson & Engquist, 1976; Swallow et al., 2009; Zacks, Swallow, Vettel, & McAvoy, 2006). A number of studies indicated that memory for coarse events is better than for fine events (Bransford et al., 1972; Gernsbacher, 1985; Johnson-Laird & Stevenson, 1970; Treisman & Tuxworth, 1974). Yet, as repeated presentation of stimulus material is crucial for learning (Ericsson et al., 1993) and repeated presentation changes basic memory processes (e.g., Reagh & Yassa, 2014), it is an important but unanswered question whether repetition affects development of memory of coarse events more than fine events or vice-versa or whether there is no difference between the two processes.

In this experiment, I will investigate how repeated practice affects memory for fine and coarse events. I will test the Deliberate Practice Framework hypothesis that repeated practice enhances the acquisition of assembly sequences. Adding evidence from event cognition, I predict that learning curves will differ for different events such that coarse, more salient events will be more successfully learned than fine, less salient events. This is because fine events are easier to confuse and more similar to each other compared to distinct coarse events. In addition, familiarity influences event cognition (Graziano et al., 1988; Jarodzka et al., 2010; Zacks et al., 2001), therefore memory processes after training may differ between domain experts and novices. Therefore, I included students from middle school in the age just before potential automotive job entry and production workers with a high degree of work experience.

This experiment consisted of two tasks, i.e., the classical event segmentation task and the virtual training task. The purpose of the initial event segmentation task was to confirm that the assembly task used as stimulus is hierarchically perceived in coarse and fine events. In the virtual training task, another sample of participants consisting of experts and novices practiced the assembly task three times in a virtual environment. After each repetition, I tested their memory. I assessed memory for coarse events by stopping viewing of the task's video either shortly *before* a coarse event boundary and asking for the correct next event; or, stopping shortly *after* a coarse event boundary and asking for the correct next fine event.

5.2 Methods

5.2.1 Car door assembly material

The car door assembly used in this experiment consisted of mounting different parts of a car door to the rack of the same door in a given sequence (see Figure 18 and Figure 7). The assembly contained typical manual operations from the production line of the Adam Opel AG, a German automotive company (e.g., picking up a work piece, screwing, etc.). Concretely, it consisted of 38 single operations.

A video of this assembly was shot from a point-of-view perspective using a head-mounted camera. On the one hand, this video was used for the initial event segmentation task (see Section 5.2.2). On the other hand, this video served as the basis for the development of a memory test (see Section 5.2.4).

Figure 18: Real door setup: The two upper pictures depict the car door rack from front and back side, respectively; the lower picture shows the main objects, screws, and tools. Source: Adam OPEL AG.

5.2.2 Event segmentation task

Participants watched the video of the car door assembly and pressed the space bar key whenever they thought that one meaningful event ended and another one began. I used this event segmentation task (Newtson, 1973) in order to empirically determine the structure of the task with respect to its coarse and fine event boundaries. I presented the video which was 7 minutes and 16 seconds long without sound. Overall, participants saw the video three times. First, they watched it without instruction, and then, they had to segment it both in fine- and coarse-grained events. The order of fine and coarse segmentation was counterbalanced across participants.

5.2.3 Virtual training task

Participants executed the virtual training task by the virtual training setup that was introduced in Section 2.2.2 and is shown in Figure 19. Participants saw the 3D simulation of the car door assembly on a monitor approximately 2 meters in front of them. Their task was to move an object shown on the screen to the correct assembly position using their hand motion tracked through a Microsoft Kinect. The correct assembly position was highlighted by a semi-transparent blue area shaped like the object in question (see Figure 4 in Section 2.2.2), e.g., a door part, a screw, or a tool. Red, orange, and green colors were given as visual feedback, respectively, in order to indicate how close the object was located with respect to its target position. When participants positioned the object correctly, they confirmed this assembly step by pressing the button on their Wii Mote controller. Then, they saw the next object. I used this so-called "easy mode" for training the participants in executing the door assembly task.

The virtual training system incorporated a more difficult mode ("advanced mode") as well. I used this mode as an additional final performance measure (see Section 5.2.6 and Section 5.3.3). In this mode, participants were asked to choose the correct subsequent object on their own using a circular menu (see Figure 4 in Section 2.2.2). As a hint for selecting the correct part, they could see the blue highlighted area that indicated the shape and target

position of the subsequent object. If a wrong part was selected, an error message appeared followed by the circular menu with the correct object in the foreground.

The door assembly within the virtual training system involved 38 assembly steps that represented an imitation of the real assembly sequence introduced in Section 5.2.1 and Figure 18.
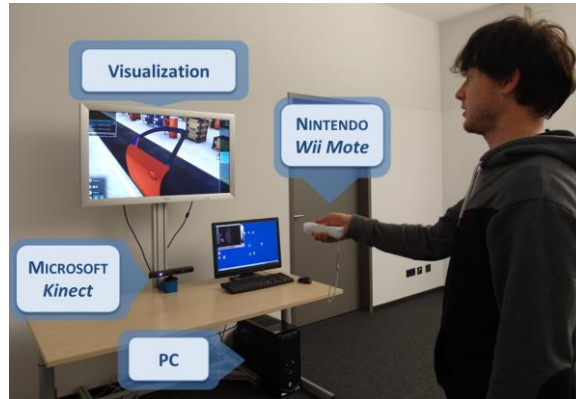
Figure 19: Virtual training system setup: Flat screen for visualization (52 inches), PC on which software was running, Microsoft Kinect for motion tracking, Nintendo Wii Mote as controller. Source: DFKI GmbH.

5.2.4 Memory test based on coarse event boundaries

In order to test memory performance for the correct assembly sequence, I created a test (similar to Swallow et al. (2009) or Zacks, Kurby, Eisenberg, and Haroutunian (2011)) based on the video of the real door assembly (Section 5.2.1) and based on the results of the event segmentation task (Section 5.2.2) which will be described in detail below (Section 5.3.1). The video stopped at time points associated with the coarse event boundaries of the door assembly and the test asked for predicting the correct next event frame. In the "predicting coarse" condition, the video stopped before a coarse event boundary; this tested memory for coarse events. In the "predicting fine" condition, the video stopped after a coarse event boundary; this tested memory for fine events. An illustration of the test is given in Figure 20.

So, depending on stop position of the video clip, i.e., before or after the coarse event boundary, there were two different conditions in the memory test. The video clips in the "predicting coarse" condition began two fine steps from the respective coarse event boundary and stopped shortly before it, i.e., shortly before the person was just about to turn back to the table in order to take the next part. The videos in the "predicting fine" condition began when the person in the video turned towards the table and stopped when she had gripped the main object from the table, so, shortly after the new coarse event began (see the overview in Figure 20). The memory test contained 14 video clips (seven "predicting coarse" and seven "predicting fine" items).

Immediately after the video clip stopped, participants saw a static picture frame depicting either the correct (target) or wrong next step (distractor) taken from the video. Target pictures depicted a screenshot of the next step. Distractor pictures in the "predicting fine" condition

depicted the assembly operation two fine steps ahead. Distractor pictures in the "predicting coarse" condition depicted one coarse step ahead.
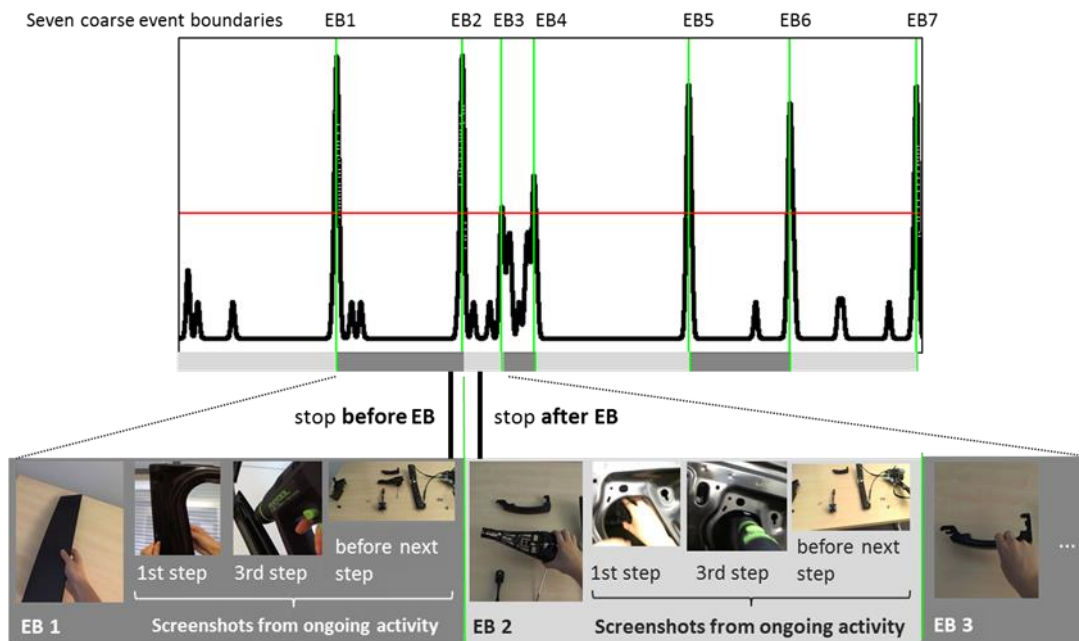
Figure 20: Memory test illustration: I schematically sketched the memory test using three consecutive coarse event boundaries from the assembly task, i.e., EB1, EB2, and EB3. Video clips were stopped either before or after a coarse event boundary (EB). The video clip that stopped before exemplary coarse event boundary, EB 2, was a coarse event item. Its respective target picture was "EB 2" and distractor picture was "EB 3". The video clip that stopped after exemplary coarse event boundary, EB 2, was a fine event item. Its target picture was "1st step after EB 2" and distractor picture was "3rd step after EB 2".

I presented each video clip twice, one time testing memory with a target and the other time with a distractor item. Order of presentation was chosen at random. Participants indicated via key press whether the shown picture was the correct next step ("old" response) or not ("new" response), respectively. The test was created using PsychoPy software (Peirce, 2007) and participants executed it on a conventional notebook PC taking approximately 15 minutes. I calculated the non-parametric Signal Detection Theory measures (Stanislaw & Todorov, 1999) sensitivity (A') and response bias (B''), see Section 5.2.7.

5.2.5 Participants

Students in the event segmentation task

For the event segmentation task, I used a sample of $N = 10$ students (5 male; age: $M = 24.6$, $SD = 4.6$) from the University of Kaiserslautern and University of Tübingen.

Experts and novices in the virtual training task

For the virtual training task, I used a sample of overall $N = 37$. Novice participants were middle school students from the Neues Gymnasium in Rüsselsheim, Germany ($N = 19$; $M_{age} = 14.9$ years, $SD_{age} = .3$) and the experts were production workers from the Volvo Trucks plant in Gothenburg, Sweden ($N = 18$; $M_{age} = 42.2$ years, $SD_{age} = 7.8$). Workers from Volvo

differed from students in having been working in production for an average of 17.2 years (*SD* = 7.3).

Experts had significantly higher self-reported manual skills compared to novices ($t(32) = -3.73$, $p < .01$), but they did not outperform with respect to their spatial ability tested by a mental rotation task ($t(29.9) = 1.19$, $p = .24$) (see Table 4). Both groups had no prior knowledge on the car door task (note that Volvo workers usually assemble trucks not cars).

5.2.6 Design and procedure

Event segmentation task

For the event segmentation task, I used the classical procedure (Newtson, 1973) for all *N* = 10 participants. The order of fine and coarse event segmentation was counterbalanced across participants. The event segmentation task took place at the University of Kaiserslautern and University of Tübingen.

Virtual training task

For the virtual training task, I adopted a within-subject design in which both experts and novices (*N* = 37) executed three training repetitions each followed by the memory test based on coarse event boundaries introduced in Section 5.2.4. Volvo Gothenburg and Neues Gymnasium Rüsselsheim compensated participants' absence from school or work, respectively.

The data assessment of experts was conducted in Volvo Trucks Factory in Gothenburg, Sweden. Novices participated one month later in the Neues Gymnasium in Rüsselsheim, Germany. Experimenters were previously trained at the DFKI. Production workers at Volvo signed an informed consent right before the experiment started. Student participants brought a consent form signed by their parents.

Table 3: Procedure of the virtual training task in a within-subject design

1. Tutorial
2. Virtual assembly training and testing (repeated 3 times)
   a. Easy mode
   b. Memory test
3. Virtual assembly training: advanced mode
4. Expertise assessment
   a. Self-reported manual skills
   b. Mental rotation test

The procedure as summarized in Table 3 was the same for all *N* = 37 participants. Each participant filled in a demographic questionnaire followed by the experimenter's oral introduction about the overall project. Each participant was then calibrated within the virtual training system. In order to get familiar with its usage, all participants performed a tutorial

consisting of seven practice assembly steps at the front spoiler of a car. After this practice trial, they performed the virtual training for the door assembly task three times in the easy mode while the experimenter recorded the time required to complete the task with help of a stopwatch. After each training repetition, participants were asked to execute the memory test (see Section 5.2.4). Next, they performed a final virtual assembly training session in the advanced mode while the experimenter noted errors of choosing the next object.

The expertise assessment consisted of a questionnaire on manual skills including six items on a 5-point Likert-scale (e.g., "I find it hard to assemble furniture by myself"). Average scores might have ranged between 1 and 5 indicating highest and lowest mean self-reported manual skills, respectively (see Table 4). Furthermore, I applied a 5-minute computer mental rotation test designed by PsychoPy software (Peirce, 2007) in order to assess potential differences between novices and experts. On a conventional notebook monitor, the experimenter showed a letter ("R" or "G") either in mirrored or normal view. Additionally, the letter could be rotated. Participants had to indicate by button press as fast as possible if the letter was mirrored or not. I calculated the sensitivity A' based on hits and false alarms for the mental rotation test (see Table 4).

Table 4: Differences in age and expertise measures between novices and experts

	Experts	Novices		
	M (SD)	*M (SD)*	*t*	*p*
Age [years]	42.2 (7.8)	14.9 (.3)	14.78	<.01
Manual skills[a]	1.5 (.45)	2.2 (.65)	-3.73	<.01
Spatial ability[b]	.98 (.04)	.95 (.07)	1.19	.24

Note. [a]Average score with 1 and 5 indicating highest and lowest self-reported manual skills, respectively; [b]Sensitivity A' based on hits and false alarms in a mental rotation test.

5.2.7 Data analysis and statistical methods

Event segmentation task

For analyzing the data in the event segmentation task, I treated time as a continuous variable. I estimated a person's perception of an event boundary as a Kernel density distributed function around the person's key press. Then, I summed up all participants' individual distributions. Simulation methods were applied to check for significance of the resulting peaks, on a 90% confidence level. Furthermore, I analyzed perceived hierarchical structure of the assembly according to hierarchical alignment and enclosure (Section 3.1.4).

I used R (R Development Core Team, 2008) for all statistical analyses and additional R package segmag (Papenmeier, 2014) for analyzing event segmentation data.

In order to understand the event structure of the car door assembly task with respect to fine and coarse event boundaries, the following event segmentation analyses (Section 3.1.4) were used. I tested for **significant event boundaries** in fine and coarse event segmentation

data, respectively. I calculated **hierarchical alignment** and **hierarchical enclosure** in order to confirm the hierarchical structure of the assembly.

Virtual training task

In order to analyze memory performance after repeated practice, I applied Signal Detection Theory measures to the memory test (Section 5.2.4). The theory is based on the calculation of the hit rate (i.e. the proportion of "old" responses to target items) and the false alarms (i.e. the proportion of "old" responses to distractor items). Both of them reflect two factors: the sensitivity, i.e. the actual cognitive ability to detect a picture as the target or distractor, and a response bias, i.e. the general tendency to respond "old" or "new" in an old/new recognition test.

I calculated the non-parametric values for sensitivity, i.e. A', ranging from .5 (no ability to distinguish between target and distractor) to 1 (perfect performance), and response bias, i.e. B'', ranging from -1 (saying always yes) to 1 (saying always no) with 0 representing no response bias. Sensitivity values less than .5 may arise from sampling error or response confusion with the minimum value being 0. Furthermore, I analyzed response times of memory test answers. Finally, I checked the performance in the virtual training task itself.

I assessed the need for linear mixed effects analysis by fitting two models, i.e., one with constant intercept for all participants and another allowing intercepts to vary across participants (Field, Miles, & Field, 2012). If the comparison of fit indices revealed significant existence of random effects, I performed a linear mixed effects analysis by the help of R package lme4 (Bates, Maechler, Bolker, & Walker, 2015). In case of absence of random effects, I computed ANOVAs.

5.3   Results

5.3.1 Fine and coarse event boundaries

Significant event boundaries

The respective event segmentation plots for both fine and coarse condition are displayed in Figure 21. I found 7 meaningful event boundaries in the coarse condition indicated by vertical (green) lines in the upper plot of Figure 21. They correspond to 7 main objects that have to be assembled successively onto the car door rack. In between those coarse event boundaries, participants perceived several fine steps, respectively, i.e., positioning the current object, inserting screws, and fixing the screws with the help of a tool. These additional fine event boundaries are indicated by the vertical (green) lines in the lower plot of Figure 21.
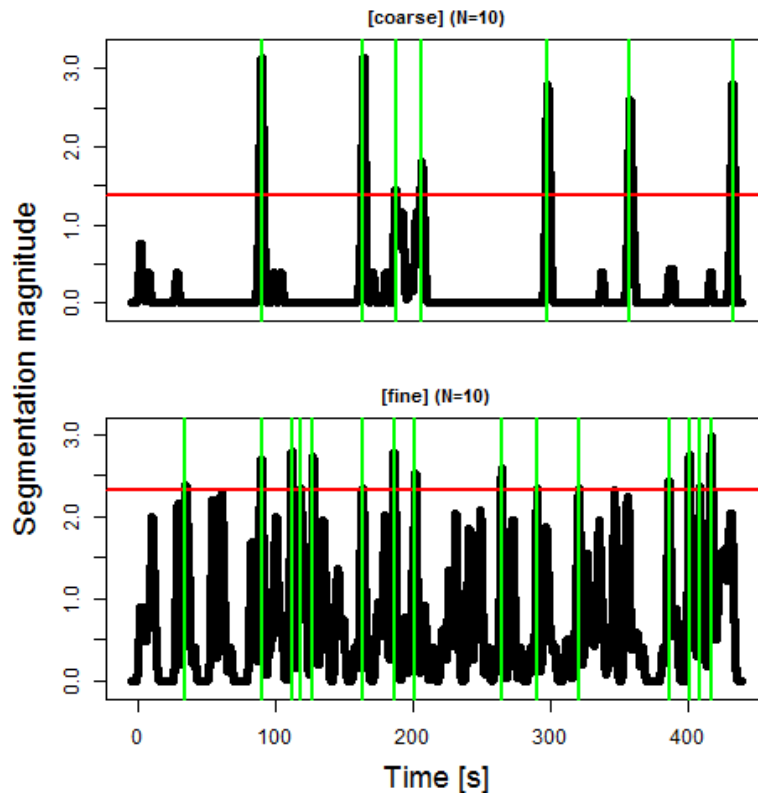
Figure 21: Event segmentation plots for the car door assembly task: Upper plot shows event segmentation behavior in coarse condition, bottom plot displays fine condition. Significant event boundaries (confidence level of 90%) are displayed as vertical (green) lines.

Hierarchical alignment and enclosure

Furthermore, participants perceived the activity according to hierarchical alignment and enclosure. First, participants observed more temporal closeness between fine and coarse event boundaries ($M = 2.7$ s, $SD = 1.1$) than expected by chance ($M = 4.8$ s, $SD = 1.2$), $t(9) = -5.24$, $p < .01$. Second, the hierarchical enclosure value was .81 ($SD = .24$) and significantly higher than a proportion of .50 assumed under the null hypothesis ($t(9) = 4.12$, $p < .01$). The significant deviation means that more than half of the nearest fine event boundaries, more specifically, 81%, *preceded* its respective coarse event boundary. These results suggest that perception of the car door assembly was hierarchically structured, i.e., several fine event boundaries were chunked under its respective coarse event boundary.

5.3.2 Memory performance after repeated practice

After I confirmed the hierarchical structure of the assembly task in question, the main aim of this experiment was to investigate whether there is any difference in acquisition of events after repeated practice based on their hierarchical level.

Sensitivity

In order to investigate memory performance, I analyzed the influence of repetition (1, 2, 3), expertise (experts, novices), and item type (predicting coarse, predicting fine) on memory

performance. The data is plotted in Figure 22. Because my test for random effects revealed individual participant as random factor, I calculated a linear mixed effects model that is summarized in Table 5 along with post hoc analyses. I found a significant interaction effect between repetition and item type ($F(1, 35) = 15.96$, $p < .01$). This suggests an improvement in memory with increasing training repetition, but only for coarse events, not for fine events.
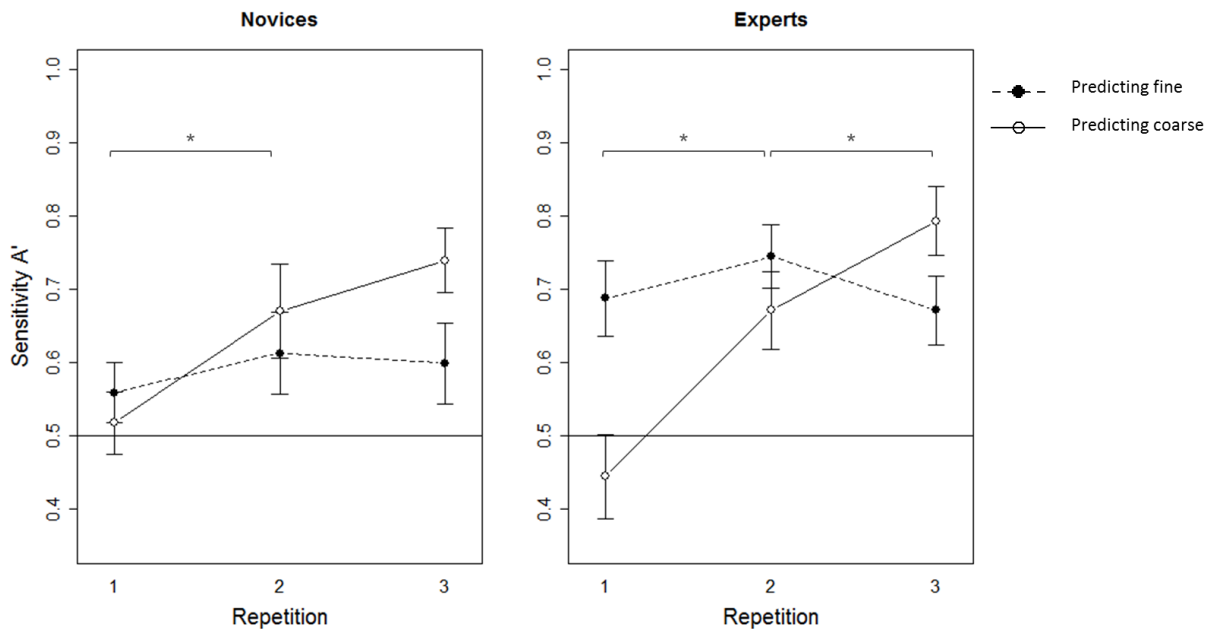
Figure 22: Learning curves after virtual training with respect to sensitivity A': The significant trends for the "predicting coarse" condition are indicated by * (95% confidence level). Error bars reflect standard errors.

Further, there was a significant interaction effect between expertise and item type ($F(1, 35) = 4.47$, $p < .05$). Experts performed generally better in predicting fine events ($M = .79$, $SD = .20$) compared to novices ($M = .59$, $SD = .22$). In contrast, memory for coarse events did not differ between experts and novices ($M = .64$, $SD = .26$ versus $M = .64$, $SD = .24$).

Thus, experts showed initial high performance for memory of fine events indicating existence of prior knowledge cued by a specific automotive-related object.

Table 5: Results of the mixed-effects model for sensitivity A' in the memory test

	b	SE b	95% CI	t	p
baseline A'	.75	.08	.57, .86	9.50	<.01
repetition	-.00	.03	-.07, .06	-.24	.81
expertise	-.17	.03	-.38,.04	-1.59	.12
item type	-.43	0.10	-0.63, -0.23	-4.16	**<.01**
item type * repetition	.18	0.05	0.09, 0.28	3.81	**<.01**
item type * expertise	.30	0.15	0.02, 0.58	2.08	**<.05**
expertise * repetition	.03	.05	-.06, .12	.59	.55
item type * repetition * expertise	.03	.05	-.06, .12	.59	.55

Response time

By using a linear mixed effects model with individual participant as random effect (see Table 6), I analyzed if response time to the items in the memory test depended on item type, expertise, repetition, and sensitivity A'. I found that time (in seconds) decreased with repetition (M_1 = 5.77 (SD = 3.00), M_2 = 4.58 (SD = 1.94), and M_3 = 3.99 (SD = 1.96); $F(2, 68)$ = 25.77, p < .01) and that the coarse events required longer response times than the fine (M_{coarse} = 4.99, SD = 2.84 and M_{fine} = 4.57, SD = 1.99; $F(1, 108)$ = 4.34, p < .05). However, this is true only for experts who differed in response time for coarse versus fine events (M_{coarse} = 6.30, SD = 3.31 and M_{fine} = 5.27, SD = 2.09); novices did not show this difference (M_{coarse} = 3.70, SD = 1.36 and M_{fine} = 3.86, SD = 1.62), $F(1, 102)$ = 9.02, p < .01 (see Figure 23).

Furthermore, response time for predicting coarse events decreased more clearly with repetition than for predicting fine events, $F(2, 102)$ = 6.78, p < .01. Slowed responses for the coarse events which accelerate with repetition indicate a higher initial difficulty of the coarse compared to the fine events which diminishes after repeated training.
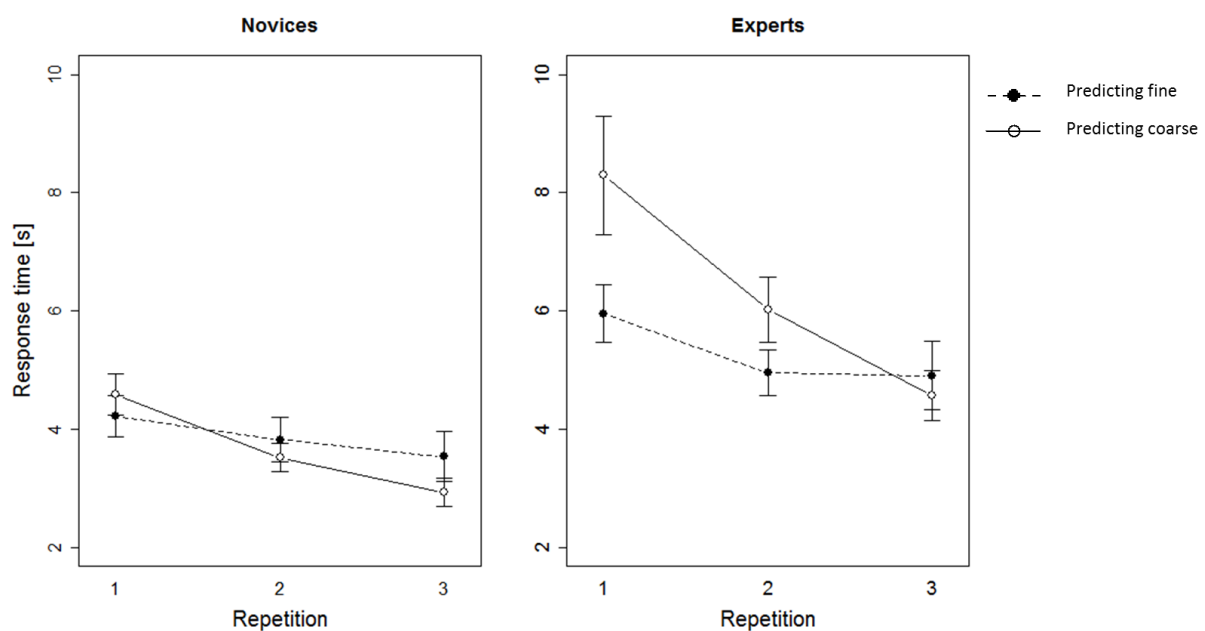
Figure 23: Response times to the memory test. Error bars reflect standard errors.

In order to be able to contrast response time with sensitivity A', I plotted response time curves in Figure 23 analogous to Figure 22. From the illustration, it becomes apparent that, overall, experts took longer regardless of repetition and item type. They show clearer negative slope in predicting coarse condition compared to novices. They show a zero slope for predicting fine events.

Table 6: Results of the mixed-effects model for response time in the memory test

	b	*SE b*	*95% CI*	*t*	*p*
baseline	6.26	1.04	4.27, 8.25	6.04	<.01
repetition	-.52	.25	-1.00, -.05	-2.10	**<.05**
expertise	-.43	1.36	-3.11, 2.24	-.32	.75
item type	3.56	1.15	1.36, 5.76	3.10	**<.01**
sensitivity A'	.08	1.15	-2.13, 2.28	.07	.94
item type * repetition	-1.46	.40	-2.22, -0.70	-3.69	**<.01**
item type * expertise	-4.02	1.57	-7.04, -1.01	-2.56	**<.05**
expertise * repetition	.22	.35	-.46, .89	.62	.54
item type * repetition * expertise	.95	.54	-.09, 1.99	1.75	.08

Response bias

Because of absence of random effects for response bias as dependent variable I computed an ANOVA (see Table 7) which revealed significant main effects of item type ($F(1,35)$ = 68.03, $p < .01$) and expertise ($F(1,35)$ = 4.16, $p < .05$) and a significant interaction between item type and expertise ($F(1,35)$ = 6.98, $p < .01$). The fine event items were more often rated as "old" ($M = -.55$, $SD = .52$) compared to the coarse event items ($M = -.02$, $SD = .57$). Experts had a general higher tendency to rate items as "old" ($M = -.38$, $SD = .59$) than novices ($M = -.20$, $SD = .60$). Novices compared to experts regarded coarse event items significantly more often as "new" ($M = .15$, $SD = .60$ versus $M = -.20$, $SD = .59$).

However, the analysis of the response bias was of minor interest. Response bias declined in conditions with high sensitivity values and was highly negative in fine condition. This indicates that participants tended toward distractor-responses in conditions in which they were less accurate in distinguishing between targets and distractors.

Table 7: ANOVA results for response bias B'' in the memory test

	df	*F*	*p*
repetition	2	.01	.91
expertise	1	4.16	**<.05**
item type	1	68.03	**<.01**
item type * repetition	2	2.45	.12
item type * expertise	1	6.98	**<.01**
repetition * expertise	2	.16	.70
item type * repetition * expertise	2	.04	.84

5.3.3 Performance in the virtual training task

I also looked at the virtual training execution time (Figure 24). I performed an ANOVA with number of repetitions as independent and execution time as dependent variable. Time for a

single virtual training session significantly decreased with repetition in novices ($F(1, 18)$ = 15.39, $p < .01$. Specifically, there is a significant increase in speed from the second to the third training ($t(18) = 3.25$, $p < .01$)), however, the experts did not get faster with more repetition ($F(1, 17) = .21$, $p = .65$).
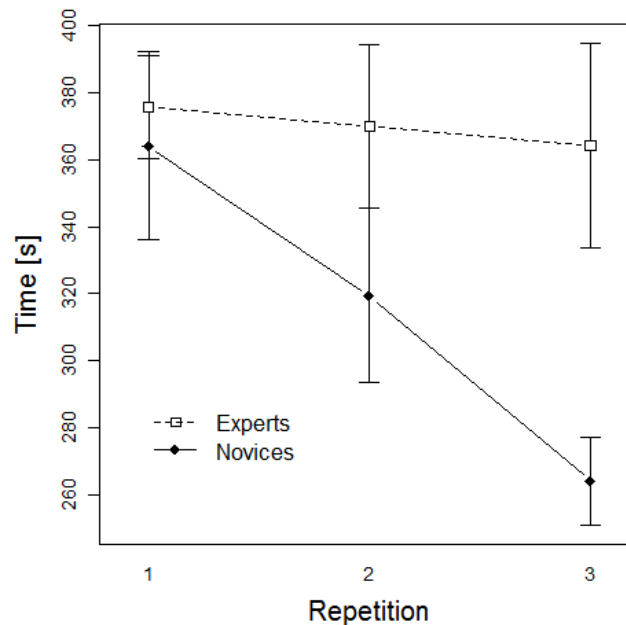
Figure 24: Time needed for the virtual training task. Error bars reflect standard errors.

In the final advanced mode training session, experts and novices made 3.9 ($SD = 1.5$) and 4.5 ($SD = 1.9$) errors, respectively ($t(33) = -0.92$, $p = .36$), when selecting the correct part out of a virtual menu. Qualitative inspection of type of errors made by the participants revealed problems with choosing the correct screw, i.e., a fine event. No participant failed to select the correct main object. Again, novices performed faster ($M = 413.5$ s, $SD = 61.5$) than experts ($M = 588.8$ s, $SD = 157.4$) in the concluding virtual training, $t(22) = 4.42$, $p < .01$.

5.4 Discussion

The main hypothesis of this experiment could be confirmed, i.e., repeating an assembly affects long-term memory acquisition of coarse and fine events differently. Coarse events benefit from repeated presentation. In contrast, I found no benefit after repeated presentation for fine events.

In the context of event cognition research, this study provides evidence that long-term memory for fine and coarse events differs and this difference is established after a short 3-repetitions training. I conjecture to explain this by differences in saliency between successive fine and coarse events, respectively. The main object is the same for consecutive fine events and they resemble each other more than consecutive coarse events. Coarse events differ from each other conceptually, i.e., because of the appearance of a new main object. This conceptual change at the coarse event boundary triggers more attention and, as a result, has higher chances on being encoded in long-term memory. Thus, repeated practice strengthens

the advantage of coarse event boundaries with respect to their long-term memory representation.

Furthermore, I found increased response times in the memory test for coarse events. In line with the Deliberate Practice Framework, longer response times may indicate more effortful experiences with coarse events which in turn promote successful learning (Ericsson, 2004). Furthermore, getting faster in predicting coarse events is in line with improved memory for coarse events.

Even if the fine events did not improve from repetition to repetition, I found an expertise effect for the fine events. That is, experts were better in memorizing fine events. This is likely due to previous experiences with several automotive objects and their assembly work over years of professional life. However, in this context, a potential speed accuracy tradeoff could be responsible for improved performance in experts. Experts were better in predicting fine events but they were slower in responding. The fact that they took more time to think about items could have resulted in better performance. However, if there was a universal speed accuracy effect, experts should outperform novices in predicting coarse events, too. This is not the case: novices and experts have approximately the same levels and curves for predicting coarse events. I conjecture to explain the difference in response time by age. The older participants reacted generally more slowly than the younger. This was true for all tasks including the virtual training task. Despite these arguments, there was an unavoidable confound between expertise and age for my goal to investigate training-relevant groups, i.e., students just before potential job training and long-term workers from automotive. Therefore, I cannot completely exclude a speed accuracy tradeoff. However, in both groups, I found the most interesting interaction effect.

This hierarchical level effect, i.e., better memory performance for coarse but not for fine events, may be further explained by the assembly's nature and the training design. First, decreased discrimination and higher response bias for fine steps across repetitions indicated that they were more likely to be confused in the memory test. Objects characterizing coarse event boundaries were so characteristic that they could be easily recognized in the video. Fine events involved smaller, less characteristic objects that were more similar to each other (e.g., screws). They were competing with each other during memory test potentially leading to memory interference (in accordance with Radvansky & Zacks, 2014, p. 37). Second, the virtual training required going through the door assembly in a fine-grained step-by-step manner without pointing to the hierarchical organization. Participants elaborated fine events in a segmented way inhibiting chunking of details (Zacks et al., 2006). The repeated execution might have reinforced confusion of details. Third, the virtual training setup is likely most suitable for communicating declarative knowledge, i.e., higher-level concepts like main assembly steps primarily represented by coarse events (Ericsson, 2008). In contrast, virtual simulations cannot teach detailed manual operations and motoric skills as effectively as hardware-based training involving real prototypes (Ericsson, 2008; Malmsköld et al., 2007).

From the results of the memory test, the conclusion could be drawn that the temporal connection of a fine event boundary towards its nearest coarse event boundary is strengthened in contrast to the other possible connection, i.e., coarse event boundary to the next fine event boundary. Prediction was enhanced for object-based coarse events from the nearest fine event boundary. However, prediction did not improve for action-based fine events from the nearest coarse event boundary. It seems that, after repetition, the connection from fine towards the next coarse event boundary is established with priority. This effect would be in line with findings that indicate that familiarity with a task increases the hierarchical structuring according to hierarchical alignment and enclosure.

The results reveal that repetition is not always beneficial for memory. Other detrimental effects of repetition have been shown by Jacoby et al. (1998) who instructed participants to detect words they read but not heard in previous study phases. Increasing the number of reading repetitions increased the difficulty in correctly rejecting a word that indeed was read but not heard. In this case, repeated presentation increased the familiarity of the read word making it more difficult to disentangle if it was additionally heard or not heard. Thus, repeated presentation may be beneficial or disadvantageous for learning depending on stimuli properties like similarity.

Since I showed that initial virtual training promotes learning of coarse assembly steps, it seems that fine assembly steps require different training strategies. Currently, training design foresees spending equal time for fine and coarse events. Instead, I propose a grouping of fine steps and an adaptation of training to expertise level. I will elaborate these aspects in Chapter 7.2.

6 Experiment 3: Cognitive potential of intellectually disabled workers in workshops for adapted work

In contrast to assembly work in the automotive domain, workers in workshops for adapted work execute highly simple and repetitive tasks. In this experiment, I study whether the event segmentation task can be used to investigate the cognitive potential of intellectually disabled workers to perform more complex and interesting tasks.

6.1 Introduction

6.1.1 Theoretical and practical relevance

Workshops for adapted work provide work places for people with different types and degrees of physical and intellectual disabilities. One of their aims is fostering workers' personal growth through manageable, interesting, and qualifying work activities. In this context, manual assembly offers multifaceted tasks that can be highly structured and simple enough but still sufficiently challenging (Richardson & Jones, 2011). In addition, manual assembly workplaces can incorporate physical and cognitive support, e.g., additional tools or interactive instructions. However, in reality many workers perform only simple, repetitive tasks (Dulaney, 1998). This could constitute a missed opportunity for individual development and yields boredom and error increment in the long run. There is not sufficient research on whether the monotonous assembly tasks offered at the workshops utilize the full cognitive potential of intellectually disabled people. In the present experiment, I propose that assessment of workers' ability to segment dynamic events into meaningful events is a way to overcome this theoretical and practical gap.

So far, assessing a worker's potential for executing assembly tasks is demanding because many different cognitive functions are involved in successful work execution (Section 3.1.2). Evidence from former research on cognitive impairments in intellectual disability (see below) provides a collection of possible deficits varying in individuals but limited insight into selected capabilities relevant for executing work in workshops. I suggest that mental representations and cognitive processes involved in assembly tasks can be comprehensively analyzed by classical event segmentation task (Newtson, 1973). Event segmentation is a complex cognitive control mechanism which combines the necessary attentional, perceptional, and memory-related processes when dividing ongoing activity into meaningful events. Importantly, event segmentation predicts actual action performance (Bailey et al., 2013). Both action and perception involve monitoring the current step in the sequence, structuring tasks in goals and sub-goals, and long-term memory for all necessary steps. Thus, capability of action perception indicates potential ability and problems of actual task performance.

The successful assessment of cognitive (dis-)abilities of the intellectually disabled group in the context of their work has significant implications for work design. First, it supports the choice of right level of task complexity. Second, it guides appropriate assistance and training activities, for instance, using upcoming computer-based methods (Section 2.2.2). Now such a measure can potentially be designed with the most recent findings on the relation between event cognition and performance (Bailey et al., 2013). By using such a measure the practical challenges in workshops for adapted work (EASPD, 2012b) and the potential of evolving technological means (Gorecky et al., 2012; Korn et al., 2013b) could be addressed. I aim to fill up this gap in psychological research with the understanding of action perception and performance in intellectually disabled people. Hence, the estimation of their assistance needs and application of appropriate support means can be guided by empirical knowledge.

6.1.2 Cognitive dysfunctions related to intellectual disability

In this section, I will shortly review potential deficits in intellectually disabled people with respect to the cognitive processes required for dynamic event perception (Section 3.1.2), i.e., visual attention, visual perception, working memory, and long-term memory.

Visual attention

The consequences of attentional deficits were shown for persons with intellectual disability (Iarocci & Burack, 1998), yet, there are large interindividual differences (Sterr, 2004). A study investigating visual information processing revealed that children have delayed visual orienting responses, e.g., in reaction to movement (Boot, Pel, Vermaak, van der Steen, & Evenhhuis, 2013).

Visual perception

Van Roon, Caeyenberghs, Swinnen, and Smits-Engelsman (2010) showed diminished performance in a group of intellectually disabled children when tracking an accelerating target red dot shown on a screen with the help of a cursor. The authors suggested a lack of anticipation for the target movement. Further deficits in anticipatory behavior of intellectually disabled persons have been reported for action planning (Crajé, Aarts, Nijhuis-van der Sanden, & Steenbergen, 2010) and movement initiation (de Campos, Cerra, Silva, & Rocha, 2014).

Working memory

Working memory plays a crucial role in intellectual functioning (Cornoldi & Giofrè, 2014). It represents the central system for storage, manipulation, and integration of different information. Working memory span may be limited (Henry, 2001) to two or three elements (Numminen, Lehto, & Ruoppila, 2001) in intellectually disabled people. Other studies treated potential impairments of working memory sub-systems, i.e., the central executive (e.g., dual

task performance), the phonological loop, and the visuo-spatial sketchpad (Baddeley & Jarrold, 2007; Lanfranchi, Baddeley, Gathercole, & Vianello, 2012).

Carretti, Belacchi, and Cornoldi (2010) aimed at investigating how well intellectually disabled people modify and update information in working memory, i.e., incorporate new information and exclude old material. Similar to the Event Segmentation Theory, they stress the importance of the ability to manage attentional resources and to dynamically adapt working memory content depending on current external changes. Their updating task differed from a classical span task by not only requiring participants to accumulate information without substitution but rather to actively control and process the content. For instance, participants were presented with a spoken list of five objects from which they had to recall the two smallest in the right order of reading. In order to perform successfully, they had to constantly compare object size of the currently read item to formerly read items in the list. In case the currently read object was small enough, previous items depicting bigger objects should be excluded and previous items still depicting potential smallest objects should be maintained. The authors provided evidence that the active attentional control and updating of memory operationalized by the mentioned task mainly discriminated intellectually disabled persons from persons with typical development (Carretti et al., 2010).

Long-term memory

Different deficits in episodic memory have been reported for intellectually disabled people (Crane & Goddard, 2008; Merrill, Lookadoo, & Rilea, 2003; Southwick et al., 2011; Stan & Mosley, 1988; Zalla et al., 2013). However, there is also evidence that long-term memory in intellectually disabled persons from workshops for adapted work improves with practice (Dulaney, 1998).

6.1.3 Overview of experiment

My research question in this experiment was whether the simple and repetitive assembly tasks offered at workshops for adapted work utilize the full cognitive potential of intellectually disabled people. In order to analyze cognitive processes related to assembly work in intellectually disabled persons from workshops for adapted work, I used the Event Segmentation Theory. It offers a comprehensive framework since it combines relevant functions from attention, perception, and memory (Kurby & Zacks, 2008; Radvansky & Zacks, 2014; Zacks et al., 2007), it provides the event segmentation task as assessment method, and it is closely connected to action performance (Bailey et al., 2013).

In order to answer the research question, I conducted a study with two groups of participants. First, the intellectually disabled group consisted of 32 workers from workshops for adapted work, with an average IQ of 64.4 (*SD* = 9.8). The second group was the control group consisting of 30 students from University of Kaiserslautern. The intellectually disabled group executed two tasks. First, they performed a classical event segmentation task on one every-day and three assembly-related activities. They segmented the videos into fine- and coarse-

grained events. I compared their event segmentation behavior with the control group and applied several event segmentation measures to the analysis.

Furthermore, the same 32 workers from workshops for adapted work performed a 7-step Lego assembly task. Beforehand, I instructed them with the help of a video and a paper manual. I assessed the errors made by them in the assembly task.

Because intellectual disability goes along with different cognitive deficits which can be linked to event segmentation (Section 6.1.2), I expected impairments in event segmentation performance as well. As event boundary perception relies on movement and conceptual changes, existing deficits of the intellectually disabled group concerning motion perception and abstract thinking (Section 6.1.2) should result in lower detection rate of event boundaries in fine and coarse conditions. Disturbed updating processes in working memory (Section 6.1.2) should also contribute to inhibited event boundary perception and extended event lengths. Further event segmentation measures assessing the hierarchical structuring should give insight into potential improvement of segmentation performance due to task familiarity (Section 3.1.5). Finally, since event perception and action performance are closely connected (Section 3.1.5), event segmentation behavior should be able to account for differences in assembly task execution measured by the Lego assembly task.

6.2   Methods

6.2.1 Participants

I received consent forms from 39 participants working in the workshops for adapted work in the Westpfalz-Werkstätten in Landstuhl, Germany. Almost all of them (N = 38) had an intellectual disability with onset during the developmental period and they were physically able to perform manual tasks. One participant had a brain damage in consequence of an accident resulting in intellectual disability.

Exclusion of incomplete segmentation data sets (N = 7) yielded to my final sample of 32 participants (13 female). On average, they were 37.6 years old (SD = 11.9) with a mean working experience in workshops of 14.0 years (SD = 11.6). Their areas of deployment were assembly (N = 27), metalworking (N = 4), or gardening (N = 1). The demographic information and performance measures are summarized in Table 8. Intellectually disabled workers were granted leave of absence during regular working time.

The control group consisted of 30 students (15 female) from University of Kaiserslautern (one student was excluded beforehand because of incomplete data) with average age of 25.2 years (SD = 3.5). They received course credit or monetary compensation for their participation.

	M	*SD*	range
Age [years]	37.6	11.9	21; 63
Working years	14.0	11.6	0.75; 39
IQ	64.4	9.8	50; 84
SRT	1.2	1.0	0.3; 4.2
Lego SE	1.7	1.9	0; 7
Lego PE	2.0	2.2	0; 7

Note. SRT = simple reaction time in seconds, SE = sequence errors,

PE = position errors.

6.2.2 Event segmentation task

In order to ensure that the intellectually disabled participants understand the classical event segmentation task procedure (Newtson, 1973), they segmented a practice video clip before the actual event segmentation. The video showed a female actor driving three nails into a block by using a hammer and screwing three screws into another block by using a screwdriver. The practice segmentation was accompanied by the following experimenter's explanations and demonstrations.

"I would like to show you videos in which you will observe a person doing a certain task. I am interested in how you perceive this activity. The activity can be divided into a sequence of steps. It is possible to divide it in many small steps but also in fewer, bigger steps. Whenever, for you, one step has ended and another has begun, please indicate it by pressing the button. We will watch a practice video together. Here, you can see a person who is driving nails into a block. If it comes to fine segmentation, some people would click each time whenever a new nail has been inserted. *Experimenter demonstrates button presses at time points in which new nails and screws are inserted.* Other people would define additional detailed steps, for instance, whenever a new nail has been taken into hand. *Experimenter has restarted the video and demonstrates the more fine segmentation by button presses at time points in which new nails and screws are both taken into hand and then inserted.* In case of a coarse segmentation, some people would only press after all nails had been inserted and before the screws were screwed in. *Experimenter demonstrates the coarse segmentation.* Other people would maybe click differently for defining big steps. So, there are different possibilities how people perceive the activity in the video and how they divide it. There is no wrong or right way. It is important that you press the button whenever you think that one step has ended and a new one has begun. I am now interested in how you would segment the video with the nails and screws. […] Do you have any more questions concerning the procedure?"

Afterwards, all participants segmented four video clips (see Figure 25) in fine- and coarse-grained meaningful events. As in the classical event segmentation task (Newtson, 1973), they pressed the space key on a keyboard whenever they thought one meaningful event ended and another began. First, they saw the "breakfast" video in which an actress is preparing breakfast in a kitchen (see e.g., Swallow, Zacks, and Abrams (2009)). In the remaining three videos, an actor was executing an assembly task, respectively. The "valve"

video depicted an assembly typical for workshops for adapted work. The "pump" video contained an assembly from a soap factory (see also Experiment 1 and screenshots in Figure 6). The "saw" video depicted assembly of small parts from a technical construction tool kit. Figure 25 shows a representative frame for each video.
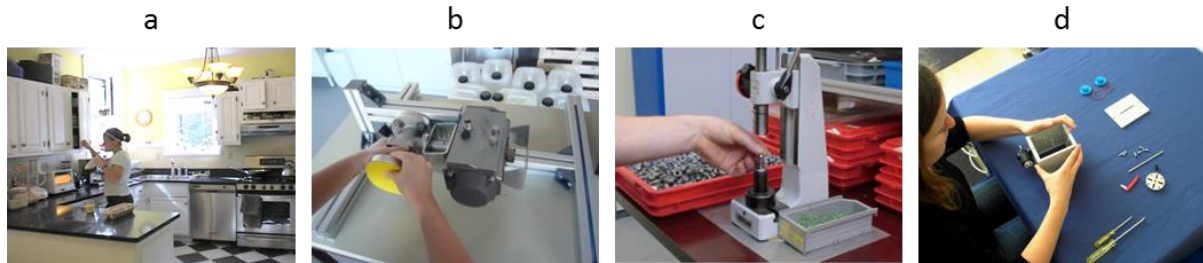
Figure 25: Video material used in the event segmentation task: (a) Breakfast (335 s): Everyday activity in which a female actor prepares breakfast; (b) Pump (195 s): Assembly task in which parts of a pump are put together; (c) Valve (93 s): Work-related task of the workshops for adapted work in which a male actor assembles small parts of a valve; (d) Saw (190 s): Female actor assembles parts of a technical construction tool kit in order to build a saw.

6.2.3 Lego assembly task

I introduced a naturalistic task, i.e., a 7-step Lego car assembly (Figure 26), consisting of conventional, different-colored Lego bricks and wheels used in previous studies (Korn et al., 2013a). The setup included 8 bricks of a Lego car lying at a table covered by a white paper sheet. First, in order to introduce the task to the participant, the experimenter lifted the sheet for 3 seconds and showed a picture of the completed Lego car on the computer screen for 5 seconds. Second, the experimenter told the participant to watch a video carefully in which he/ she will see how to build the car step by step. The experimenter replayed the video and showed a pictorial summary similar to Figure 26 on a paper sheet for 7 seconds. Then, she instructed the participant to "assemble the Lego car in the same sequence as was shown".
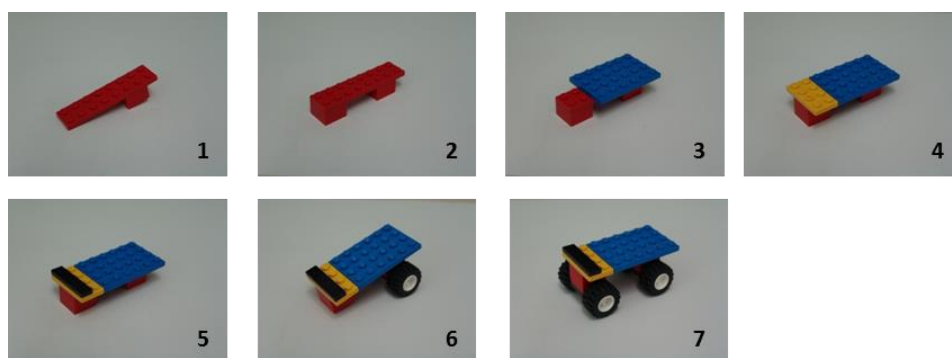
Figure 26: Seven steps to assemble a Lego car (adapted from Korn et al. (2013)).

All participants executed the assembly by themselves with the experimenter noting the accomplishment and sequence of steps on an observation sheet without giving additional help. In case a participant asked a question, the experimenter said that all what he/she needs was lying on a table and he/ she should please try as good as he/ she could (procedure adapted from Schwartz, Segal, Veramonti, Ferraro, & Buxbaum, 2002). One out of all participants remained inactive even with three-time encouragement.

I assessed both the number of sequence errors and the number of position errors made by intellectually disabled participants. I defined sequence error as an incomplete step, that is, when the person did an assembly step earlier or later or not at all compared to the original sequence shown previously. A position error was defined as misplacement of a part, that is, when the subject assembled the Lego brick not at the exact location with respect to the other bricks. Both error scores may range from 0 to 7, respectively.

6.2.4 General ability assessment

Simple reaction time

I assessed the mean simple reaction time for each intellectually disabled participant by instructing him/ her to press the space key as soon as they saw a pictogram appearing on the screen. After 10 practice pictures, which did not enter final data analysis, participants viewed 36 different pictograms of objects, buildings, and animals twice and in full random order. Before each presentation, a fixation cross appeared for the duration of 500 ms. Then, the pictogram appeared after a randomly determined interval between 1 and 2 s.

General cognitive ability

Intellectually disabled participants completed the nonverbal Wiener Matrizen-Test-2 (WMT-2) consisting of 18 matrices (Formann et al., 2011). The experimenter noted the participant's answers. For every intellectually disabled individual, I computed the IQ score as the indicator of their general cognitive ability.

6.2.5 Procedure and design

This experiment was approved by the local ethics commission of the University of Kaiserslautern.

The experiment with intellectually disabled participants took place in a separated room at the workshops site. Each participant came for two sessions within two weeks. In the first session, they saw and practiced event segmentation with the practice video clip and performed the event segmentation task successively for the four videos (breakfast, valve, pump, and saw) in fine and coarse grains, respectively, counter-balanced across participants. Written event segmentation instructions were presented on the screen. To ensure adequate understanding of the task (despite potential literacy problems), the experimenter read aloud the instruction to the participant. Videos were shown on a Notebook PC running PsychoPy software (Peirce, 2007). Participants could use the keyboard for indicating a new event or, alternatively, use the mouse in case they showed difficulties with the keyboard in the practice phase. All participants performed the event segmentation task with the keyboard. The whole session took 1 hour.

In the second session, I showed intellectually disabled participants how to build the Lego car using a Notebook PC and a paper sheet. After they completed the assembly on their own, they were asked to do the simple reaction task on a Laptop PC using the space bar since all of them managed usage of the keyboard before. Finally, they answered the WMT-2. The second session also took one hour.

The control group consisting of student participants completed only the event segmentation task. The process was similar to the intellectually disabled participants without the read aloud instructions by the experimenter and without a practice phase. I collected data of the control group at the University of Kaiserslautern. The session took approximately 35 minutes.

6.2.6 Data analysis and statistical methods

Segmentation data of the intellectually disabled group and the control group were analyzed by the whole range of existing event segmentation measures introduced in Section 3.1.4. Overall, I used the following measures. For each participant, I computed the **number of events** in fine and coarse condition, respectively, as well as their difference and ratio. I tested for **significant event boundaries** for both grains and both groups. I investigated the **segmentation agreement** between the intellectually disabled group and the control group by, first, using point-biserial correlations, second, correlating both groups' histograms, third, analyzing pair-wise kappa's, and, fourth, computing the differences in both groups' segmentation magnitudes. I calculated **hierarchical alignment** and **hierarchical enclosure** in order to investigate participants' ability to hierarchically structure activities.

Again, I used R (R Development Core Team, 2008) for all statistical analyses and additional R package segmag (Papenmeier, 2014) for the continuous segmentation analyses, i.e., plotting segmentation magnitudes, determining significant event boundaries, and subtracting groups' segmentation data.

6.3 Results

6.3.1 Event segmentation ability

Number of events

I compared segmentation behavior in fine versus coarse conditions. Correspondent to the instructions, the intellectually disabled group segmented descriptively more fine ($M = 16.8$, $SD = 14.1$) than coarse ($M = 11.6$, $SD = 15.8$) events, however, without significance ($t(61.36) = 1.58$, $p = .12$). Yet, they showed the expected difference concerning event length. They defined shorter fine ($M = 19.8$ s, $SD = 13.7$) than coarse events ($M = 42.5$ s, $SD = 39.7$), $t(38.42) = -3.60$, $p < .01$. The controls showed both expected differences significantly. In the fine condition, they segmented more events ($M = 24.5$, $SD = 15.1$) with shorter event length ($M = 10.9$ s, $SD = 5.2$) compared to the coarse condition where they defined less events ($M =$

7.3, SD = 3.6) with larger event length (M = 32.6 s, SD = 16.0), $t_{number\ of\ events}$(31.96) = 7.66, p < .01 and $t_{length\ of\ events}$ (39.18) = -11.00, p < .01.

Table 9: Event segmentation results aggregated over four videos

Segmentation measure	Persons with ID (*N = 32*)	Controls (*N = 30*)	
	M (*SD*)	*M* (*SD*)	*t, p*
Number of events			
coarse	11.6 (15.8)	7.3 (3.6)	-1.72, .09
fine	16.8 (14.1)	24.5 (15.1)	2.49, *
Difference of fine – coarse events (F –C)	3.97 (5.37)	12.37 (7.45)	5.06, **
Ratio of fine / coarse events (F / C)	2.5 (1.82)	3.6 (1.60)	2.58, *
Event length			
coarse	42.5 (39.7)	32.6 (16.0)	-1.59, .12
fine	19.8 (13.7)	10.9 (5.2)	-4.05, **
Segmentation agreement[a] [group: controls]			
coarse	.24 (.15)	.61 (.12)	9.28, **
fine	.40 (.13)	.71 (.11)	10.10, **
Segmentation agreement[a] [group: persons with ID]			
coarse	.60 (.12)	.54 (.09)	2.18, *
fine	.62 (.10)	.62 (.05)	.36, .72
Mean pairwise correlation[b] [within group]			
coarse	.03*	.14*	-
fine	.03*	.19*	-
Mean pairwise correlation[b] [between groups]			
coarse		.05*	-
fine		.06*	-
Hierarchical alignment[ce]			
Observed mean distance [s]	3.50 (3.90)	1.30 (.91)	
Expected mean distance [s]	7.00 (5.85)	3.54 (2.26)	
Observed – expected mean distances	3.68 (5.24)	2.30 (2.13)	-1.70, .10
Hierarchical alignment[de] – eCALC	.55 (.14)	.63 (.08)	-2.92, **
Hierarchical enclosure[e]	.56 (.29)	.67 (.21)	3.12, **

Note. ID = Intellectual disability. eCALC = exponential function calculation. [a]Point-biserial correlation between individual and group. [b]Cohen's kappa with significance tests by constructing bootstrap confidence intervals. [c]Computation based on Zacks et al. (2001). [d]Computation based on exponential transformation of distances (see text below). [e]Analysis involved subgroup of N=20 intellectually disabled participants with a correct difference of F – C > 0. *p < .05, **p < .01.

Furthermore, I calculated the difference between number of fine and coarse events for each participant. The controls ($M_{difference\ F-C}$ = 12.4, $SD_{difference\ F-C}$ = 7.5) showed a higher difference than the intellectually disabled group ($M_{difference\ F-C}$ = 4.0, $SD_{difference\ F-C}$ = 5.4), t(60) = 5.06, p < .01 (compare Figure 27). The differences in number of events between both groups was further evident for each grain, i.e., fine (t(60) = 4.18, p < .01) and coarse segmentation (t(60)

= -3.00, $p < .01$), respectively (see Table 9). From these initial results, it seems that the intellectually disabled group did not segment the activities as clearly in accordance with two hierarchical levels as the control group did. Figure 27 illustrates the reduced difference in number of events between fine and coarse condition when comparing intellectually disabled participants with the controls.
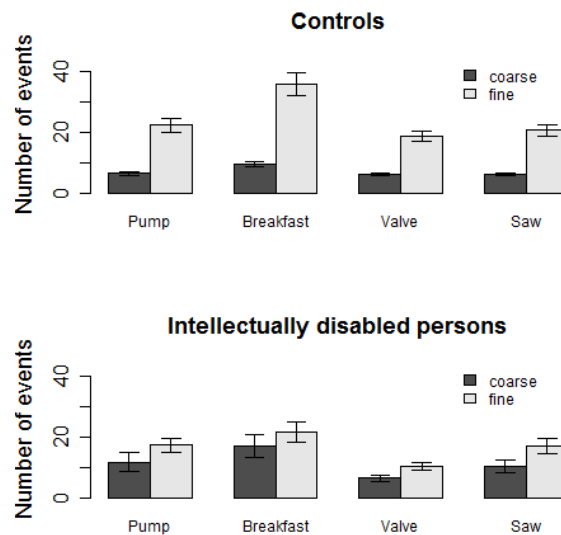

Figure 27: Mean number of events per video for each group. Error bars reflect standard errors.

In order to analyze this finding in more detail, I also calculated the ratio of number of fine to coarse events per subject. Again, there were significant group differences between the control group ($M_{\text{ratio F/C}} = 3.60$, $SD_{\text{ratio F/C}} = 1.60$) and the intellectually disabled group ($M_{\text{ratio F/C}} = 2.50$, $SD_{\text{ratio F/C}} = 1.82$), $t_{\text{ratio F/C}}$ (60) = 2.58, $p < .05$. Whereas the control group showed a ratio value similar to previously reported values (Zacks et al., 2001), the mean value of the intellectually disabled group was diminished. In more detail, I looked at the individual level (see Figure 28) in order to reveal potential interindividual differences within the intellectually disabled group. Figure 28 illustrates that there were intellectually disabled participants whose ratio values were similarly high compared to those of the controls. However, $N = 12$ participants showed ratio values around 1 or even lower. This means that they pressed more frequently or equally frequently in the coarse compared to the fine condition.

Repeating the previous group comparison after excluding those $N = 12$ "outliers" made the significant group differences in ratio and difference between the controls and the intellectually disabled participants disperse ($t_{\text{difference}}(48) = 1.04$, $p = .30$ and $t_{\text{ratio}}(48) = 1.17$, $p = .25$). These results suggest that the intellectually disabled group can be divided into two sub-groups, i.e., a group of $N = 20$ persons with correct understanding of different hierarchical levels and a group of $N = 12$ persons with a misconception of hierarchical organization.
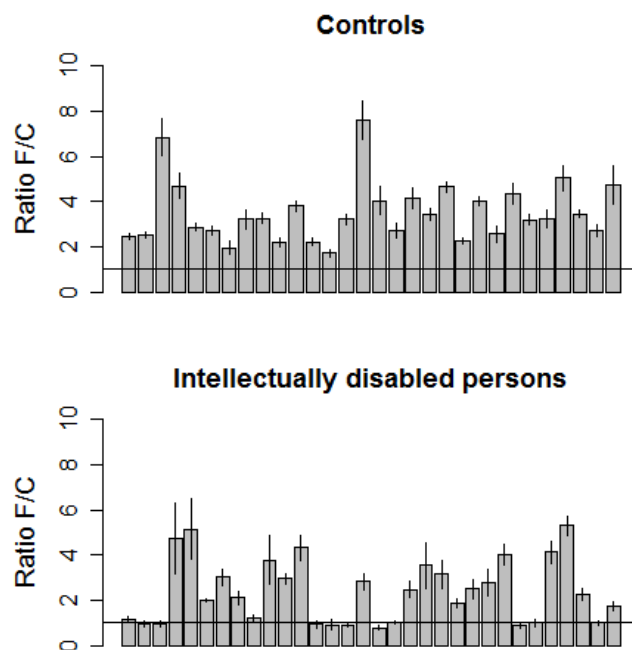
Figure 28: Ratio of fine to coarse events for each individual from both groups. A ratio of 1 is highlighted as a black horizontal line, respectively. Error bars reflect standard errors.

Significant event boundaries

In addition to number of events, I was interested in the exact locations of event boundaries. Figure 29 displays segmentation behavior across time in the pump video for both groups during the fine condition. The upper and middle segmentation plots (Figure 29) indicate significant chronological correspondence *within* groups highlighted by green lines. I confirmed these chronological correspondences within groups by correlating individual segmentation behavior with own group's segmentation. The mean point-biserial correlation over all intellectually disabled participants was $r = .60$ in the coarse and $r = .62$ in the fine condition. The control group showed a mean point-biserial correlation of $r = .61$ in the coarse and $r = .71$ in the fine condition (see also Table 9).

Despite this evidence for within-group agreement, the upper and the middle plots suggest that the controls clearly agreed on location of event boundaries whereas the intellectually disabled group was less consistent and noisier. These observations were further supported by the mean pairwise kappas within pairs of the same group. These pairwise kappas were significant and substantial for the control group ($k_{coarse} = .14$, $k_{fine} = .19$). Intellectually disabled participants showed significant kappas close to zero ($k_{coarse} = .03$, $k_{fine} = .03$). Thus, the intellectually disabled participants agreed on common event boundaries. However, they were not as homogeneous as the control group.

Figure 29: Event segmentation plot for the pump task in the fine condition: The controls (upper plot) and the intellectually disabled group (middle plot). Significant event boundaries (confidence level of 95%) are displayed as vertical (green) lines. The lower plot depicts the difference when subtracting controls and intellectually disabled participants. Vertical (green) lines represent significant differences (confidence level of 99%).

The findings just described applied for the coarse condition (see Figure 30), too, and were consistent across all four videos. Participants from the control group agreed upon more significant event boundaries than participants from the intellectually disabled group. Segmentation plots of the intellectually disabled group were noisier.

In addition to the within-group correspondences, I was interested if event boundaries were similar *between* both groups. I will present these results in the next section.
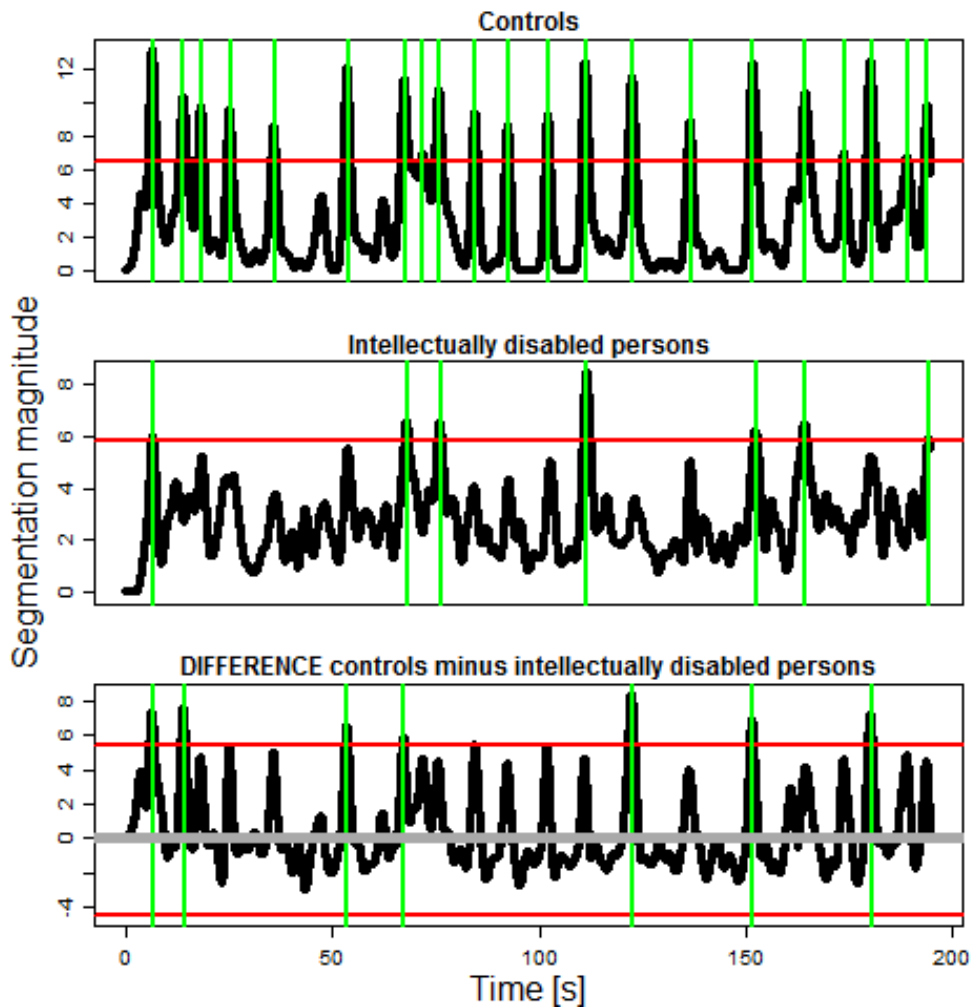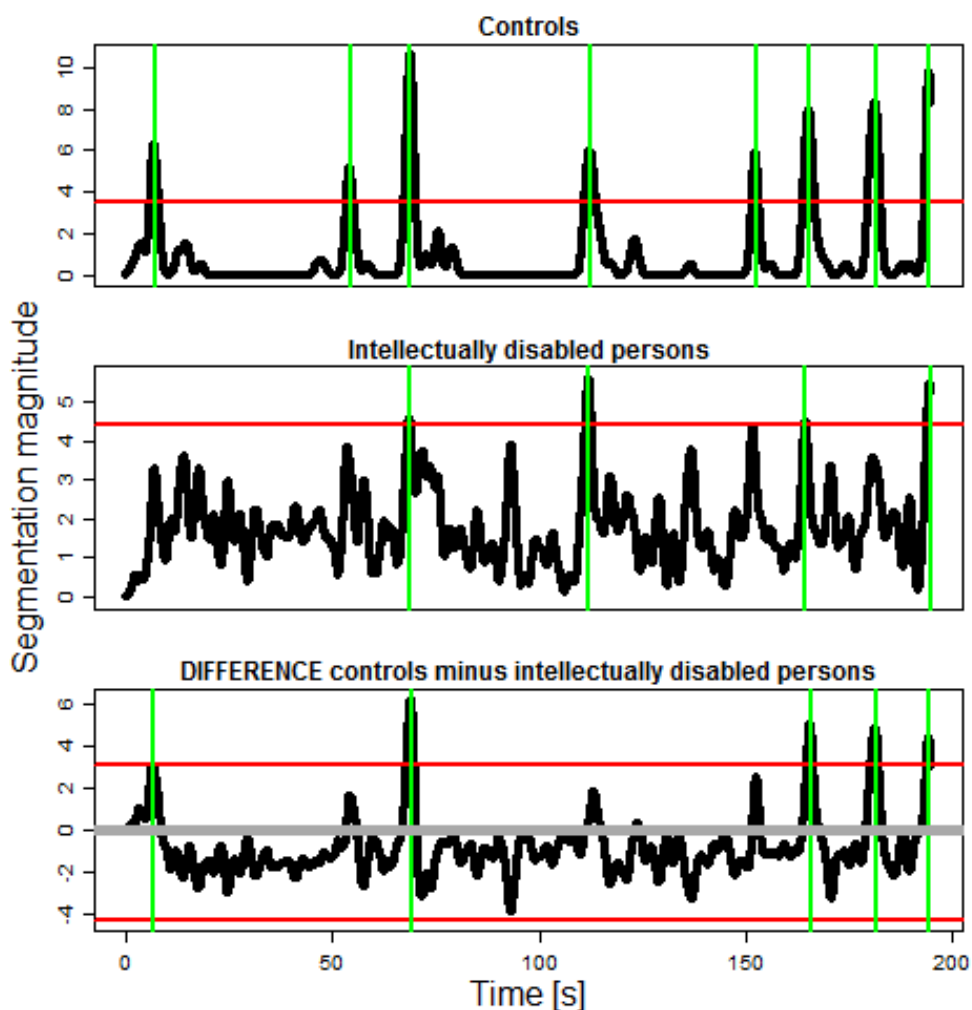
Figure 30: Event segmentation plot for the pump task in the coarse condition: The controls (upper plot) and the intellectually disabled group (middle plot). Significant event boundaries (confidence level of 95%) are displayed as vertical (green) lines. The lower plot depicts the difference when subtracting the controls and the intellectually disabled participants. Vertical (green) lines represent significant differences (confidence level of 99%).

Segmentation agreement

Initial inspection of Figure 29 and Figure 30 suggests an agreement between the controls and the intellectually disabled persons concerning event boundaries because each event boundary found for the intellectually disabled group has a corresponding event boundary in the control group. In the following, I will analyze this agreement between groups quantitatively.

To test if the intellectually disabled participants pressed at similar locations compared to the control group, I looked at the groups' segmentation agreement between individuals from the intellectually disabled group with the control group calculated by point-biserial correlations. I found meaningful mean point-biserial correlations between persons from the intellectually disabled group and the control group as a whole (r_{coarse} = .24, r_{fine} = .40). Correlating both groups' histograms using Pearson's r confirmed the overlap between groups (r_{coarse} = .41, $p <$ .01 and r_{fine} = .48, $p <$ .01). Further evidence for chronological correspondence between the

intellectually disabled participants and the controls came from between-group pairwise kappas (k_{coarse} = .05, k_{fine} = .06) which were significantly above 0.

Despite the overlap, subtracting segmentation plots (compare lower plots in Figure 29 and Figure 30) enables illustration of disagreement represented by segmentation magnitude differences. As can be seen, there were event boundaries that the control group perceived but the intellectually disabled group did not, i.e., significant positive differences. In contrast, there were no event boundaries that intellectually disabled persons perceived compared to the controls, i.e., no difference was significantly negative. Despite the already discussed dependence of the difference method on given segmentation magnitudes and group sizes, it provided an additional graphical indication of exact time points of segmentation agreement and disagreement.

As reported in the previous section, intellectually disabled participants were not homogeneous within their group. Therefore, my final question was whether they were more similar to each other than to the control group. For this, I analyzed whether pairs of participants from the same group chose boundaries that were more similar than pairs of participants from different groups by subtracting kappas taken from pairs from the same group and pairs from different groups. For intellectually disabled participants the differences were not statistically significant neither for fine (k = -.03) nor coarse segmentation (k = -.02) indicating that intellectually disabled participants did choose boundaries that were not more similar to their own group than to control group. For the controls, the differences were statistically significant for fine (k = .13) and coarse segmentation (k = .10) indicating that agreement within the control group is higher than their agreement with the intellectually disabled group. Thus, the intellectually disabled persons cannot be seen as similar to each other concerning their segmentation behavior. Rather, it is a heterogeneous group of different "segmenters".

Hierarchical alignment and enclosure in the intellectually disabled group compared to the control group

I quantified ability to hierarchically structure dynamic activities by hierarchical enclosure and alignment, respectively. Since interrelating coarse with fine event boundaries makes sense only for participants who correctly pressed more often in the fine than in the coarse condition, I excluded those N = 12 intellectually disabled participants who did not meet this requirement from all further computations. The analysis of hierarchical enclosure showed that, in the control group, more than half of nearest fine event boundaries were hierarchically enclosed to their respective coarse event boundary (M = 0.67, SD = .21), t = 9.09, $p < $ .01. Intellectually disabled participants had a significantly lower enclosure value (M = 0.56, SD = .29) than the controls, $t(140)$ = 3.12, $p < $ .01 (Table 9), and it was not significantly different from 0.5 (t = 1.74, p = 0.08). This was initial evidence for a detriment in the intellectually disabled group with respect to hierarchical perception.

Further, I analyzed the hierarchical alignment between fine and coarse event boundaries per subject, i.e., if they were temporally close to each other. High temporal closeness between fine and coarse event boundaries is an indicator of ability to hierarchically structure dynamic content. First, I calculated the observed mean distance in time between coarse and nearest fine event boundaries. Then, I computed the expected mean distance under the assumption that the key presses in the coarse and the fine condition were independent. As can be seen in Table 9, observed mean distance was significantly smaller than expected mean distance for both intellectually disabled participants ($M_{observed\ distance}$ = 3.50, *SD* = 3.90 versus $M_{expected\ distance}$ = 7.00, *SD* = 5.85, *t*(79) = -5.74, *p* < .01) and the controls ($M_{observed\ distance}$ = 1.3, *SD* = .91 versus $M_{expected\ distance}$ = 3.54, *SD* = 2.26, *t*(119) = -11.24, *p* < .01) indicating that for both of the groups segmentation was more temporally aligned than would be expected by chance.

I compared the intellectually disabled participants and the controls for hierarchical alignment by calculating the difference between observed and expected mean distance for each participant (Zacks et al., 2001). The higher this value gets, the better the hierarchical alignment. I found a descriptively higher average alignment value for the intellectually disabled group (*M* = 3.68, *SD* = 5.24) compared to the controls (*M* = 2.30, *SD* = 2.13) without significance (*t*(22.6) = -1.70, *p* = .10). Nevertheless, this group result was hard to explain because I would have expected that the controls show higher hierarchical alignment than intellectually disabled participants. The potential problem could be the assumption that this model makes about the relationship between distances and alignment. The distances between coarse and nearest fine event boundaries are treated linearly, i.e., a fixed proportionality constant is assumed between distance and alignment.

However, intuitively, the further away a nearest fine event boundary is from its coarse event boundary, the less likely it is related to it. Therefore, I postulate an exponential relationship between distance and alignment: with increasing distance between the coarse and the nearest fine event boundary temporal closeness exponentially decays to zero. Consequently, I transformed observed distances ("d") according to an exponential function: e^-d.

Figure 31 displays exemplary, empirical fine event boundaries (f_1 to f_{13}) of a representative participant across time. The curve illustrates exponential distance values on the y axis depending on location of coarse event boundaries which are c_1 to c_3. *Observed* mean exponential distance is, therefore, the average of y values at the locations of c_1 to c_3 representing the exponential distance between f_4 and c_1, f_8 and c_2, and f_{11} and c_3. In order to derive the *expected* mean exponential distance, I assumed that – like in the classical computation (Zacks et al., 2001) – coarse event boundaries were defined by pure guessing. This yields to an expected mean exponential distance that equals the area under the upper curve in Figure 31, called "a" here (J. M. Zacks, personal communication, November, 2014). Observed and expected exponential mean distances can be combined into one formula by including "a" as factor: e^-d*a.

Finally, I introduced a smoothing factor of 1/75 so that the alignment calculation fits a decline to close-to-zero for distances higher than 7.5 s (J. M. Zacks, personal communication,

November, 2014). The underlying assumption is that nearest fine and coarse event boundary with a distance exceeding values of 7.5 s cannot be interpreted as temporally close anymore. Typical mean empirical distances have been around 2-3 s (see Section 3.1.4 and Experiments 2 and 3 in the present dissertation). So, I am defining zero alignment starting from distances that are more than twice as large as the "usual" distances.
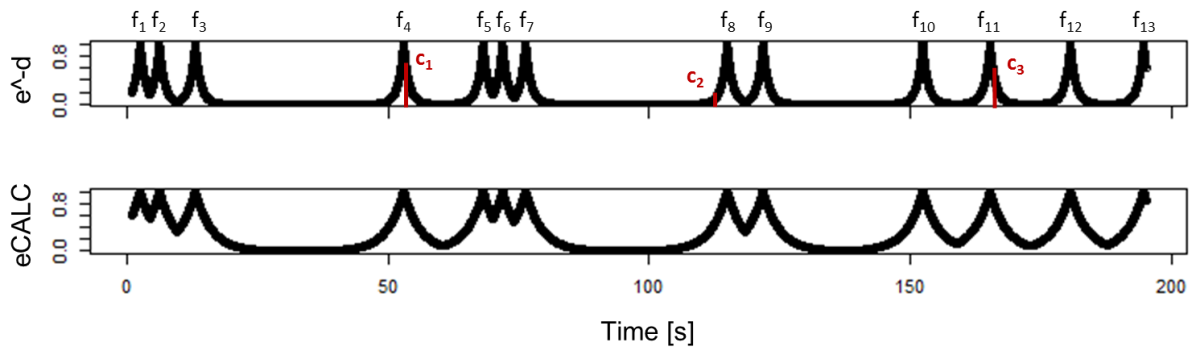
Figure 31: Illustration of the hierarchical alignment with exponential function calculation, eCALC = e^-d*a/75: The upper plot shows perceived fine (f_1 to f_{13}) and coarse (c_1 to c_3) event boundaries of one participant. The observed mean exponential distance is the average of exponential distances from coarse to their nearest fine event boundaries (see y values at time points c_1, c_2, and c_3 indicated by vertical (red) lines). The expected mean exponential distance under the assumption of independence between fine and coarse event boundaries is the area under the upper black curve, "a". The lower plot displays the curve of the final eCALC method: Combining observed and expected mean distances together with a smoothing factor of 1/75 results in the final calculation term for hierarchical alignment, e^-d*a/75.

The new hierarchical alignment measure with exponential function calculation (eCALC), that is, e^-d*a/75, may range from 0 (no alignment) to 1 (perfect alignment). It positively correlates with the classical computation by Zacks et al. (2001); that is, $r = .44$, $p < .05$ for the controls and $r = .69$, $p < .01$ for the intellectually disabled participants (see Table 10). This overlap of the eCALC method with the classical hierarchical alignment measure indicates its validity.

I compared hierarchical alignment for both groups with eCALC and found, as expected, significantly higher hierarchical alignment with eCALC in the controls ($M = .63$, $SD = .08$) compared to intellectually disabled participants ($M = .55$, $SD = .14$), $t(50) = -2.92$, $p < .01$. A general shortcoming of the hierarchical alignment is its dependence on number of key presses, i.e., higher numbers of key presses statistically yield to decreased alignment values (Zacks et al., 2001). In order to evaluate whether this statistical relation holds for eCALC, too, I performed a simulation (see next section). Anticipatory results, the statistical disadvantage for higher key presses applies indeed for eCALC. Therefore, I tested if this relation worked in favor or against the just reported group difference in hierarchical alignment. Although the control group showed higher number of key presses (Table 9), they reached a higher score in hierarchical alignment than the intellectually disabled group.

Hence, the found group effect in hierarchical alignment is valid, i.e., the intellectually disabled group perceived less temporal closeness than the control group. Together with impaired hierarchical enclosure, these results indicate reduced ability to hierarchically structure tasks into goals and sub-goals.

Hierarchical alignment in the intellectually disabled group for each video

Since hierarchical alignment improves with task familiarity (Zacks et al., 2001), I was interested in a potential improvement of hierarchical alignment for videos depicting assembly tasks compared to the rather unfamiliar breakfast video (Figure 25). I found no significant effect of video on hierarchical alignment with eCALC for the intellectually disabled participants, $F(3, 72) = 0.71$, $p = .55$. However, it could be that the dependence on number of key presses dilutes a potential effect. Since the videos differ in length, they also yielded different numbers of key presses. In order to control the interference of key presses, I simulated the relationship between key presses and hierarchical alignment. Afterwards, I will present a graphical way to compare hierarchical alignment between videos with different video length and, consequently, different numbers of key presses.

First, I verified the negative correlation between hierarchical alignment with eCALC and number of key presses by a simulation in which I assumed a fixed temporal closeness of the coarse and its nearest fine event boundaries. The simulation was executed for each of the four videos and will be described using the pump video as example. (Therefore, for the following explanations, please refer to the upper right graph of Figure 32 labeled as "pump".)

Given a time length equal to the pump video (195 s), I simulated random key presses in fine condition ranging in number from 3 to 60[2] (incrementally increasing in steps of 3). I randomly defined 1/3[3] of fine event boundaries as being a "nearest fine event boundary". The location of the corresponding coarse event boundary was set by a Kernel density function with SD = 1 s (this value represents a given hierarchical alignment which is high). Then, I computed the values for the hierarchical alignment between the notional fine and coarse event boundaries according to eCALC. For each of the chosen numbers of fine key presses, I repeated this simulation and computation 100 times.

The resulting average alignment values are displayed below (see upper (green) points in Figure 32). Additionally, I plotted a smoothing line connecting these points by using locally-weighted polynomial regression according to Cleveland (1981) and the respective R function (R Development Core Team, 2008). The resulting upper (green) line shows that despite the fixed distance setting between nearest fine and coarse event boundaries (SD = 1 s), alignment decreased with number of key presses. I repeated the simulation. This time, I choose a low given temporal closeness between nearest fine and coarse event boundaries (setting SD = 5 s) resulting in average alignment values depicted as bottom (red) points. As should be expected, this computation resulted in a lower hierarchical alignment value (see red line in Figure 32). The relationship between number of key presses and corresponding alignment value pointed in the same direction, i.e., a negative correlation. In sum, no matter if persons perceive high or low hierarchical alignment, their alignment value will automatically decrease with increasing number of key presses.

---

[2] The chosen spans cover the empirical ranges found in this experiment, respectively.
[3] This value approximates the ratio found in the present experiment and in prior reports (Zacks et al., 2001).

The resulting bottom (red) and upper (green) lines for each video depict marginal conditions of low and high hierarchical alignment. Next, I added the empirical values from intellectually disabled person to the plots, respectively (Figure 32). If alignment was high, than individual points should be close to the upper (green) line; if alignment in a video is low, the points would be close to the bottom (red) line. Furthermore, if alignment gets better with familiarity than the plots for the assembly videos should contain more close-to-the-upper empirical points compared to the plot for the breakfast video. The inspection showed the expected pattern: the unfamiliar breakfast video led to many values that are close to the bottom (red) line. In contrast, for all assembly videos more participants had good alignment values.
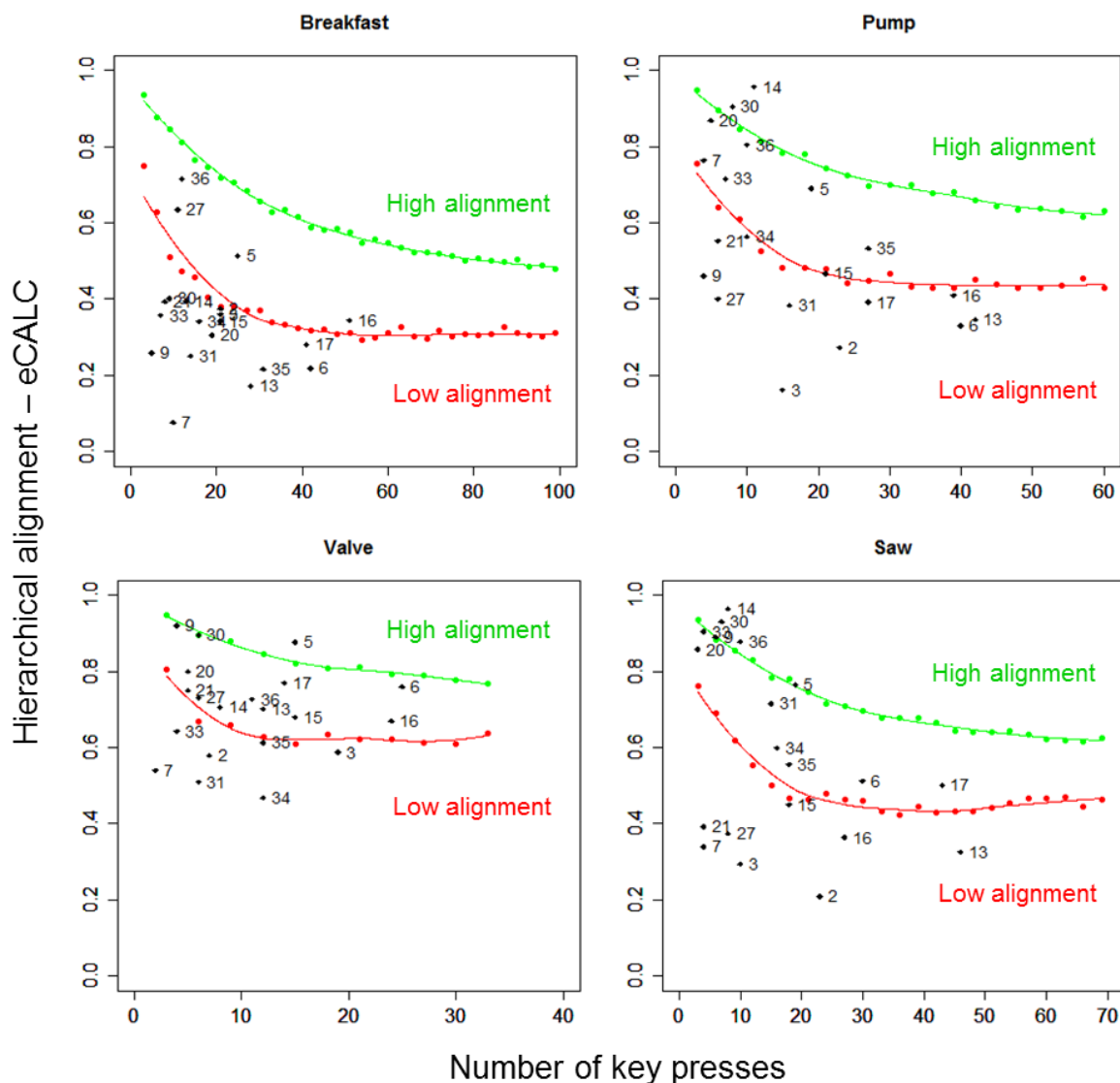
Figure 32: Hierarchical alignment with exponential function calculation (eCALC) is dependent on key presses: The simulation shows that given high (upper, green line) and low (lower, red line) hierarchical alignment values decline with decreasing numbers of key presses. Empirical values of the intellectually disabled group are added as (black) dots and subject identification.

Furthermore, there were interindividual differences. Some participants had predominantly good alignment values, i.e., in at least two videos they are close to the green line; for instance, participants with the numbers 30, 20, 14, 5, and 36 (see Figure 32). Other

participants show predominantly low hierarchical alignment, e.g., participants with the numbers 21, 30, 3, and 17 (see Figure 32). Thus, the ability to hierarchically structure is both person- and task-dependent.

6.3.2 Segmentation ability, assembly performance, and IQ

After I presented the detailed analyses of event segmentation behavior in the intellectually disabled participants and demonstrated that it is a heterogeneous group, I aimed at investigating whether the found interindividual differences in event segmentation account for differences in Lego assembly performance.

First, the following Figure 33 displays the errors in the Lego assembly task, i.e., sequence and position errors (see also Table 8). Both distributions indicate that, overall, the Lego assembly was a manageable task for the intellectually disabled group. The majority of them showed zero or only one error. This was likely due to the detailed instructions given beforehand (Section 6.2.3). Nevertheless, there was limited variation so that correlation analyses could be performed.

The correlation analyses (Spearman's rank correlations) revealed that the ratio between fine and coarse events was most promising to account for differences in action execution, i.e., the number of sequence errors. The better the conception of hierarchical organization, the fewer errors participants made. However, the correlation scarcely missed significance ($r = -.35$, $p = .05$). Then, I analyzed another measure of hierarchical organization, i.e., the hierarchical alignment (note that the correlation was performed for a sub-group of $N = 20$). The relationship with sequence errors was along the same lines ($r = -.17$, $p = .49$) but not significant. Table 10 summarizes all correlation analyses. In sum, I could not find a clear connection between different event segmentation measures and the performance in the Lego assembly task.
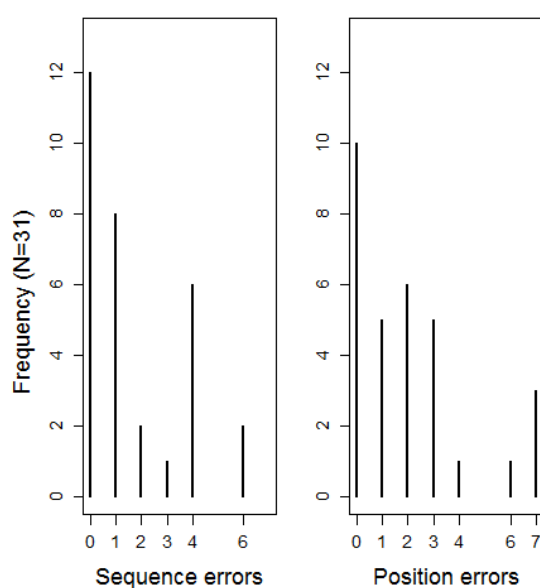

Figure 33: Lego assembly task performance: Sequence and position errors.

Finally, I investigated the relation between event segmentation measures and intelligence. As can be seen from Table 10, level of IQ correlated in the expected directions with event segmentation. The higher the IQ, the higher the difference ($r = .44$, $p < .05$) and the ratio ($r = .52$, $p < .05$) between fine and coarse events, respectively. Furthermore, IQ was positively associated with performance in the Lego assembly task. The higher the IQ, the lower the number of sequence ($r = -.40$, $p < .05$) and position errors ($r = -.53$, $p < .05$), respectively.

Table 10: Spearman Rank correlation matrix (N = 32 intellectually disabled persons)

	IQ	Age	SRT	Lego SE	Lego PE	Diff F – C	Ratio F / C	SA [F]	SA [C]	HA	HA eCALC
Age	.09										
SRT	-.54**	.20									
Lego SE	-.40*	.04	.11								
Lego PE	-.53*	-.21	.28	.55*							
Diff F – C	.44*	.02	-.42*	-.27	-.25						
Ratio F / C	.52*	.00	-.46**	-.35(*)	-.19	.87**					
SA[a] [F]	.34(*)	-.14	-.53**	-.07	-.15	.32(*)	.34(*)				
SA[a] [C]	.15	.00	-.21	.12	-.03	.18	.22	.68**			
HA[bc]	-.03	.04	.04	.19	.36	-.67**	.05	-.02	.27		
HA[c] eCALC	.22	-.10	.02	-.17	.13	-.09	.43(*)	.35	.51*	.69**	
HE [c]	-.34	.17	-.07	-.06	-.09	-.38(*)	-.19	.05	-.14	.24	.03

Note. Diff = Difference, SE = Sequence errors, PE = Position errors, SA = Segmentation agreement, F = Fine, C = Coarse, HA = Hierarchical alignment, eCALC = exponential function calculation. [a]Point-biserial correlation of individuals from the intellectually disabled group with the controls. [b]Computation of hierarchical alignment is based on Zacks et al. (2001). [c]Analysis involved subgroup of N=20 intellectually disabled participants who pressed more often in fine than in coarse event segmentation (F – C > 0). (*)$p < .10$. *$p < .05$. **$p < .01$.

6.4 Discussion

In this experiment, I investigated event perception and action execution in intellectually disabled assembly workers. I aimed at evaluating whether the simple and repetitive assembly tasks offered at workshops for adapted work utilize their full cognitive potential. First of all, it became clear that the group of intellectually disabled people is not homogeneous but interindividual differences are predominant. Despite their meaningful heterogeneity (which will be discussed later), I came to the conclusion that intellectually disabled participants are capable of performing more complex and interesting tasks than they have been executing in their every-day lives. The first empirical support for this claim is good performance in the Lego assembly consisting of seven assembly steps. After participants received a detailed instruction, i.e., video and paper directions, around half of them executed the task without errors. Hence, they showed to have the potential to cope with a more complex task given instructional support. The second finding pointing at the cognitive potential of intellectually disabled participants is their performance in the event segmentation task. Despite a number of difficulties (see below), they showed significant agreement with the

control group with respect to the location of event boundaries. In other words, intellectually disabled participants have the cognitive potential to detect event boundaries.

However, I also found problems during the event segmentation task. First, intellectually disabled participants repeatedly neglected the end of an event, i.e., they maintained an old event model without updating it. Second, the intellectually disabled group had problems with the conception of hierarchical organization of dynamic activities. This was evident in one sub-group of participants who pressed the key button more frequently during the coarse compared to the fine condition. It remains open if their misconception was due to lack of conceptual knowledge about different hierarchical levels or just misunderstanding of the task instructions. The other group who managed to press according to two different grains showed a diminished hierarchical alignment and enclosure between coarse and fine event boundaries compared to the control group. However, the hierarchical structuring got better when tasks were familiar. The strengths and problems observed during event perception likewise suggest strengths and problems in understanding and executing structured activities (Bailey et al., 2013; Kurby & Zacks, 2008; Zacks et al., 2007). Therefore, I believe that investigating the mentioned event segmentation ability provides the foundation for developing appropriate assistance and training strategies.

Firstly, the (occasionally) successful detection of event boundaries can be interpreted as promising cognitive prerequisite for application of training and assistance means. It points to the existence of schemas in long-term memory of the intellectually disabled group based on their repeated experience with assembly tasks (Dulaney, 1998). Concretely, their long-term memory may contain basic knowledge about assembly activities being executed sequentially. Despite recognizing the step-by-step structure, the intellectually disabled group seems to need support in reliably detecting the distinct sequential steps, for instance, through salient highlights at event boundaries. Besides understanding of sequential structure, the intellectually disabled participants are required to perceive different hierarchical levels. I could not find selective impairments for neither fine nor coarse grain in contrast to prior reports (Zalla et al., 2013). My experiment revealed difficulties in both grains. These problems prompt appropriate ideas to communicate the goals and sub-goals of activities. One way to improve the hierarchical structuring could be by repeatedly practicing assembly tasks since familiarity with a task has the potential to increase hierarchical alignment (Zacks et al., 2001). I will enlarge upon concrete assistance and training strategies based on insights into cognitive processes in Section 7.2.

As already mentioned, I found substantial group heterogeneity in event segmentation ability which is in line with research on cognitive processes in intellectual disability (Section 6.1.2). The prevalence of interindividual differences is further in line with observations from practice where workers vary in need for individual support (Section 2.1.2). I could explain the interindividual differences found during event segmentation by differences in intelligence level. This suggests that, with the help of the event segmentation task, we can assess cognitive ability. I will further discuss this finding with respect to the usage of the event

segmentation task as ecological diagnostic method in the context of computer-based systems in Section 7.2.

The assessment of event segmentation behavior provides knowledge on strengths and weaknesses when intellectually disabled participants structure activities into meaningful events. Furthermore, it prompts means to address their problems and reinforce their strengths. However, the assessment is limited in explaining the detailed cognitive mechanisms behind impaired event segmentation. For instance, disturbed event boundary detection could be due to problems in initial stages of event segmentation, i.e., visual attention, or due to lack of integration of knowledge from long-term memory in order to establish new events. Since intellectually disabled people may show a variety of possible cognitive dysfunctions (as shown in Section 6.1.2), the detailed cognitive mechanisms behind problems in event segmentation can be only solved by further empirical investigations. For instance, prediction experiments requiring workers to guess the correct next action after a video stopped (similar to Zalla, Labruyère, Clément, & Georgieff, 2010) or reconstruction of assembly instructions (similar to Baron-Cohen, Leslie, & Frith, 1986) could provide hints on the existence of appropriate situation models.

The close connection between perception and action reported in the literature (Bailey et al., 2013) suggests that the found strengths and weaknesses in event segmentation are relevant for assembly performance, too. However, in my experiment, there was limited empirical support for this relationship. Event segmentation ability was only weakly associated with assembly performance assessed by the Lego assembly task. Specifically, the segmentation agreement measure used by Bailey et al. (2013) did not account for differences in action performance. Rather, measures of hierarchical organization, i.e., ratio of fine to coarse events as well as hierarchical alignment, showed expected tendencies to account for sequence errors during assembly. A theoretical explanation for the advantage of these hierarchical organization measures could be that, in contrast to key presses alone or to the segmentation agreement measure, they rely more on an understanding of goals and sub-goals. This understanding is also crucial for task execution. I empirically showed that there were participants who were generally good in hierarchical structuring. However, I could measure the hierarchical alignment only in a reduced sample of $N = 20$ participants which decreased the power for correlation analyses. The limitation of the segmentation agreement measure will be discussed in Section 7.1.2. Finally, a general shortcoming that was responsible for finding no substantial correlations between perception and action was the low variability among the intellectually disabled group in the Lego assembly task. The amount of provided instructional support was likely detailed enough so that the sample executed the task with a few errors. Despite the missing correlations of event segmentation measures and Lego assembly performance, the present experiment gave insights into theoretically and practically relevant cognitive processes in workers from workshops for adapted work.

7 General discussion

7.1   Theoretical contributions

In the following section, I will answer the three research questions and elaborate my main contributions with respect to the theoretical and applied research in event cognition (Section 7.1.1). Afterwards, I will reflect upon the used methodology (Section 7.1.2) and discuss how my findings affect future research plans (Section 7.1.3).

7.1.1 Major findings

In order to answer the research questions introduced in Section 3.3.1, I conducted the three experiments which were described in the previous chapters. The first experiment addressed the fact that assembly workers process their tasks mostly offline rather than during online perception. Whereas online event segmentation takes place fast and automatically during perception of an activity, offline event segmentation takes place during deliberate elaboration of a task with no time constraints. This differentiation led to the first research question:

> **R1. Are event boundaries during offline event segmentation similar to event boundaries during online event segmentation?**

I showed that event boundaries during offline event segmentation are at similar temporal locations as event boundaries during online event segmentation (Chapter 4). Hence, the principle of segmenting events is a basic process guiding both perception and offline elaboration of dynamic activities. Furthermore, the quantitative overlap I found between online and offline event boundaries demonstrates that event boundaries for offline elaboration could have corresponding event boundaries in event perception (Swallow et al., 2009). I concluded that the easy to perform classical online event segmentation task (Newtson, 1973) yields valid offline event boundaries for assembly tasks and, therefore, used it for my further experiments.

The number of chosen offline event boundaries was a combination between fine and coarse event boundaries from online event segmentation. Thus, higher-level, object-based as well as lower-level, action-based information was used for defining offline event boundaries. Furthermore, the chosen granularity was interindividually different. The variety of possible offline event boundaries may indicate different representations in working memory by the persons that created the instructions. Either their own expertise level or the expertise level of their imagined users of instructions may have influenced their working memory representations. For instance, because experts are already familiar with various details of assembly tasks, creating instructions for expert users might result in fewer offline event boundaries, as a result of leaving out details and focusing on higher-level changes. As was

shown in Experiment 2 (see the following paragraph), omitting lower-level information might be indeed a reasonable strategy for creating instructions for experienced workers. To sum up, the temporal locations of online and offline event boundaries correspond. The eventual choice of offline event boundaries is a combination of fine and coarse online event boundaries and this choice is person-dependent.

In Experiment 2, I investigated whether it is valid to make the assumption that experts know the fine level details for instruction creation. I conducted experiments to compare how experts and novices process the fine and the coarse information of assembly tasks, respectively:

R2. How does memory for events develop when repeatedly practicing the sequence of events, both in novices and experts?

Acquiring the sequence of coarse events was successful after repeated presentation but memory for the sequence of fine events did not improve. That is, the performance of participants was influenced by the interaction between training repetition and level of information, i.e., coarse verses fine events (Chapter 5). In the context of event cognition research, this study provides evidence that long-term memory for fine and coarse events differs and this difference is established after a short 3-repetitions training. Fine event boundaries involved smaller, less characteristic changes, so, consecutive fine events were more similar to each other. These similarly represented fine events were competing with each other resulting in potential memory interference (Radvansky & Zacks, 2014, p. 37).

Experts had a high performance in memorizing fine events in comparison to novices, even after only one repetition. I suggest that knowledge on these low-level actions is acquired predominantly with the help of own actions and actual experiences with the hardware. These actions resemble each other across different assembly tasks, i.e., orienting, positioning, and using screws and tools. The daily experiences could lead to a high amount of procedural knowledge that is transferrable to other assembly tasks. In contrast, object-based information, i.e., the sequence of main objects to be assembled, is very specific and has to be acquired all over again, with repetition and effort. So, the main assembly sequence represents declarative knowledge that is initially present in neither experts nor novices. Only with training, can both expert and novice participants improve memory for these coarse events. Trainees acquired these events with priority, so, initial training for novices and experts should support this cognitive process and highlight the main sequence information. Since experts already showed a high performance in fine events, it could be argued that they should get a more concise presentation of the fine events rather than going through each step in detail.

In accordance with findings on improved hierarchical alignment when familiar with a task, I interpret the improved prediction performance for coarse events as a temporally closer relation between the nearest fine and the coarse event boundary. The concept of hierarchical alignment does not specify if closeness refers to the direction from the fine to the next coarse or from the coarse to the next fine event boundary. In my experiment, I could show that the

former direction applies, i.e., from nearest fine to coarse. Because I stopped videos either shortly before or shortly after a coarse event boundary and tested prediction performance of participants, I could investigate both temporal directions. If the connection was strengthened between the nearest fine and the coarse event boundary (to illustrate, see example in Figure 7: from end of "screwing" to the next coarse event boundary), than performance in predicting coarse should improve more clearly than performance in predicting fine. If the connection was strengthened between the coarse and the upcoming fine event boundary (example in Figure 7: from initial coarse event boundary to end of "positioning"), than performance in predicting fine should improve more clearly than performance in predicting coarse. I empirically confirmed that the direction of temporal closeness in long-term memory that is established after initial training is the one from fine event to coarse event, which is in accordance with the enclosure concept formulated for event perception.

As event segmentation is based on multiple cognitive processes, I further extended the research mentioned above to investigate cognitive ability of a group of intellectually disabled participants. I investigated the influence of cognitive abilities and daily experiences on event segmentation more deeply in the third experiment. Persons working in workshops for adapted work are familiar with assembly tasks, too. However, their daily tasks are mostly very simple and monotonous. This led to the following research question:

R3. Do the simple and repetitive assembly tasks offered at workshops for adapted work utilize the full cognitive potential of intellectually disabled persons?

First of all, I found a substantial group heterogeneity in event segmentation ability. This high amount of interindividual differences during event segmentation confirms that the intellectually disabled people vary concerning their cognitive (dys-)functions (Section 6.1.2). Having said that, the results of the experiment suggest that, in general, they are capable of performing more complex and interesting tasks than they have been executing so far. The first evidence is their good performance in a complex task with Lego bricks with prior detailed instructional support. Second, they showed significant agreement with the control group with respect to locations of event boundaries. Third, ability to hierarchically structure dynamic activities was better for assembly tasks compared to an unfamiliar yet common activity, namely, a breakfast making task.

Further analyses of the event segmentation data provided detailed information on strengths and weaknesses during event perception. Because event segmentation is a basic cognitive function important for action understanding (see Experiment 1), memory (Sargent et al., 2013), and action execution (Bailey et al., 2013), I suggest interpreting the empirical findings on event segmentation with respect to the workers' cognitive potential to actually perform and their potential need for support. For instance, the (occasionally) successful detection of event boundaries at the right temporal location can be interpreted as a promising cognitive prerequisite. It points to a basic understanding of the sequential nature of tasks, potentially due to existing knowledge from long-term memory (Dulaney, 1998). Despite this capability, they also showed a number of difficulties.

First, they repeatedly neglected the end of an event during the event segmentation task indicating a disturbed updating process. This weakness prompts instructional systems that draw their attention towards the end of old events and beginning of new events. Second, the intellectually disabled group had problems with the conception of hierarchical organization of dynamic activities evident either in a low ratio of fine to coarse events or in diminished hierarchical alignment and enclosure values. These problems suggest appropriate means to communicate the goals and sub-goals within activities. However, I could show that structuring was flexible to some extent, i.e., structuring was better in assembly tasks compared to the unfamiliar video. In sum, the strengths and problems observed during event perception likewise suggest strengths and problems in understanding and executing structured activities because of the close connection of perception, memory, and action (Bailey et al., 2013; Kurby & Zacks, 2008; Zacks et al., 2007). I will go into different assistance and training strategies addressing the need for support in Section 7.2.

A final remark addresses the just mentioned improved hierarchical alignment of intellectually disabled participants in the assembly tasks. An alternative explanation to the familiarity effect, is the fact that assembly tasks *per se* are highly structured in comparison to the task shown in the breakfast video. Supporting evidence comes from relatively high alignment and enclosure values throughout this thesis. On the one hand, the assembly task from Experiment 2 resulted in an enclosure value of .81; on the other hand, the assembly tasks from Experiment 3 led to observed distance values around 1 s for the control group. In contrast to values reported in the literature, i.e., enclosure values around .40 to .67 (Hard et al., 2006) and observed distances of 1.7 to 2.8 s (Swallow et al., 2009; Zacks et al., 2001), the used assembly tasks may prompt especially high temporal closeness and hierarchical enclosure compared to other activities. Thus, assembly tasks seem to foster event segmentation with a clear chunking pattern. This would explain why the understanding of goals and sub-goals is improved in highly structured tasks for intellectually disabled people. Therefore, I conclude that assembly tasks are a good choice of potentially manageable tasks for intellectually disabled workers in workshops for adapted work.

Besides the answers to my specific research questions, I will summarize my general contributions to the area of event cognition research in the following.

The Event Segmentation Theory is an ecologically valid and comprehensive framework

Throughout this thesis, I demonstrated that the Event Segmentation Theory offers a comprehensive and ecologically valid framework for the investigation of cognitive processes in the applied field of assembly work. I demonstrated the range of applications of the event segmentation task for investigating cognitive processes during perception, understanding, and practicing assembly tasks. In this section, I will review the empirical arguments.

First, the output of the online event segmentation task is correlated to the event boundaries from the offline event segmentation. Hence, I confirmed that this task is an ecologically valid

way to assess time points that are important during offline understanding of assembly tasks. In my experiment, I showed that there is a close connection between online event segmentation and the paradigm of instruction creation. This, in turn, confirms the practical relevance of the event segmentation task for investigating both perception and understanding of assembly tasks. With the help of the IBES tool that was developed within this thesis, the event boundaries can be easily utilized for creating instructions that are structured around these important strategic points.

Second, the output of the event segmentation task, i.e., the hierarchical organization of object-based coarse events and action-based fine events, provided a useful framework to test memory within a prediction paradigm. Concretely, I stopped a video around coarse event boundaries and asked participants to predict next actions in order to test memory on the assembly sequence based on similar recommendations by Ericsson (2008). With the help of this memory test, I could test declarative knowledge for information differing in level of hierarchy.

Third, the event segmentation task allowed the investigation of assembly work-related cognitive processes in the intellectually disabled group. Based on these insights concrete assistance means can be derived. Furthermore, performance in the task was able to account for differences in general cognitive ability. Based on my empirical findings using a variety of available event segmentation measures, I will recommend two of them as potential diagnostic measures (see Section 7.1.2). The initial experience with this task in the context of special workshops confirmed its ecological validity. First, most intellectually disabled participants were able to participate in the event segmentation task. Second, watching a video had generally positive effects on their assembly performance (see the video instruction before the Lego assembly task). In Section 7.1.3, I will introduce the approach of an event segmentation training in order to establish the utilization of this assessment method. Using an event segmentation based measure is practically useful because it addresses the lack of appropriate assessment in workshops for adapted work (EASPD, 2012b) and the potential to be integrated into computer-based systems (Section 7.2)(Gorecky et al., 2012; Korn et al., 2013b).

The role of long-term memory in event segmentation

In the context of the Event Segmentation Theory, long-term memory comes into play twofold (see the red arrow in Figure 34 and refer also to Section 3.1.2). First, in the context of knowledge acquisition, experiences pass from working memory into episodic long-term memory encoded as respective situation models. In Experiment 2, I investigated the transfer of sequential events from working memory to long-term memory and showed that there are memory differences between consecutive fine event versus consecutive coarse event representations after repeated presentation. I demonstrated that, when repeatedly going through assembly tasks, people memorize the main sequence of coarse events instead of the sequence of details. Hence, it could be that situation models for coarse events are

established with priority compared to situation models for fine events, and that this advantage of coarse events is amplified with repetition. This finding suggests a concretization of the Deliberate Practice Framework (Ericsson et al., 1993). The framework's claim that repetition is beneficial for learning initially only holds for declarative knowledge about the sequence of higher-level, coarse events.

Second, the Event Segmentation Theory stresses the importance of long-term memory when *perceiving* dynamic information (Section 3.1.1 and Section 3.1.2). Knowledge from long-term memory influences the working memory processing of new experiences. In Experiment 1, participants were likely guided by conceptual knowledge when thinking about events offline because they defined event boundaries that resembled online coarse event boundaries which are guided by top-down processes from memory (Zacks et al., 2009). They also used finely represented information but not all of these fine events from perception were used when thinking about events offline.

The empirical results provide an indication which lower-level changes are not processed as important strategic points for instruction creation. Imagine four consecutive fine events depicting screwing first, second, third, and fourth screw consecutively. Whereas some participants perceived four fine event boundaries for each screwing event during online perception, the offline event segmentation led to a summarized representation of these four events. The number of changes between these four events was only one, respectively, i.e., "next screw", and this was apparently not enough to justify a new offline event boundary. Based on prior studies (Huff et al., 2014), I postulate a quantitative relationship for offline event segmentation. Only increasing the number of changes to greater than one, e.g., the appearance of a new screwdriver between the 2nd and 3rd screw, would result in the definition of an offline event boundary.
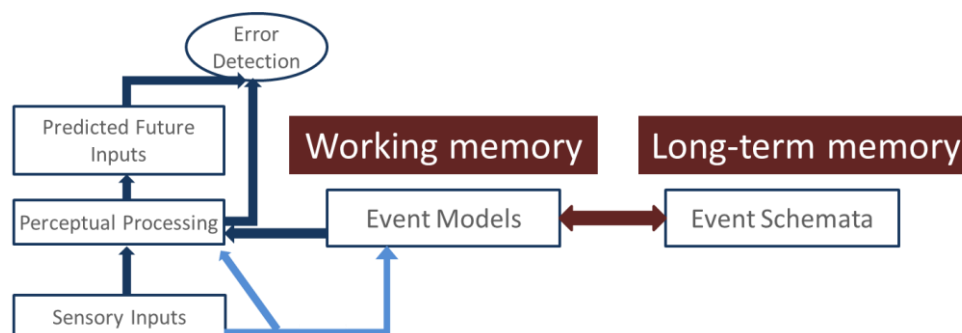
Figure 34: Bi-directional exchange between working memory and long-term memory during event segmentation.

Experiment 2 provided further insights into the influence of long-term memory on working memory processes. Despite extensive prior knowledge of experts in automotive manufacturing, the sequence of object-based coarse events has to be learned with the same deliberate practice as in novices. In Experiment 3, intellectually disabled people segmented with a higher hierarchical alignment in the familiar tasks. I already discussed that this might be due to the structured nature of assembly tasks. Nevertheless, it provides evidence that there is fundamental knowledge about activities being executed in a sequential manner

which is in line with the claim of existent schematic knowledge in long-term memory (Dulaney, 1998; Zalla et al., 2013).

7.1.2 Methodological considerations

Furthermore, I pursued two methodological goals within this thesis introduced in Section 3.3.2. First, there was a lack of assessment methods for offline event segmentation and their respective event boundaries. This led to the following aim:

M1. Developing a tool for assessing offline event segmentation by using an instruction creation paradigm

I presented the IBES tool in Chapter 4 which is the first software tool that makes it possible to create instructions semi-automatically based on a video of a task. The resulting instructions rely on event boundaries important for understanding and learning. The input the IBES tool requires is a sequence of static frames of a task's video. Its output is a ready-to-use instruction manual containing text and pictures. In order to create this manual, participants execute four steps. First, they define an appropriate structure for the task. Second, they choose those static frames from the video that are most illustrative. Third, they add textual descriptions. Fourth, the manual can be printed and added by manual overlays, if necessary. In the following, I will discuss the potential usage of the IBES tool in instructional design, cognitive psychology, and practical applications.

Researchers interested in instructional design may use this tool to analyze desirable characteristics of instruction manuals by letting actual users create them. This may support the development of suitable Augmented Reality instructions (Bleser et al., 2015). The IBES tool can be used along the lines of "turning users into designers" (Daniel & Tversky, 2012, p. 303). Furthermore, the easy creation of manuals within the IBES tool promotes manipulation of manuals differing in structure. For instance, the structure of the manual could be either fine- or coarse-grained, or even completely violate the human event structure. Furthermore, the amount of textual and graphical content could be varied. In sum, researchers in instructional design have a tool to investigate both the *creators* and the *users* of instructions.

For instance, it could be argued that experts are able to structure activities in different granularities depending on specific aims. They can create instructions focusing on the superordinate relations or instructions incorporating details that may be especially important for novices. In a further evaluation step, their different versions of instructions could be used by experts and novices, respectively, in order to validate if novices actually benefit from fine-grained training while experts prefer a coarse-grained training.

Research in psychology may use the IBES tool to further explore situation models in long-term memory (Section 7.1.3). It may also support research on theory of mind. For instance, Killingsworth, Saylor, and Levin (2005) were interested if their participants would create different instructions given that they made them either for computers or for humans. Hence, they showed that participants defined more segments for computers because they attributed

limited reasoning capabilities to them. An important contribution of this newly created IBES tool for research is that it supports log files with which investigators can analyze the offline segmentation or the instructional design process as a whole.

For practical applications, there is generally little software support targeted to instruction generation. Here, the IBES tool can fill a practical gap. Currently, producing efficient instruction manuals requires an effortful, labor-intensive process involving creation of meaningful structure for the assembly steps and the choice of appropriate media. Engineers and trainers typically use existing data from the engineering process, e.g., graphical product models and planned production sequence data from CAD software and import this information to word processing or image editing programs, to provide additional, manually edited descriptions and graphics. Technical writers and editors use these documents as a starting point, and may exploit more sophisticated and expensive desktop publishing tools for more powerful functionalities for graphic design and media creation. In contrast to this complex process, the IBES tool only requires a video of the actual assembly task and a computer to run the software. It supports the production of ready-to-use manuals that incorporate multimedia and are meaningfully structured according to event boundaries.

The second methodological goal addressed the lack of an extensive overview of event segmentation measures that are suited to analyze the event segmentation ability in intellectually disabled people. Therefore, I aimed at:

M2. Evaluating and refining existing event segmentation measures with respect to their suitability for intellectually disabled persons

In this thesis, I applied a range of event segmentation measures introduced in Section 3.1.4 to the intellectually disabled group in order to extensively describe their event segmentation behavior as well as to evaluate the feasibility and validity of the available measures. My main methodological contribution was the refinement of the hierarchical alignment measure. The new method is more suitable for the intellectually disabled group, it is theoretically sounder, and provides a more convenient way of computation and interpretation. It will be discussed below.

The basic measure, number of key presses, in both fine and coarse segmentation was an initial hint whether intellectually disabled participants can follow instructions and can segment according to instructions. However, the number of key presses is greatly influenced by video length. Furthermore, interindividual differences in the number of key presses are high, even within the control group. Therefore, this number is no reliable indication for event segmentation performance. In contrast, the ratio of key presses in fine and coarse segmentation provides a better way to interpret the event segmentation ability. This method incorporates the relation between key presses in fine and coarse condition in one measure. Hence, it provides an initial assessment of hierarchical perception. Together with the existence of a comprehensive reference value of 3 (based on prior empirical findings), this measure allows easy interpretation.

Segmentation agreement was intuitively a promising measure because it captures the individual overlap with a control group that consists of presumably good "segmenters" (represented by the group of students in this thesis). This measure was used successfully in prior studies to analyze individual ability to perform the event segmentation task and relate this ability to action performance (e.g., Bailey et al., 2013). In contrast, in this thesis, segmentation agreement could neither account for differences in assembly performance nor explain differences in IQ. As noted in prior work, the segmentation agreement calculation is influenced by number of key presses (Kurby & Zacks, 2011). Frequent key presses go along with an increased chance to "hit" an event boundary perceived by the control group. Consequently, persons in the intellectually disabled group who pressed frequently but rather randomly could achieve similar values compared to persons who pressed not frequently but at time points actually corresponding to the controls. Even if the segmentation agreement calculation has been scaled according to prior work (Kurby & Zacks, 2011), this adaptation was insufficient to correct for the key presses in this sample.

The classical hierarchical alignment model (Zacks et al., 2001) assumes a linear relation between distance and alignment. The investigation in my experiment revealed this assumption as problematic in the sample of the intellectually disabled participants. Concretely, it could be shown that if nearest fine and coarse event boundary are far away from each other, the null model is insufficient to correct for this extensive distance. Rather, the calculation model had to be adjusted. I proposed a hierarchical alignment with an exponential function calculation abbreviated as eCALC. It contains the assumption that the further away a nearest fine event boundary is from its coarse event boundary, the less likely it is related to it. In other words, with increasing distance between coarse and nearest fine event boundary temporal closeness exponentially decays to zero. The positive correlation with the classical calculation supported its validity for both the intellectually disabled group and the control group. Therefore, this method should be further exploited in future studies.

In addition, the eCALC method is more convenient because it does not require separate calculation of the *expected* distances and their consecutive subtraction from *observed* distances. Furthermore, eCALC goes along with standardized values ranging from 0 to 1 that support comparability across different stimulus material and different empirical investigations. In the context of intellectual disability, a shortcoming of the hierarchical alignment computation was that a suitable calculation requires the difference between key presses in fine minus coarse condition to be greater than 0. If the requirement of right conception of coarse and fine granularity is met, it is a method that provides a detailed insight into the ability to hierarchically structure dynamic information. To sum up, the hierarchical alignment computation with eCALC provides a theoretically sound and convenient way to investigate hierarchical alignment. In contrast to the classical method, it is applicable for people with intellectual disability as well.

Finally, I review the determination of significant event boundaries based on treating time continuously. Based on my initial experiences with the R package segmag (Papenmeier,

2014), I would generally recommend a sample size of at least 20 in order to assess valid event boundaries in typical samples. As can be seen from the event segmentation data in Experiment 2, a sample size of 10 is adequate to find characteristic peaks. However, they do not exceed critical cutoffs with conventional significance levels of 95% or 99%. Furthermore, in the context of using the difference method it has to be noted that the computation is highly sensitive for group differences in key presses. Therefore, the initial prerequisite is that group sizes should approximately equal. However, significant differences may appear even if graphical inspection of separate segmentation plots shows significant event boundaries for both groups. Therefore, a more strict confidence level and additional graphical evaluation are recommended when using the difference method.

7.1.3 Future directions

The offline event segmentation, as it was performed in Experiment 1, provided a way to assess working memory representations during offline elaboration of dynamic activities. The chosen paradigm of offline elaboration was instruction creation. The IBES tool could be exploited and further developed for other areas of offline event segmentation. For instance, assessment of long-term memory representations of events, i.e., situation models, is still an area of open empirical questions. The IBES tool's instruction paradigm could be adapted towards a memory assessment tool in order to use it for free recall or recognition studies. Consequently, the video cues would have to be omitted. In retrospect, participants would have to segment a black video stream by using the IBES tool, with a time line as orientation. Alternatively, a picture frame could be shown to participants. Their task would be to localize the right time point where the action happened by using again the IBES tool. Furthermore, the software tool allows assessment of verbal descriptions, that is, collection of qualitative free recall data. Thus, the IBES tool offers ways to combine traditional memory paradigms with offline event segmentation. These different memory measures can be used to disentangle how schemata influence long-term memory processes (Brewer, 1981).

It would be also interesting to further investigate the existence and nature of situation models in intellectually disabled participants. Experiments using prediction paradigms (e.g., Huff et al., 2014; Zacks et al., 2011) could be used to investigate conceptual knowledge. In the context of workers from workshops for adapted work, intellectually disabled participants could be asked to watch videos depicting assembly tasks that suddenly stop. Afterwards, they have to predict the correct next action, as was previously done in an autistic sample (Zalla et al., 2010). Similar to the study by Zalla et al. (2010), distractor frames could be varied depicting more or less likely next events or showing the right versus false temporal order. This would enable a detailed understanding of potential conceptual problems of intellectually disabled people.

Despite the close connection between perception and action (Bailey et al., 2013), I found only weak correlations. The potential methodological reasons with respect to the instructed Lego assembly task have been discussed in the experimental discussion in Section 6.4. In

order to further elaborate the association between perception and action, the reverse hypothesis could be postulated and tested. If the assumption that action and perception are based on the same basic ability to structure activities holds, the repeated execution of correct event segmentation trials should beneficially affect action performance. Concretely, an event segmentation training that incorporates cues for event boundaries could be introduced into the workshops for adapted work. Afterwards, assembly performance tests could reveal if the event segmentation training resulted in action improvement. Furthermore, the event segmentation training design could be varied concerning different variables in order to find most appropriate means. For instance, highlighting fine and coarse event boundaries differently, varying the exact timing of cues, choosing different ways of communicating the cues involving graphical highlights, auditory signals, or language cues, and so on. Furthermore, cues may be unspecific ("there is an event boundary") or specific ("there is an event boundary: object X"). In sum, I believe that scientific investigations on the link between action and perception are theoretically relevant and practically promising with respect to improving quality of work and life.

7.2   Practical contributions

In Chapter 2, I have introduced the applied field of research which is the human-machine interaction in computer-based assistance and training systems for assembly, specifically, instructional support in automotive industry and workshops for adapted work. I argued that these systems have to be adaptive to the individual user's need, experience, and cognitive potential. In order to be adaptive, two important questions have to be answered. In the present chapter, I will revisit the questions from Section 2.2.3 and present the practical implications that I derive from my empirical findings.

1. **What are relevant work-related user characteristics and how to integrate their assessment into a technical system? (User assessment)**

The empirical evidence from this thesis is that general cognitive ability and expertise influence processes from event cognition. In Experiment 3, I used the classical event segmentation task to assess cognitive potential in intellectually disabled workers. I argue that assessment based on the event segmentation task is an alternative to other diagnostic measures. I already mentioned the disadvantages of available tests with respect to their integration into computer-based systems (Section 2.2.3). Some of them depend on real world interactions. Others need detailed explanation and consistent presence of an experimenter, they incorporate artificial, ecologically invalid tasks, or their acceptance by users is low. To the contrary, the event segmentation task presented within this thesis can be easily integrated into computer-based systems and it addresses the mentioned limitations.

First, the technical requirements to integrate the task are reasonable. The diagnostic material consists of videos depicting assembly tasks that can be shown within the computer-based system, e.g., on the monitor (compare the assistance system in Figure 5). An input device like mouse or keyboard has to be added. The analysis of event segmentation data can be

supported by the system according to pre-defined calculations. The ratio of key presses during fine versus coarse event segmentation was promising in order to explain differences in performance and general cognitive ability. Second, need for the support of supervisors is limited. In order to execute the event segmentation task, written explanations are usually used. In the group of intellectually disabled people an additional personal introduction is required. However, with repeated execution, users could cope with the event segmentation task independently. The task requires understanding the action of key pressing and the concepts of fine versus coarse grain so that it is accessible to a wide range of users. A respective event segmentation training suggestion has been sketched above. Third, the event segmentation task contains no artificial material but videos that depict the actual work content. This is important for its ecological validity and the acceptance by the users. In addition, watching the videos constitutes an opportunity to increase familiarity with assembly tasks and to learn. To sum up, the event segmentation task suggests being a valid, convenient, computer-based diagnostic assessment method to investigate assembly work-related cognitive ability in workshops for adapted work.

Informed by the results from the event segmentation task, the system can estimate the need for cognitive support. Based on that, the following question was to elaborate:

2. How should tailored instructions be designed? (Adaptation guidelines)

As a basic guideline, the system's instructions should match the overall human event understanding, i.e., the system's segmentation of assembly steps should go along with human event boundaries (for an appropriate computer algorithm refer to Petersen & Stricker, 2012). The Event Segmentation Theory gives further hints on adaptive assistance and training means. Concretely, the instructional support may address different cognitive processes involved when making sense of assembly tasks and also involved when performing assembly tasks. The following list illustrates a sample of adaptation guidelines:

- Visual attention: Assistance systems should provide attentional guidance in assembly tasks by salient spatial cues at the relevant position (Stork & Schubö, 2010). This thesis adds the suggestion for cues at event boundaries, especially, for intellectually disabled workers. These cues would enhance the classical event segmentation task and constitute an event segmentation training. Furthermore, salient cues should also highlight the end of an old and the beginning of a new assembly step in both instructional videos and real-time step-by-step guidance. There is a wide range of cues possible including verbal and non-verbal ones. In the context of event segmentation processes, the cues could contain, for instance, information about the level of change, i.e., fine or coarse event boundaries.

- Working memory: According to the Event Segmentation Theory, event models can be processed at different grains simultaneously. In order to support the correct hierarchical representation of object-based coarse events and their respective action-based fine events, instructions could depict the respective higher-level event for each action (similar to the concept of Zacks & Tversky, 2003). Furthermore, training

systems with gesture-based interaction (similar to Figure 3) could make the grouping of fine events clearer by requiring users to execute only one gesture to trigger all fine events instead of triggering each fine event separately. Particularly, experts should be able to skip a detailed training mode in favor of a concise version containing only the main sequence of objects.

- Long-term memory: An assistance system could store a sample of similar "situation models" that have been executed before. In order to support the establishment of a new event model in the context of a new assembly task, the instructional system could then present this "prior knowledge". An assistance system could furthermore incorporate learning performance tests in order to derive the state of knowledge and, accordingly, reduce or adapt the support (see the concept of fading in Eiriksdottir & Catrambone, 2011). Furthermore, incorporating testing elements per se can have positive effects on learning (Roediger & Karpicke, 2006). For instance, Hegarty, Kriz, and Cate (2003) showed that the challenge to predict the behavior of a mechanical system increased the participants' understanding compared to only passive presentation of the mechanical system.

- Visual perception and anticipation: The just mentioned concept of prediction can be also helpful for training the loop process between perceptions, predictions, and error detection (see the Event Segmentation Theory in Figure 6). Animations of what is coming next (Hegarty et al., 2003) or requests to predict the next action could be useful for intellectually disabled persons to foster their anticipatory perception and behavior.

Based on cognitive processes important in event segmentation, existing literature in instructional design, and my empirical findings, I presented an overview of selected design means like cueing event boundaries, structuring events, and challenging the users. Hence, depending on the users' cognitive processes and needs, several design adaptations – inspired by the underlying event structure of assembly tasks – are on hand. They need to be evaluated in future work.

The future technological outlook is a human-machine interaction with advanced systems that are more and more able to anticipate, learn, adapt, and collaborate with humans instead of only passively following the users' control (Evenson et al., 2010), for instance, in advanced human-robot collaboration. This leads to systems that intrude into our stream of information and action. Radvansky and Zacks (2014) noted that segmenting becomes more demanding in *interactive* events compared to passive observation of events. However, the detailed underlying cognitive processes and potential practical implications are open for empirical investigations. In sum, I believe that the Event Segmentation Theory will remain a relevant framework for scientifically solving practical challenges in the context of advanced human-machine interaction.

7.3   Conclusion

My aim was to contribute to research on event segmentation processes with an applied focus on assembly workplaces. In Experiment 1, I demonstrated quantitatively that event segmentation processes not only come into effect when online observing activities but also when offline thinking about the activities. Thus, when people deliberately process assembly tasks, they are guided by similar event boundaries compared to event segmentation during online perception. In Experiment 2, I focused on the output of the event segmentation processes, i.e., perceiving assembly tasks as hierarchical sequence of object-based coarse and action-based fine events. The experiment showed that repeatedly practicing such an assembly task leads to an advantageous acquisition of the sequence of coarse events and no learning benefit for fine events. Experts outperformed novices in memory for action-based fine events. Thus, when workers practice assembly tasks, their learning performance depends upon the hierarchical structure of the task and prior experiences. In Experiment 3, I investigated the event segmentation processes, for the first time, in a group of intellectually disabled employees from the assembly workshops. Event segmentation data correlated with general cognitive ability. I suggest interpreting the empirical findings on strengths and weaknesses in event segmentation with respect to the intellectually disabled group's cognitive potential to perform more interesting tasks than the current repetitive ones. Hence, understanding the event segmentation processes in intellectually disabled people can prompt appropriate assistance means at the workplace and improve their quality of life.

I suggest that assembly workplaces can benefit from these empirical findings with respect to the development of user-adaptive computer-based assistance and training systems. First, the event segmentation task can be integrated as diagnostic user assessment of the individual need for support. Second, the empirical observations prompt adaptation guidelines regarding presentation of the structure of events and emphasis on event boundaries.

I believe that the practical usefulness of research in cognitive psychology, especially in event cognition, also applies for other domains of human-machine interaction. More and more context-sensitive and "intelligent" systems will step in our daily lives by providing assistance and support in an increasingly autonomous way, for instance, advanced driver assistance, human-robot collaboration, or home automation systems. They can only provide suitable information, suggestions, and automated actions, if they incorporate general knowledge about the nature of human activities and if they tailor their assistance to the users' individual perception and understanding of interactive events.

8 References

Adamczyk, P. D., & Bailey, B. P. (2004). If not now, when?: The effects of interruption at different moments within task execution. *Proceedings of the Conference on Human Factors in Computing Systems (CHI '04)*, New York, USA, 271–278.

Ainsworth, S., & VanLabeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction*, *14*(3), 241–255.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: APA.

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, *6*(4), 355–385.

Baddeley, A., & Jarrold, C. (2007). Working memory and Down syndrome. *Journal of Intellectual Disability Research*, *51*, 925–931.

Bailey, H. R., Kurby, C. A., Giovannetti, T., & Zacks, J. M. (2013). Action perception predicts action performance. *Neuropsychologia*, *51*(11), 2294–2304.

Baron-Cohen, S., Leslie, A., & Frith, U. (1986). Mechanical, behavioral and intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology*, *4*, 113–125.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-8.

Biocca, F., Tang, A., Owen, C., & Xiao, F. (2006). Attention funnel: Omnidirectional 3D cursor for mobile Augmented Reality platforms. *Proceedings of the Conference on Human Factors in Computing Systems (CHI '06)*, Montreal, Canada, 1115–1122.

Black, J. B., & Bower, G. H. (1979). Episodes as chunks in narrative memory. *Journal of Verbal Learning and Verbal Behavior*, *18*(3), 309–318.

Bleser, G., Damen, D., Behera, A., Mura, K., Miezal, M., Gee, A., … Stricker, D. (2015). Cognitive learning, monitoring and assistance of industrial workflows using egocentric sensor networks. *PLoS ONE*, *10*(6), 1–41.

Boltz, M. (1992). Filmed narratives. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 90–105.

Boot, F., Pel, J., Vermaak, M., van der Steen, J., & Evenhhuis, H. (2013). Delayed visual orienting responses in children with developmental and/or intellectual disabilities. *Journal of Intellectual Disability Research*, *57*(12), 1093–1103.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, *3*, 193–209.

Brewer, W. F. (1981). Role of Schemata. *Cognitive Psychology*, *13*, 207–230.

Brough, J. E., Schwartz, M., Gupta, S. K., Anand, D. K., Kavetsky, R., & Pettersen, R. (2007). Towards the development of a virtual environment-based training system for mechanical assembly operations. *Virtual Reality*, *11*(4), 189–206.

Carretti, B., Belacchi, C., & Cornoldi, C. (2010). Difficulties in working memory updating in individuals with intellectual disability. *Journal of Intellectual Disability Research*, *54*(4), 337–345.

Charness, N., Kelley, C. L., Bosman, E. A., & Mottram, M. (2001). Word-processing training and retraining : Effects of adult age, experience, and interface. *Psychology and Aging*, *16*, 110–127.

Charness, N., Tuffiash, M., Krampe, R., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology*, *19*(2), 151–165.

Cleveland, W. (1981). LOWESS: A program for smoothing scatterplots by robust locally weigthed regression. *The American Statistician*, *35*, 54.

Cornoldi, C., & Giofrè, D. (2014). The crucial role of working memory in intellectual functioning. *European Psychologist*, *19*(4), 260–268.

Crajé, C., Aarts, P., Nijhuis-van der Sanden, M., & Steenbergen, B. (2010). Action planning in typically and atypically developing children (unilateral cerebral palsy). *Research in Developmental Disabilities*, *31*, 1039–1046.

Crane, L., & Goddard, L. (2008). Episodic and semantic autobiographical memory in adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(3), 498–506.

Daniel, M.-P., & Tversky, B. (2012). How to put things together. *Cognitive Processing*, *13*(4), 303–19.

De Campos, A. C., Cerra, L. C., Silva, F. P., & Rocha, N. A. (2014). Bimanual coordination in typical and atypical infants: Movement initiation, object touching and grasping. *Research in Developmental Disabilities*, *35*(10), 2416–2422.

Dulaney, C. L. (1998). Automatic processing: Implications for job training of individuals with mental retardation. *Journal of Developmental and Physical Disabilities*, *10*(2), 175–184.

Eiriksdottir, E., & Catrambone, R. (2011). Procedural instructions, principles, and examples: How to structure instructions for procedural tasks to enhance performance, learning, and transfer. *Human Factors*, *53*(6), 749–770.

Endsley, M. (2013). Situation awareness. In J. Lee & A. Kirlik (Eds.), *The Oxford Handbook of Cognitive Engineering* (pp. 88–124). Oxford: Oxford University Press.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, *79*(10), 70–81.

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine*, *15*, 988–994.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The Role of Deliberate Practice in the Acquisition of Expert Performance. *Psychological Review*, *100*(3), 363–406.

European Association of Service Providers for Persons with Disabilities. (2012a). *Sheltered workshops in the EU*. Retrieved from http://www.easpd.eu/sites/default/files/sites/default/files/factsheet_sheltered_workshops _in_the_eu.pdf

European Association of Service Providers for Persons with Disabilities. (2012b). *The spectrum of work and employment*. Retrieved from http://www.easpd.eu/sites/default/files/sites/default/files/factsheet_spectrum_of_work_a nd_employment.pdf

Evenson, S., Rheinfrank, J., & Dubberly, H. (2010). Ability-centered design: From static to adaptive worlds. *Interactions*, *17*(6), 75–79.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London: SAGE Publications Ltd.

Fleishman, E. A. (1972). On the relation between abilities, learning, and human performance. *American Psychologist*, *27*(11), 1017–1032.

Formann, A., Waldherr, K., & Piswanger, K. (2011). *Wiener Matrizen-Test 2. Manual*. Göttingen: Beltz Test GmbH.

Gagne, R. M., & Dick, W. (1983). Instructional psychology. *Annual Reviews of Psychology*, *34*, 261–295.

Gernsbacher, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology*, *17*, 324–363.

Gorecky, D., Campos, R., & Meixner, G. (2012). Seamless Augmented Reality support on the shopfloor based on cyber-physical-systems. *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '12)*, San Francisco, USA.

Gorecky, D., Mura, K., von Falkenhausen, I., Apold, J., & Arlt, F. (2013). Spielebasiertes Training gestalten und integrieren. a*tp edition - Automatisierungstechnische Praxis*, *55*(5), 888–895.

Goto, M., Uematsu, Y., & Saito, H. (2010). Task support system by displaying instructional video onto AR workspace. *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR '10)*, Seoul, South Corea, 83–90.

Graziano, W. G., Moore, J. S., & Collins, J. E. (1988). Social cognition as segmentation of the stream of behavior. *Developmental Psychology*, *24*(4), 43–50.

Hard, B. M., Lozano, S. C., & Tversky, B. (2006). Hierarchical encoding of behavior: translating perception into action. *Journal of Experimental Psychology: General*, *135*(4), 588–608.

Hard, B. M., Recchia, G., & Tversky, B. (2011). The shape of action. *Journal of Experimental Psychology: General*, *140*(4), 586–604.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition and Instruction*, *21*(4), 209–249.

Henderson, S., & Feiner, S. (2011). Exploring the benefits of Augmented Reality documentation for maintenance and repair. *IEEE Transactions on Visualization and Computer Graphics*, *17*(10), 1355–1368.

Henry, L. (2001). How does the severity of a learning disability affect working memory performance? *Memory*, *9*(4-6), 233–247.

Hermawati, S., Lawson, G., D'Cruz, M., Arlt, F., Apold, J., Andersson, L., … Malmsköld, L. (2015). Understanding the complex needs of automotive training at final assembly lines. *Applied Ergonomics*, *46*(Part A), 144–157.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, *8*(11), 494–500.

Huff, M., Meitz, T. G. K., & Papenmeier, F. (2014). Changes in situation models modulate processes of event perception in audiovisual narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1377–1388.

Huff, M., Papenmeier, F., & Zacks, J. M. (2012). Visual target detection is impaired at event boundaries. *Visual Cognition*, *20*(7), 848–864.

Iarocci, G., & Burack, J. (1998). Understanding the development of attention in persons with mental retardation: Challenging the myths. In J. Burack, R. Hodapp, & E. Zigler (Eds.), *Handbook of mental retardation and development* (pp. 349–381). Cambridge: Cambridge University Press.

Jacoby, L. L., Jones, T. C., & Dolan, P. O. (1998). Two effects of repetition: Support for a dual-process model of know judgments and exclusion errors. *Psychonomic Bulletin & Review*, *5*(4), 705–709.

Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction*, *20*(2), 146–154.

Jipp, M., Wagner, A., & Badreddin, E. (2008). Individual ability-based system design of dependable human-technology interaction. *Proceedings of the World Congress International Federation of Automatic Control (IFAC'08)*, Seoul, South Corea, 14779–14784.

Johnson-Laird, P., & Stevenson, R. (1970). Memory for syntax. *Nature*, *227*(412).

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*(1), 23–31.

Killingsworth, S., Saylor, M. M., & Levin, D. T. (2005). Segmenting action for computers and humans: Possible links to intentional understanding. *Proceedings of the International Workshop on Robot and Human Interactive Communication (ROMAN '05)*, Nashville, USA, 196–201.

Korn, O., Schmidt, A., & Hörz, T. (2013a). Augmented manufacturing: A study with impaired persons on assistive systems using in-situ projection. *Proceedings of the International Conference on Pervasive Technologies Related to Assistive Environments (PETRA '13)*, Island of Rhodes, Greece, 1–8.

Korn, O., Schmidt, A., & Hörz, T. (2013b). The potentials of in-situ-projection for augmented workplaces in production. A study with impaired persons. *Proceedings of the Conference on Human Factors in Computing Systems (CHI '13)*, Paris, France, 979–984.

Krampe, R., & Ericsson, K. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General*, *125*(4), 331–359.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*(2), 72–9.

Kurby, C. A., & Zacks, J. M. (2011). Age differences in the perception of hierarchical structure in events. *Memory & Cognition*, *39*, 75–91.

Lanfranchi, S., Baddeley, A., Gathercole, S., & Vianello, R. (2012). Working memory in Down syndrome: Is there a dual task deficit? *Journal of Intellectual Disability Research*, *56*, 157–166.

Lassiter, G. D., & Slaw, R. D. (1991). The unitization and memory of events. *Journal of Experimental Psychology: General*, *120*(1), 80–82.

Leven, A. (2007). *Postponed plans - Prospective memory and intellectual disability* (Doctoral dissertation). Retrieved from DiVA institutional repository for theses. (diva2:17136)

Lienert, G. (1976). *Drahtbiegeprobe (DBP)*. Göttingen: Hogrefe.

Maguire, M., Brumberg, J., Ennis, M., & Shipley, T. (2011). Similarities in object and event segmentation: A geometric approach to event path segmentation. *Spatial Cognition & Computation: An Interdisciplinary Journal*, *11*(3), 254–279.

Malmsköld, L., Örtengren, R., Carlson, B. E., & Svensson, L. (2007a). Preparatory virtual training of assembly operators – An explorative study of different learning models. *Proceedings of the Swedish Production Symposium*, 1–8.

Malmsköld, L., Örtengren, R., Carlson, B. E., & Svensson, L. (2007b). Virtual Training – Towards a Design Framework. *Proceedings of the World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, Chesapeake, USA, 6299–6307.

Mayer, R. E. (2005). *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied*, *11*(4), 256–265.

Merrill, E. C., Lookadoo, R., & Rilea, S. (2003). Memory, language comprehension, and mental retardation. *International Review of Research in Mental Retardation*, *27*, 151–189.

Migliore, A. (2010). Sheltered workshops. In J. Stone & M. Blouin (Eds.), *International Encyclopedia of Rehabilitation*. Buffalo: Center for International Rehabilitation Research Information and Exchange.

Moreno, R. (2007). Optimising learning from animations by minimising cognitive load: Cognitive and affective consequences of signalling and segmentation methods. *Applied Cognitive Psychology*, *21*, 765–781.

Moskaliuk, J., Bertram, J., & Cress, U. (2013). Training in virtual environments: Putting theory into practice. *Ergonomics*, *56*(2), 195–204.

Mura, K., Petersen, N., Huff, M., & Ghose, T. (2013). IBES: A tool for creating instructions based on event segmentation. *Frontiers in Psychology*, *4*, 994.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology*, *28*(1), 28–38.

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, *12*(5), 436–450.

Nilsson, S., & Johansson, B. (2006). User experience and acceptance of a mixed reality system in a naturalistic setting: a case study. *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR '06)*, Santa Barbara, USA, 247–248.

Numminen, H., Lehto, J. E., & Ruoppila, I. (2001). Tower of Hanoi and working memory in adult persons with intellectual disability. *Research in Developmental Disabilities*, *22*, 373–387.

Papenmeier, F. (2014). segmag: Determine event boundaries in event segmentation experiments. R package version 1.2.2.

Parasuraman, R. (1990). Event-related brain potentials and human factors research. In J. W. Rohrbaugh, R. Parasuraman, & R. Johnson (Eds.), *Event-related brain potentials: Basic issues and applications* (pp. 279–300). New York: Oxford University Press.

Peirce, J. (2007). PsychoPy - Psychophysics Software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13.

Petersen, N., & Stricker, D. (2012). Learning task structure from video examples for workflow tracking and authoring. *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR '12)*, Atlanta, USA, 237–246.

R Development Core Team. (2008). *R: A language and environment for statistical computing*.

Radvansky, G. A., & Copeland, D. E. (2006). Walking through doorways causes forgetting: Situation models and experienced space. *Memory & Cognition*, *34*(5), 1150–1156.

Radvansky, G. A., & Zacks, J. M. (2014). *Event cognition*. Oxford: Oxford University Press.

Reagh, Z. M., & Yassa, M. A. (2014). Repetition strengthens target recognition but impairs similar lure discrimination: Evidence for trace competition. *Learning & Memory*, *21*, 342–6.

Richardson, M., & Ball, L. J. (2009). Internal representations, external representations and ergonomics: towards a theoretical integration. *Theoretical Issues in Ergonomics Science*, *10*(4), 335–376.

Richardson, M., & Jones, G. (2011). Identifying the task characteristics that predict children's construction task performance. *Applied Cognitive Psychology*, *25*(3), 377–385.

Richardson, M., Jones, G., Torrance, M., & Baguley, T. (2006). Identifying the task variables that predict object assembly difficulty. *Human Factors*, *48*(3), 511–525.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255.

Royston, P., Altman, D. G., & Sauerbrei, W. (2006). Dichotomizing continuous predictors in multiple regression: A bad idea. *Statistics in Medicine*, *25*, 127–141.

Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., … Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition*, *129*, 241–255.

Schank, R., & Abelson, R. (1977). *Scripts, plans, goals, and understanding*. Hillside: Lawrence Erlbaum.

Schwan, S., & Garsoffky, B. (2004). The cognitive representation of filmic event summaries. *Applied Cognitive Psychology*, *18*(1), 37–55.

Schwan, S., & Garsoffky, B. (2008). The role of segmentation for perception and understanding of events. In T. F. Shipley & J. M. Zacks (Eds.), *Understanding events - From Perception to Action* (pp. 391–414). New York: Oxford University Press.

Schwartz, M., Segal, M., Veramonti, T., Ferraro, M., & Buxbaum, L. (2002). The Naturalistic Action Test: A standardised assessment for everyday action impairment. *Neuropsychological Rehabilitation*, *12*(4), 311–339.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701–703.

Shneiderman, B. (1989). Reflections on authoring, editing, and managing hypertext. In E. Barrett (Ed.), *The society of text* (pp. 1–22). Cambridge: MIT Press.

Sonnentag, S., & Kleine, B. M. (2000). Deliberate practice at work: A study with insurance agents. *Journal of Occupational and Organizational Psychology*, *73*(1), 87–102.

Southwick, J., Bigler, E., Froehlich, A., DuBray, M., Alexander, A., Lange, N., & Lainhart, J. (2011). Memory functioning in children and adolescents with autism. *Neuropsychology*, *25*(6), 702–710.

Spanjers, I. A. E., van Gog, T., & van Merriënboer, J. J. G. (2010). A theoretical analysis of how segmentation of dynamic visualizations optimizes students' learning. *Educational Psychology Review*, *22*(4), 411–423.

Stan, E., & Mosley, J. (1988). Semantic encoding by mildly retarded and non-retarded individuals. *Journal of Intellectual Disability Research*, *32*(5), 371–382.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149.

Sterr, A. (2004). Attention performance in young adults with learning disabilities. *Learning and Individual Differences*, *14*(2), 125–133.

Stork, S., & Schubö, A. (2010). Human cognition in manual assembly: Theories and applications. *Advanced Engineering Informatics*, *24*(3), 320–328.

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236–257.

Tomporowski, P. D., & Hayden, A. M. (1990). Employment of individuals with developmental disabilities. In G. Reid (Ed.), *Problems in Movement Control* (pp. 133–158). North-Holland: Elsevier Science Publishers

Treisman, A., & Tuxworth, J. (1974). Immediate and delayed recall of sentences after perceptual processing at different levels. *Journal of Verbal Learning and Verbal Behavior*, *13*, 38–44.

Tversky, B., Zacks, J. M., Lee, P., & Heiser, J. (2000). Lines, blobs, crosses and arrows: Diagrammatic communication with schematic figures. In M. Anderson, P. Cheng, & V. Haarslev (Eds.), *Diagrams 2000* (pp. 221–230). Berlin, Heidelberg: Springer.

Tversky, B., Zacks, J. M., Morrison, J. B., & Hard, B. M. (2011). Talking about events. In J. Bohnemeyer & E. Pederson (Eds.), *Event representation in language and cognition* (pp. 216–227). New York: Cambridge University Press.

Van Gog, T., Paas, F., & Sweller, J. (2010). Cognitive Load Theory: Advances in research on worked examples, animations, and cognitive load measurement. *Educational Psychology Review*, *22*, 375–378.

Van Roon, D., Caeyenberghs, K., Swinnen, S. P., & Smits-Engelsman, B. C. (2010). Children with a learning disorder show prospective control impairments during visuomanual tracking. *Research in Developmental Disabilities*, *31*, 195–202.

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems, Man, and Cybernetics*, *22*(4), 589–606.

Vignais, N., Miezal, M., Bleser, G., Mura, K., Gorecky, D., & Marin, F. (2013). Innovative system for real-time ergonomic feedback in industrial manufacturing. *Applied Ergonomics*, *44*, 566–574.

Visier, L. (1998). Sheltered employment for persons with disabilities. *International Labour Review*, *137*(3), 347–365.

Webel, S., Bockholt, U., Engelke, T., Peveri, M., Olbrich, M., & Preusche, C. (2011). Augmented Reality training for assembly and maintenance skills. *BIO Web of Conferences*, *1*(00097), 1–4.

Wehmeyer, M., Smith, S., & Palmer, S. (2004). Technology use and people with mental retardation. *International Review of Research in Mental Retardation*, *29*, 291–337.

Whitney, C., Huber, W., Klann, J., Weis, S., Krach, S., & Kircher, T. (2009). Neural correlates of narrative shifts during auditory story comprehension. *NeuroImage*, *47*, 360–366.

Wobbrock, J. O., Kane, S. K., Gajos, K. Z., Harada, S., & Froehlich, J. (2011). Ability-based design: Concept, principles, and examples. *ACM Transactions on Accessible Computing*, *3*(3), 1–36.

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science*, *28*, 979–1008.

Zacks, J. M., Kumar, S., Abrams, R. A., & Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition*, *112*(2), 201–216.

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction error associated with the perceptual segmentation of naturalistic events. *Journal of Cognitive Neuroscience*, *23*(12), 4057–4066.

Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, *138*(2), 307–327.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, *133*(2), 273–293.

Zacks, J. M., Speer, N. K., Vettel, J. M., & Jacoby, L. L. (2006). Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychology and Aging*, *21*(3), 466–482.

Zacks, J. M., Swallow, K. M., Vettel, J. M., & McAvoy, M. P. (2006). Visual motion and the neural correlates of event perception. *Brain Research*, *1076*(1), 150–162.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21.

Zacks, J. M., & Tversky, B. (2003). Structuring information interfaces for procedural learning. *Journal of Experimental Psychology: Applied*, *9*(2), 88–100.

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, *130*(1), 29–58.

Zalla, T., Labruyère, N., Clément, A., & Georgieff, N. (2010). Predicting ensuing actions in children and adolescents with autism spectrum disorders. *Experimental Brain Research*, *201*(4), 809–819.

Zalla, T., Labruyère, N., & Georgieff, N. (2013). Perceiving goals and actions in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*(10), 2353–2365.

Zook, N., Davalos, D., DeLosh, E., & Davis, H. (2004). Working memory, inhibition, and fluid intelligence as predictors of performance on Tower of Hanoi and London tasks. *Brain and Cognition*, *56*(3), 286–292.

Zwaan, R., Langston, M., & Graesser, A. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*(5), 292–297.

Lebenslauf

Persönliche Daten

Name: Katharina Sebastian, geb. Mura

Heimatanschrift: Marienallee 2
01099 Dresden

Ausbildung

09/2011 – 09/2015	Technische Universität Kaiserslautern
09/2014 – 12/2014	Washington University in St. Louis, USA
09/2010 – 06/2011	Graduate School of Economic and Social Sciences, Universität Mannheim
10/2004 – 07/2010	Universität Mannheim Studium der Psychologie mit Nebenfach Mathematik, Diplom
10/2008 – 02/2009	Universität Kardinal Stefan Wyszynski in Warschau, Polen
08/1995 – 06/2004	Marie-Luise-Kaschnitz Gymnasium, Völklingen

Berufliche Erfahrungen und Praktika

10/2010 – 08/2011, 03/2013 – 07/2015	Researcher am Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI GmbH), Innovative Fabriksysteme, Kaiserslautern
09/2011 – 02/2013	Researcher bei Technologie-Initiative SmartFactoryKL e.V., Kaiserslautern
04/2010 – 07/2010	Praktikum bei Continental Automotive GmbH, Babenhausen
04/2008 – 07/2008	Praktikum bei Sozial- & Gesundheitspsychologie, Universität Zürich
08/2007 – 02/2010	Studentische Hilfskraft bei Berufsgenossenschaft Nahrungsmittel und Gaststätten (BGN), Mannheim
12/2005 – 07/2007	Wissenschaftliche Hilfskraft „Biologische Psychologie", Universität Mannheim