Justus-Liebig-Universität Gießen
Fachbereich 06
Institut für Sportwissenschaft

Dynamic Capacity Allocation in Motor-Cognitive Dual-Tasking

- probed by Semantic Auditory Stimuli

Dissertation submitted in fulfilment of the requirements
for the degree
Dr. rer. nat.
by

Jelena Müller (M.Sc.)

Table of Contents

List of Figures

List of Tables

Abbreviations

A

ANOVA .... *analysis of variance*

C

Calc .... *calculation phase of the calculation task*

CalcForce ....*event combining the calculation phase and force tracking*, *event combining calculation phase and force tracking*

CalcGoProbe .... *event combining the calculation phase with a Go probe stimulus*

CalcNoGoProbe .... *event combining the calculation phase with a NoGo probe stimulus*

CalcOnly .... *event during the calculation phase without probe*

CalcOnlyForce .... *event combination of calculation phase and force tracking but no probe stimulus during triple tasking*

CalcProbe .... *event during the calculation phase with probe*

CalcProbeForce .... *event combintaion of calculation phase, probe stimulus and force tracking during triple tasking*

CLC .... *calculation task*

Comp .... *comparison phase of the calculation task*

CompForce .... *event combining comparison phase and force tracking*

CompGoProbe .... event combining the comparison phase with a Go probe stimulus

CompNoGoProbe .... *event combining the comparison phase wit a NoGo probe stimulus*

CompOnlyForce .... *event combination of comparison phase and force tracking but no probe stimulus during triple tasking*

CompProbeForce .... *event combination of comparison phase, probe stimulus and force tracking during triple tasking*

D

dir .... *direction probe stimulus*

DT .... *dual task*

DTcp .... *dual task execution of cognitive and probe task*

DTmc .... *dual task execution of motor and cognitive task*

DTmp .... *dual task execution of motor and probe task*

E

ERA .... *event-related analysis*

F

FFT .... *fast fourier transformation*

F_{max} .... *individual maximum force*

fMRI .... *functional Magnetic Resonance Imaging*

FRC .... *force tracking task*

FSR .... *force sensing resistor*

G

GoProbeForce .... *event combining a Go probe stimulus with force tracking*

H

HSD .... *honest significant difference*
Hz .... *Hertz*

I

int .... *intensity probe stimulus*
IPS .... *intraparietal sulcus*

M

M1 .... *primary motor cortex*
MRM .... *Multiple Resource Model*

N

NoGoProbeForce .... *event combining a NoGo probe stimulus with force tracking*
num .... *numerical probe stimulus*

P

PC .... *processing capacity*
PC_{100} .... *total processing capacity*
PCU .... *processing capacity utilization*
PCU_{100} .... *total processing capacity utilization*
PCU_{Task} .... *processing capacity utilization required for task execution*
PostReact .... *interval subsequent to probe response during which force tracking performance is measured*
PostStim .... *interval subsequent to probe stimulus but prior to response during which force tracking performance is measured*
PreStim .... *interval prior to probe stimulus during which force tracking performance is measured*
PRT .... *probe reaction time task, probe reaction time task*
pSTG .... *posterior superior temporal gyrus*

Q

qua .... *quantitative probe stimuli*

R

R1 .... *frist line of research*
R2 .... *second line of research*
R3 .... *third line of research*
rmANOVA .... *repeated-measures analyses of variance*
RMSE .... *root mean square error*
ROI .... *regions of interest*
RT .... *reaction time*

S

S1 .... *primary somatosensory cortex*
S2 .... *secondary somatosensory cortex*
spa .... *spatial probe stimuli*
S-R-mappings........ *Stimulus-Response-mappings*
ST .... *single task*
STc .... *single task execution of cognitive calculation task*
STm .... *single task execution of motor force tracking task*
STp .... *single task execution of probe reaction time task*
SW test .... *Shapiro-Wilk test*

T

TMS........ *transcranial magnetic stimulation*
TR........ *time regimes*
TT .... *triple task*

1 Introduction

Real life imposes various demands on individuals, often requiring the execution of multiple concurrent tasks. A typical scenario involves, for example, walking across a crosswalk while engaging in text messaging. Both tasks requiring attention to be executed. Performance in such tasks can be evaluated along two primary dimensions: time and accuracy. The time dimension reflects the duration required for task completion, whereas accuracy pertains to the quality of task execution, quantified by the frequency and severity of occurring errors during task execution. For the combined tasks of street crossing while texting, both concurrent activities can be analyzed within these two performance dimensions. Time until text completion and occurrence of typing errors account for time and accuracy dimension for the texting task. The duration for road crossing and incidence of accidents measure performance in the road crossing task along the two dimensions. These inquiries gain significance when the simultaneous execution of multiple tasks leads to a decline in performance along one or both dimensions. While delays or errors in text messaging may be negligible, failures to perceive oncoming traffic during pedestrian crossings could precipitate severe accidents, underscoring the critical implications of performance decrements in multitasking scenarios.

Although this illustration draws directly from real-life scenarios, assessing performance in multiple executed tasks may present nontrivial challenges. Basic research in cognitive psychology has generated a wealth of meticulously designed experiments conducted in laboratory settings. Within these experimental frameworks, performance metrics related to time and accuracy of execution can be controlled and quantified to investigate multitasking capabilities. However, it is important to note that these experiments often involve relatively simple reaction time tasks, which may not fully capture the intricacies of real-world complexities.

Furthermore, basic psychological studies usually assume that the processing regime is induced more or less automatically by the respective task constellation and then remains relatively static. However, this assumption could be questioned, as a dynamically responsive system could offer advantages, particularly for responding adequately to execution errors. This highlights the need to examine the cognitive processing more closely and, above all, with sufficiently high temporal resolution.

The objective of work described here is to establish a paradigm that allows for self-organized execution of multiple tasks under conditions that more closely resemble real-world scenarios, allowing for conclusions that remain valid beyond the confines of a laboratory. In the first experiment, a cognitive and a motor task are combined in a self-organized task setting for participants. An additional auditory reaction time task further probes task integration at specific times during task execution. The experiment investigates general phenomena such as performance decline during multitasking execution, as well as specific effects related to the semantics of stimuli and the time-dependent nature of the tasks. Findings from this experiment will be used for a second, similar study. Here, the methodology is primarily adapted to exclude possible interferences in the response behavior of participants. Additionally, semantics of probe stimuli are slightly altered to further test specifics of their interference with motor and cognitive tasks. In a third experiment, a Go-NoGo paradigm is added to the auditory response task to test findings about the cognitive processing steps involved in inhibiting responses during task processing.

The unique nature of all three experiments is that performance is not solely recorded and averaged on a blockwise basis, but studied with a higher temporal resolution at several discrete times during task execution. This provides a deeper insight into targeted events during the execution of continuously executed multiple tasks.

In the subsequent chapters, attention is analyzed from a cognitive psychological standpoint as an intermediary in cognitive processes (James, 1950) for performing a specific task. Capacity theory (Kahneman, 1973) is employed to describe and quantify attention as a limited resource. The suggestion that this limited capacity can manifest in terms of time and accuracy deterioration is then visually modeled in this work. Potential outcomes thereof are examined within the frameworks of both capacity theory and the multiple resource model as proposed by Wickens (2002). Additionally, traditional bottleneck theories (e.g. Broadbent, 1958; Treisman, 1964) are explored, emphasizing different stages of information processing. This exploration provides the basis for hypothesizing variable capacity allocations across different processing stages, which can be quantitatively assessed using time and accuracy metrics.

2 Basic Assumptions about Attention

Can I have your full attention?

In common understanding, attention is often perceived as a metaphysical construct primarily under voluntary control and direction, as exemplified in the preceding statement. Thus, attention mediates the cognitive processing of incoming sensory information, enabling the selection of appropriate responses to perceived stimuli to execute tasks (James, 1950; Pashler, 2004). In cognitive psychology, there is no strict definition of what constitutes a task. According to Koch et al. (2018), a task can be described as a self-instructed or instructed cognitive goal, accompanied by requisite motor and cognitive representations necessary to achieve said goal. Tasks can encompass a wide range of activities, from simple key responses to stimuli as seen in the psychological refractory period (PRP) paradigm, to visuomotor tracking tasks, mental operations like mental arithmetic, and the execution of complex movements such as performing a pirouette. Crossing the road or writing a text message are also suitable for being classified as tasks. Task completion requires adequate and specific information processing. The central objective of attention research is to identify the specific attentional processes required to complete a given task. However, despite its seemingly inherent nature, the study and precise quantification of attention remain fundamental challenges in cognitive research.

In this chapter, a new approach to the distribution of attentional processes during task execution is presented, which serves to create the paradigm later described in this thesis.

2.1 Attention as Processing Capacity Utilization per Time

Attention is often thought to have a quantitative dimension; it can be spent. The reservoir from which attention is drawn and directed to specific stimuli processing that initiates task execution is called capacity. Each task requires a specific amount of attributed capacity to be executed. Yet, capacity is a limited resource (Kahneman, 1973). The absolute amount of cognitive processing capacity (PC) in each individual is an unknown variable. However, processing capacity utilization (PCU) can be measured to some extent.

The basic assumption for human behavior being that the goal is to execute tasks flawlessly. If errors occur during task execution, this can be interpreted as proof that processing capacity is depleted. Insufficient processing capacity can reveal itself in two dimensions: time and quality. Time required for task completion can be prolonged, and/or the quality of task execution, i.e., of its results, may suffer. The capacitive requirements that a task places on central processing can therefore be described in two dimensions by the area between the time and the processing capacity utilization ($Time_{Task}$ x PCU_{Task}) as depicted in Figure 1.

Task execution is not impaired if the processing capacity utilization over time for a task stays below the general amount of processing capacity ($PCU_{Task} < PCU_{100}$) of the central nervous system.

This way of thinking is based on the physical law $P = \frac{\delta W}{dt}$, where Power equals Work per time (Nolting, 2018, p. 253); or Work can be described by Power multiplied by time. Accordingly, in the model described here, PC_{100} as equivalent to Work, would result from multiplying PCU_{Task} and $Time_{Task}$ ($PC_{100} = PCU_{Task} \times Time_{Task}$).

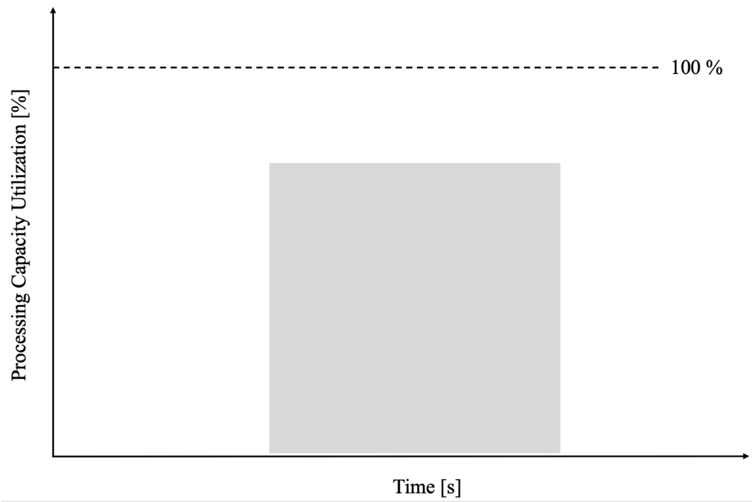

Figure 1: Capacitive requirements for task processing can be described by the area between its time needed for task execution and processing capacity utilization (PCU) for the task. The general limit of processing capacity utilization is 100% (PCU_{100}).

In real-life scenarios, single tasks are executed in very few moments. Therefore, the responsibility lies within the individual to organize incoming stimuli based on their relevance or prompting nature. It may therefore well be that only a certain proportion of total processing capacity (PC_{100}) is available for processing one task because a specific amount must be distributed to another task. How much PCU is made available per time by the central nervous system can be explained by different models. According to Kahneman's capacity theory, an unspecific reservoir containing the PC_{100} would distribute PC according to the tasks to be processed by means of an allocation policy, which is determined in particular by arousal (Kahneman, 1973), which is induced by complexity or novelty of task stimuli (Berlyne, 1960). This implies that both tasks compete for capacity allocation of the PC_{100} (see Figure 2). Wickens' (2002) multiple resource model would assume that the PC_{100} is divided into different reservoirs, which are specifically available for different modalities. If two tasks draw from the same resource, they would have to compete for capacity. The latter case is also depicted in Figure 2, where the 100% would not refer to PC_{100} but rather to 100% of a specific reservoir being accessed by both tasks.

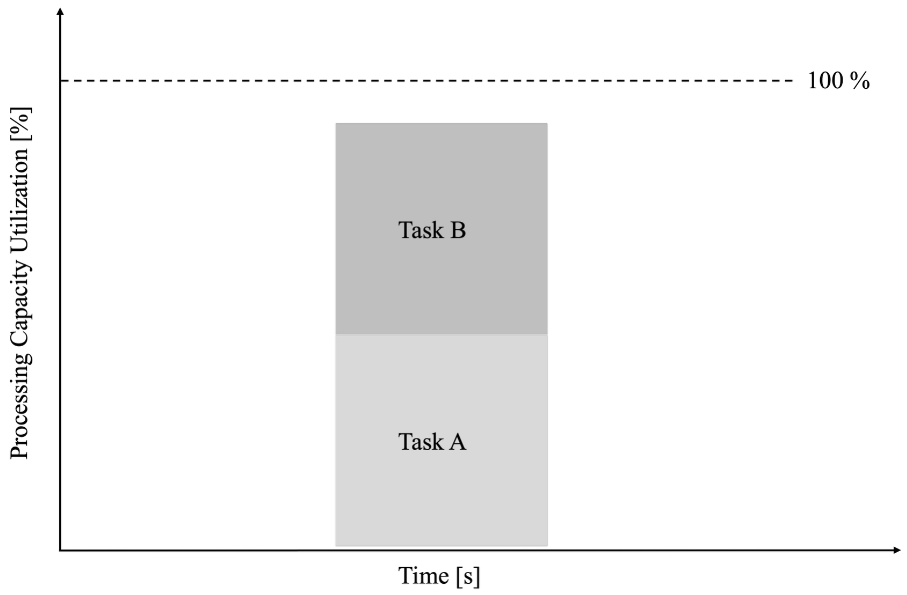

Figure 2: Two tasks (A and B) drawing from the central capacity (PCU_{100}) and therefore competing for it, as explained by the capacity theory (Kahneman, 1973), or two tasks drawing from the same resource under the multiple resource model (Wickens, 2002).

If the tasks however, draw capacity from different resources, they do not compete for the same processing capacity under the multiple resource theory (see Figure 3) and can be executed undisturbedly and concurrently.

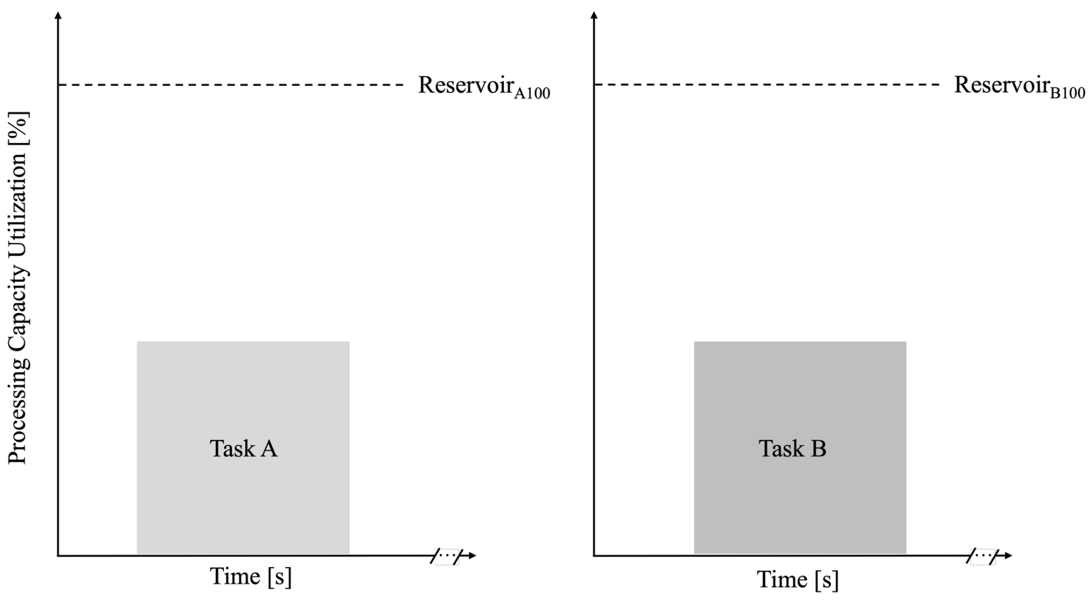

Figure 3: Two tasks (A and B) drawing capacity from different resources (Reservoir A_{100} and Reservoir B_{100}) as explained by the multiple resource theory (Wickens, 2002). Here, the 100% of each reservoir refers to the full amount of capacity allocated to a specific resource.

Theories rooted in a structural bottleneck propose that task interference arises because two processes necessitate the same structure and thus cannot be executed concurrently (Broadbent, 1958; Deutsch & Deutsch, 1963; Norman, 1969; Treisman, 1964). The specific processes that impede each other depend on that structure. However, in the capacity model (Kahneman, 1973), interference is not specific or based on rigid structures. It arises whenever there is not enough capacity available or released from the limited reservoir for a given process to function adequately. This prompts further exploration into allocation policy. Factors influencing whether an activity places high or low demands on capacity may include its complexity, novelty, urgency, and relevance to current goals (Berlyne, 1960). The amount of available capacity is likely controlled through mechanisms involving arousal levels, cognitive resources, and attentional processes (Kahneman, 1973). As for the rules guiding the allocation policy, they may involve prioritizing tasks based on their importance, timing, and potential consequences, as well as dynamically adjusting allocations based on real-time feedback and changing task requirements (Lund, 2001). Of course, this continuous closed feedback loop causes a high processing overhead when allocating capacity and is therefore rather inefficient.

In comparison to the capacity theory, the multiple resource model (Wickens, 2002) is similarly rigid as the structural bottleneck models, as it only permits interference in four specified dimensions (modalities, codes, stages and responses) and does not account for interference occurring anywhere in between. On the other hand, it also facilitates the specific testing of capacity distribution across different dimensions. However, the model assumes that distinct resources exist due to separate processing locations in the brain. This notion is overly simplistic, as some task types that initially appear to engage different processes, show significant coordination between brain regions, making clear localization difficult. An example for this is action comprehension, where somatosensory and motor brain regions, as well as regions for linguistic processing can be active when describing actions through language (Tettamanti et al., 2005). The multiple resource model thus seems somewhat rigid, as it overlooks the fact that most tasks, particularly more complex ones with real-world applicability, are not confined to discrete brain regions.

The multiple resource model with its four specific dimensions is not considered for the model presented in this work. Of course, it must be assumed that different resources are only available to certain stimuli and response behaviors. However, this specificity is initially negligible. As for the introduced model in this thesis, it appears irrelevant from which reservoir the available PC for a task is drawn. An unspecific limit of PCU_{100} is assumed which could derive from one or multiple reservoirs. To what extent a universal reservoir can account for task interferences is one of the subjects of this work. Furthermore, the model presented here assumes a two-dimensionality, whereas both capacity theory and the multiple resource model think one-dimensionally to a certain extent. Both well-known models assume that a lack of capacity always results in task errors. These errors are usually described as a decrease in performance and are therefore not defined in more detail. Implicitly describing a combination of increased task execution time and an accuracy trade-off. In the newly introduced model, time and accuracy as described by PCU represent two separate dimensions, which are, however, linked with an internal dependency in that they describe the area of capacity required for a task.

2.2 Operationalization of Capacity Allocation for Task Execution

A task is executed undisturbedly if enough PCU per time can be attributed to it. Thus, capacity allocation can be operationalized in two dimensions: time taken for task execution and processing capacity utilization, as measured by the quality of task results, i.e., erroneous responses.

If task execution has less time available than normally required, PCU must be increased so that the total area of allocated capacity remains the same, thus enabling undisturbed execution (see Figure 4). A task can have such a fixed time regime if only a certain amount of time is available for task execution. This might be the case if a stimulus requires a response during a certain time span before the next stimulus is presented. If the increase in PCU cannot be accommodated for, because PCU_{100} is exploited, the task becomes prone to measurable erroneous response behaviors. Such as a missed response to an auditory stimulus.

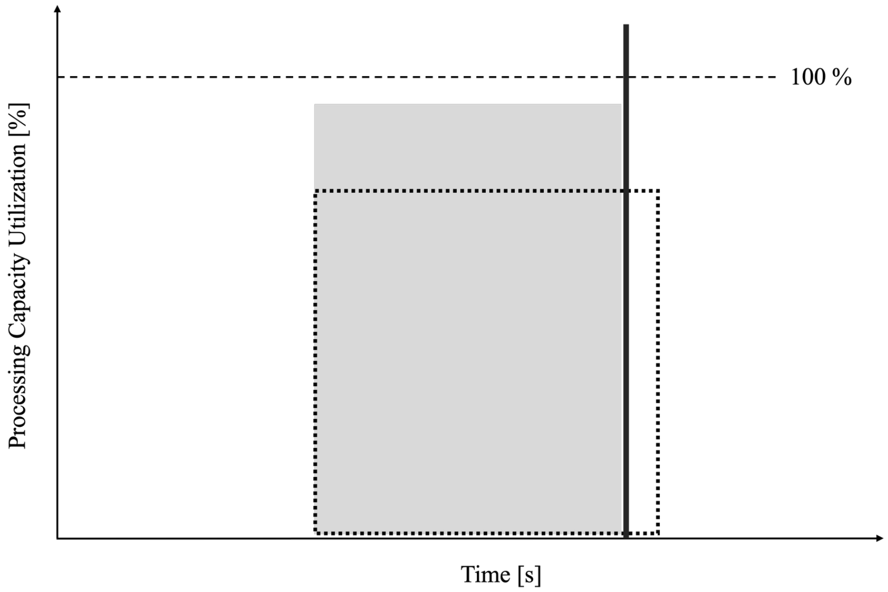

Figure 4: PCU for a task is increased to compensate for decreased time available for task completion to ensure undisturbed execution. Execution is successful because PCU_{Task} < PCU_{100}.

On the other hand, if there is no time limit for a task, it can simply be extended in the time dimension in the absence of sufficient PCU to avoid a tradeoff in accuracy. Accordingly, an extension of the processing time of the task can then be measured. These considerations become relevant if, for example, there is not enough PCU available for a task because it must compete with another task for the PC_{100}. Or if two tasks are intertwined in terms of time, so that they must be organized accordingly.

The use of the psychological refractory period (PRP) paradigm, first discovered by Telford (1931) and later described by Welford (1952), has sufficiently demonstrated that a prolongation of such a task processing time is exactly the case. Welford (1952) would assume that stimulus processing takes some time, so that when two stimuli are presented in close temporal succession, the response to the second stimulus is delayed due to a temporal bottleneck caused by only one channel for stimulus processing (Broadbent, 1954). Another explanation could be that not enough PCU per time is available to respond to both stimuli in such close temporal succession. Thus, the response to the second stimulus is expanded, to guarantee, that PCU_{100} is not exceeded (Figure 5).

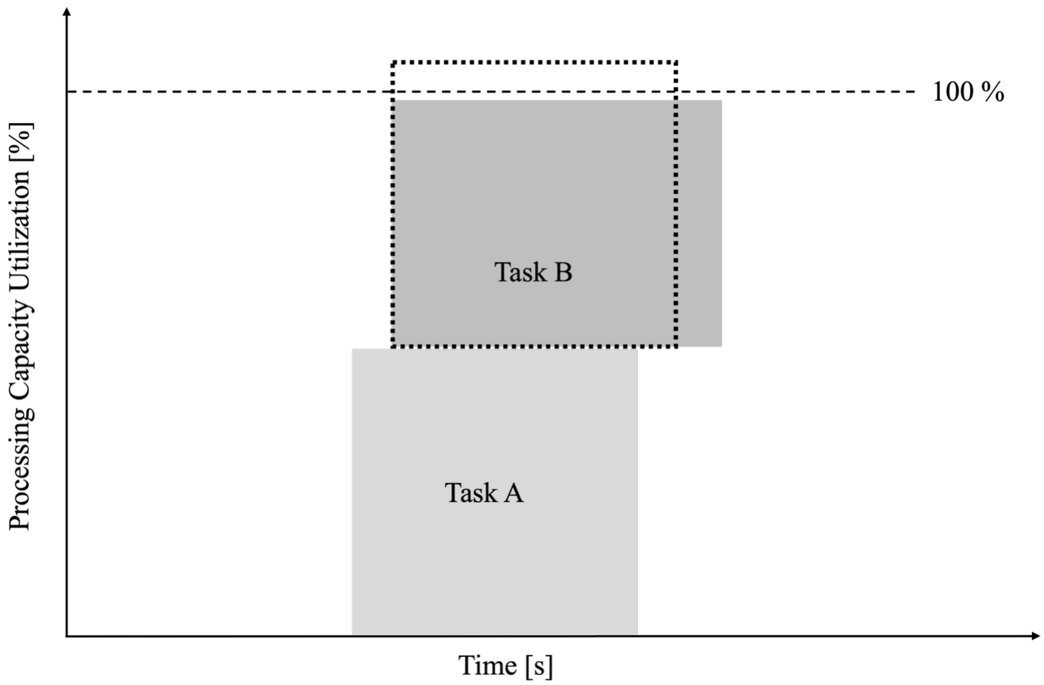

Figure 5: Task A can be executed undisturbed, while Task B is elongated in time as shown by the PRP paradigm.

Under the PRP paradigm, there seems to be a first-come-first-served principle based on a structural bottleneck, as the response to the first stimulus is not impaired in the time dimension. The combination of two concurrent tasks can now be viewed from the perspective that there are two different time regimes, namely that tasks can either be extended in the time dimension or they cannot. Whether this is due to innate task constraints or instructions is neglected in this moment. These considerations give rise to further possibilities for PCU allocation when working on several tasks simultaneously.

2.3 Processing Capacity Utilization Allocation during Multiple Task Execution

When executing two tasks concurrently, PC_{100} must be divided during the time for task execution between both tasks to meet their requirements as described by the area between PCU_{Task} x $Time_{Task}$. By this means, the sum of the areas for both tasks (PCU_A x $Time_A$, PCU_B x $Time_B$) should lie beneath PC_{100} at all times to avoid errors that would result if PC_{100} is exploited by task demands. However, as some tasks have time limits, there are several possible combinations with those different time regimes. Two single-choice reaction-time tasks, as combined in the PRP paradigm, could theoretically both be prolonged in this model, as they have a flexible time regime. However, the results from the paradigm account for a prolonged response time only for the stimulus presented after a specific stimulus onset asynchrony (as was depicted in Figure 5). When combining a task with a fixed and a flexible time regime, it seems feasible that the task with the flexible time regime is processed with a delay. The potential outcome if this task had been initiated first and, in accordance with the first-come-first-served principle, should continue uninterrupted, can only be speculated upon. Nonetheless, if preventing errors is the primary objective, extending the task with the flexible time regime is also advisable if its underlying stimulus was presented before the task stimulus of the task operating under a fixed time regime (see Figure 6).

In this scenario, it can be inferred that Task A can initially proceed without disruption until Task B is initiated. Once Task B commences, the available PCU becomes inadequate to maintain below the PC_{100} threshold unless the task durations are extended. To prevent processing errors due to exceeding the PC_{100}, task A is then extended due to its flexible time regime to precisely compensate for the area required by Task B, ensuring that both tasks remain below the PC_{100} limit. In other words, Task B partially displaces Task A to ensure that an adequate PCU × Time area is available for its execution. Task A can then append the area that was displaced.

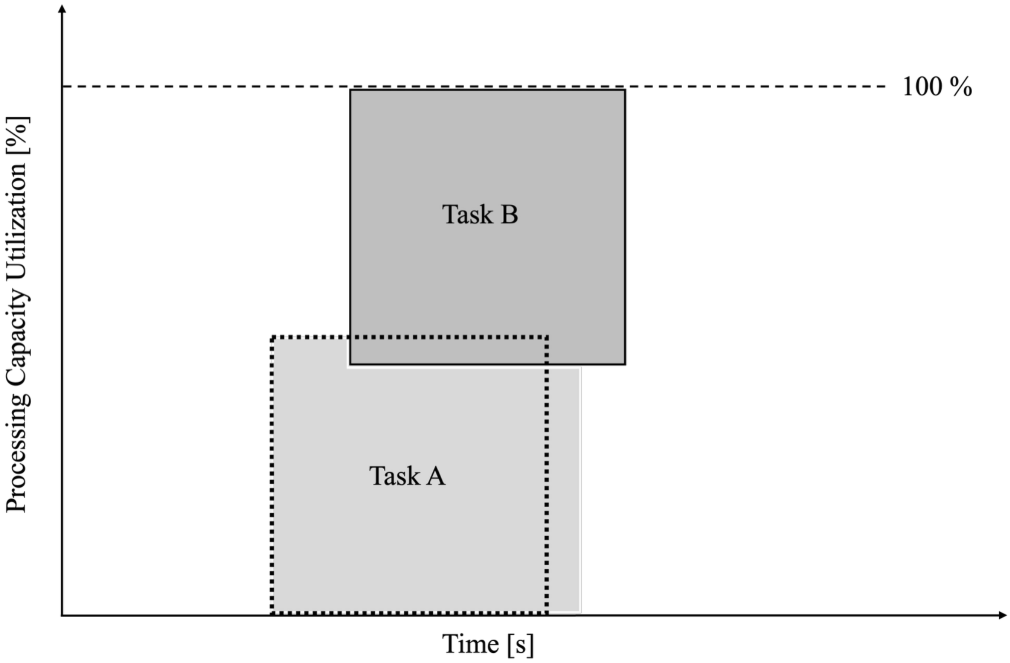

Figure 6: Dual-task execution of Task A with a flexible time regime and Task B with a fixed time regime. Although Task B is initiated second, Task A (flexible time regime) is extended in the time dimension from the time of stimulus onset of Task B on, to prevent PCU_{100} from being exceeded and thus the occurrence of errors.

When two tasks with a fixed time regime are combined, the occurrence of a decline in task accuracy in one of those tasks is inevitable if the sum of PCU for both tasks exceeds PCU_{100}. (see Figure 7). This could only be prevented if tasks were executed serially.

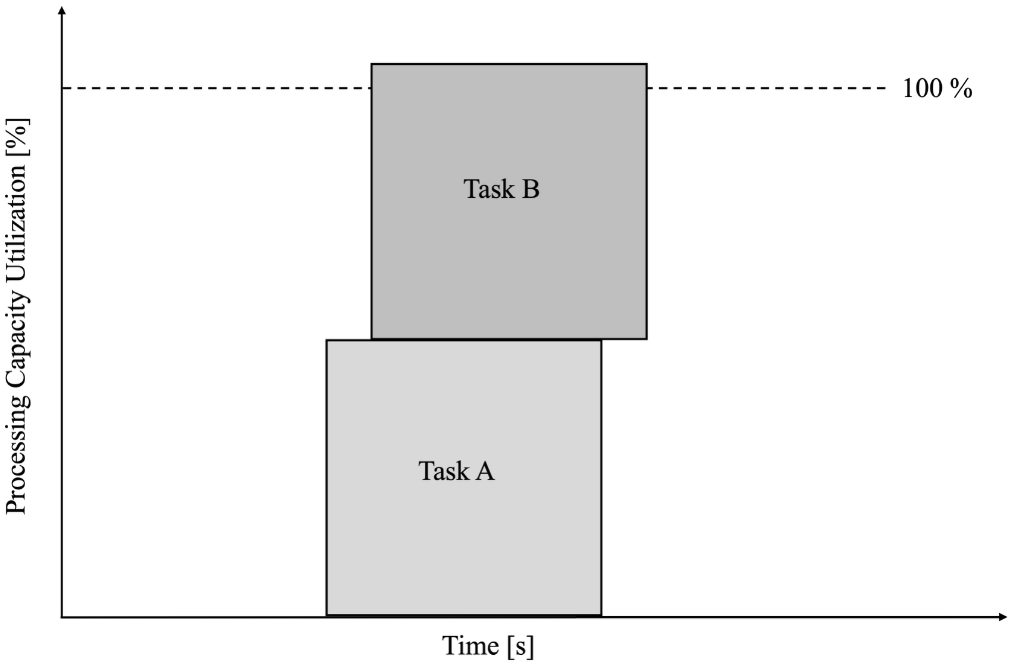

Figure 7: Dual-task execution of two tasks with fixed time regimes. The first initiated Task A will run unaffected, while Task B is prone to errors according to PCU_{100} depletion.

With this established model, a foundation for possible behavior during task combination with different time regimes is introduced. So far, the individual processing steps during task execution have been shown as grey areas. In the next subchapter, however, the individual processing steps will be analyzed in more detail and will be integrated into the model.

2.4 Processing Capacity Utilization During Different Processing Stages

One of the central concepts in cognitive psychology and neurology is the information processing approach, where computer science principles serve to unravel cognitive processes. This is important because perception of, and response to, stimuli are not immediate results of stimulation but occur as a result of processing information over time (Lund, 2001).

Therefore, the gray area for a task that resulted from the processing capacity utilization during the time needed for task execution can be filled with three basic processing steps of information-handling operations (Attneave, 1959). These are perception of the stimulus, response selection to plan and generate an adequate output, and its response execution. Adding to these three central processing steps, the residual PRP paradigm introduces yet another processing step.

In earlier research, the PRP paradigm has only considered cases where the stimulus onset asynchrony is short enough to place the second stimulus before response execution to the first. In the residual PRP paradigm, the stimulus for the second task is given after response execution to the first stimulus is registered (Wirth et al., 2018). The reaction time to the second stimulus increases the closer its stimulus occurs after response execution to the first stimulus. Since an overlap of response selection phases is not possible, a capacity-limited effect monitoring hypothesis seems to explain these findings (Jentzsch et al., 2007). With this effect monitoring paradigm (also called backward crosstalk effect), another component is added to the information processing model. Our actions, or responses to stimuli, produce environmental effects that must be monitored to compare the intended effects to the produced effects. An ideomotor approach assumes that motor actions cannot be directly initiated through accessing muscle activity, but rather through codes of perceptual changes. Thereby, response selection is accompanied by the anticipation of sensory effects (Wirth et al., 2018). The latter must be monitored and compared to the actual sensory effects. There is evidence that these codes remain active for some time interval after response execution (Desantis et al., 2014). This process that remains active and may take up capacity even after response execution is called effect monitoring.

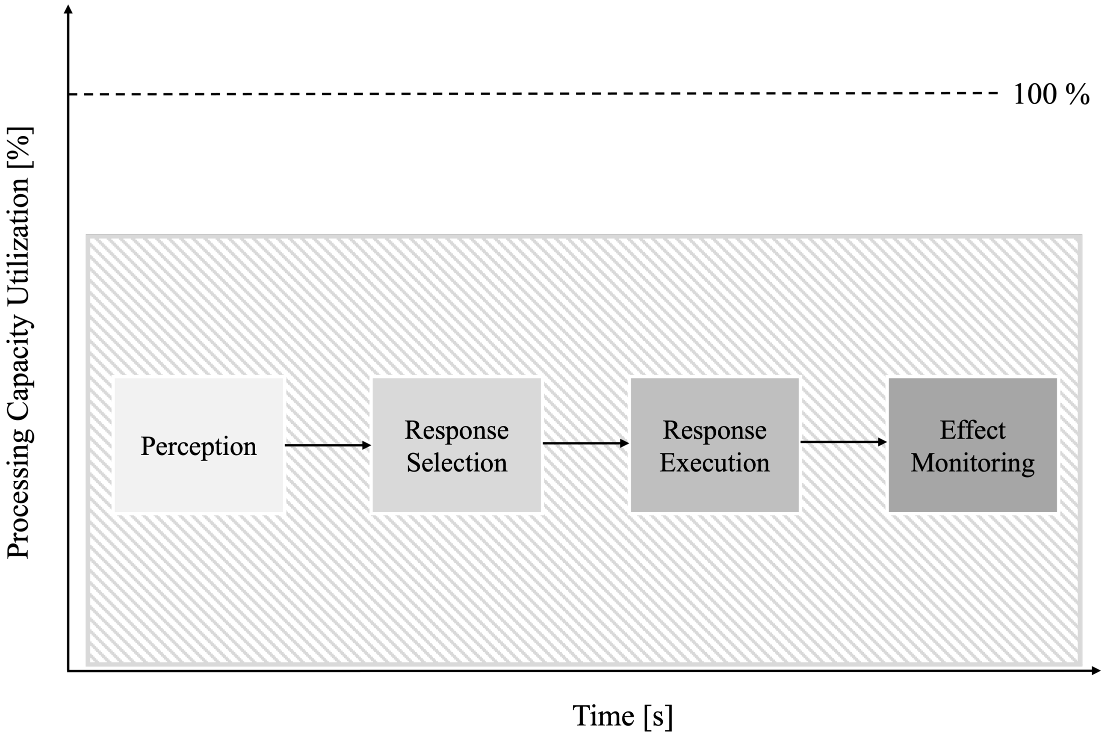

Figure 8: Central processing steps required for task execution that are described by the area between processing capacity utilization and time.

Information processing theories, which focus on structural bottlenecks and are typically measured through the PRP paradigm by response times, describe how task interference occurs over time. A key question that emerges from the previous chapters is why errors occur during task processing, even when the processing time is extended. If it is assumed that errors should be avoided due to their negative consequences and that extending the processing time is feasible, performance decrements should not be observed in both dimensions (time and PCU) simultaneously. However, this decline is consistently observed across multitasking studies. Therefore, a potential explanation for this phenomenon must first be identified.

The previously established model with two task dimensions and the division into different processing stages assumes that a maximum of 100% of PC can be distributed to a task (Figure 9, A). The different processing steps themselves can also place different demands on the PCU_{Task} x $Time_{Task}$, which are expressed by differently sized areas. The cognitive processing of several simultaneously executed tasks can be visualized as in Figure 9 (B), where the PC_{100} would then be shared by both tasks throughout the different processing stages. A certain amount of time multiplied by PCU is required to process both tasks. Ideally, both tasks should be accommodated within the available capacity. However, if tasks cannot be sufficiently extended along the time dimension, the total capacity will inevitably be exceeded, resulting in errors.

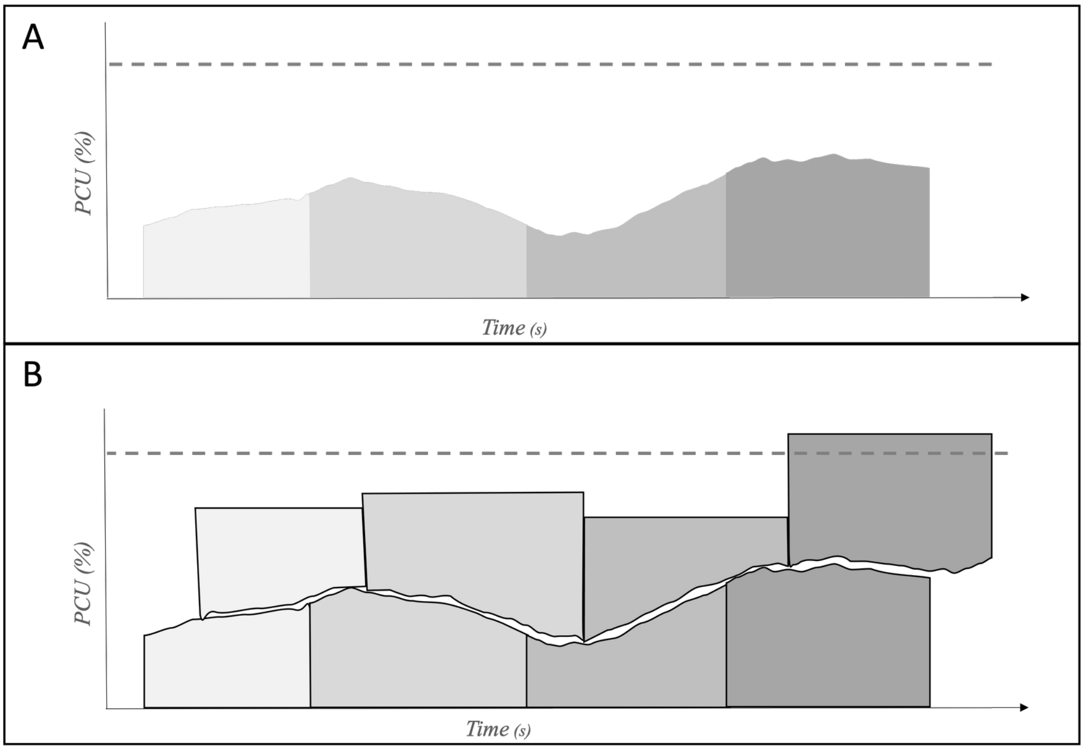

Figure 9: Model of task processing along the two dimensions time and PCU during different processing stages for a task (A).Combination of two tasks being concurrently processed (B). Grayscale represents the different processing steps of perception, response selection, response execution, and effect monitoring (from left to right). Processing phase length is suggested as equally long for the use of this model.

One explanation for this could be that even tasks with a flexible time regime cannot be extended endlessly. If the flexibility in the temporal dimension is exhausted to a maximum that does not suffice to provide enough capacity for the error-free execution of two tasks, errors could still occur in one or both tasks. However, whether this happens in the task with the fixed or flexible time regime, or in the task initiated first or second, cannot yet be specifically predicted. Another possible explanation suggests that certain processing stages of a task are subject to a fixed time regime, whereas others can be extended flexibly within a task itself. This would explain why certain error types occur during specific phases. Similar to Fitts Law (Fitts, 1954), a trade-off between time and accuracy, i.e., PCU_{Task} could exist, which is ideally balanced in such a way that a minimum of adjustments are necessary in both dimensions.

To enable more precise assumptions based on the here introduced time regimes model, the special features of different tasks with specific characteristics in the different processing stages must be addressed. The next chapter, therefore, describes the special features of processing tasks with more cognitive or motor characteristics. In addition, tasks are also considered that can be assigned to reaction time tasks by means of simple S-R-mappings, but which require a specific processing effort due to a semantic processing of stimuli and can therefore be used to specifically probe processing steps of motor or cognitive tasks.

3 Capacity Utilization under Different Task Characteristics

To understand the extent to which different task characteristics have different demands on the distribution of capacity in the two dimensions of time and PCU, it must first be clearly defined that the greatest distinction here must be made primarily between motor and cognitive tasks.

This distinction may not always be sharply defined. Take, for instance, the example of crossing a road, previously introduced as primarily a motor task. As discussed in Chapter 2.4, sensory feedback from effect monitoring outcomes is integrated into the subsequent cycle of information processing, guiding the next step. In this sense, motor tasks exhibit a higher degree of inner dependency, where each action is contingent on the successful completion of the prior one. On the other hand, writing a text message, often categorized as a cognitive task, also demonstrates some level of sequential dependency, as the correct execution of one letter influences the next in forming a word. However, in such cognitive tasks, monitoring primarily determines whether corrections or continuing with the next letters are necessary, potentially resulting in a lower degree of interdependence between steps. This contrast allows for a classification of motor tasks as more continuous, whereas cognitive tasks may be viewed as more discrete in nature, as will be further discussed in the next subchapter.

Furthermore, when texting, one could further distinguish between the cognitive component, namely the decision of which letters to choose, and the motor component, i.e., the actual typing of the letters. One can imagine that the selection is rather discrete in time, whereas the typing extends over time and requires continuous monitoring and error correction. This shows that even within one task, cognitive and motor components with different inner dependencies can be distinguished.

Therefore, in cognitive tasks (or task components), perception and response selection are certainly more decisive for the process of error-free task execution. Accordingly, these processing steps should also make up a larger proportion of the total area described by PCU_{Task} x $Time_{Task}$ (see Figure 10A). The increase in the area for a processing stage could result from both the PCU and the time dimension, if a flexible time regime is allowed. For motor tasks, in contrast to cognitive tasks, response execution and response monitoring are crucial to success. Accordingly, these two processing steps account for a large part of the total task requirements in terms of PCU_{Task} and $Time_{Task}$ (see Figure 10B).

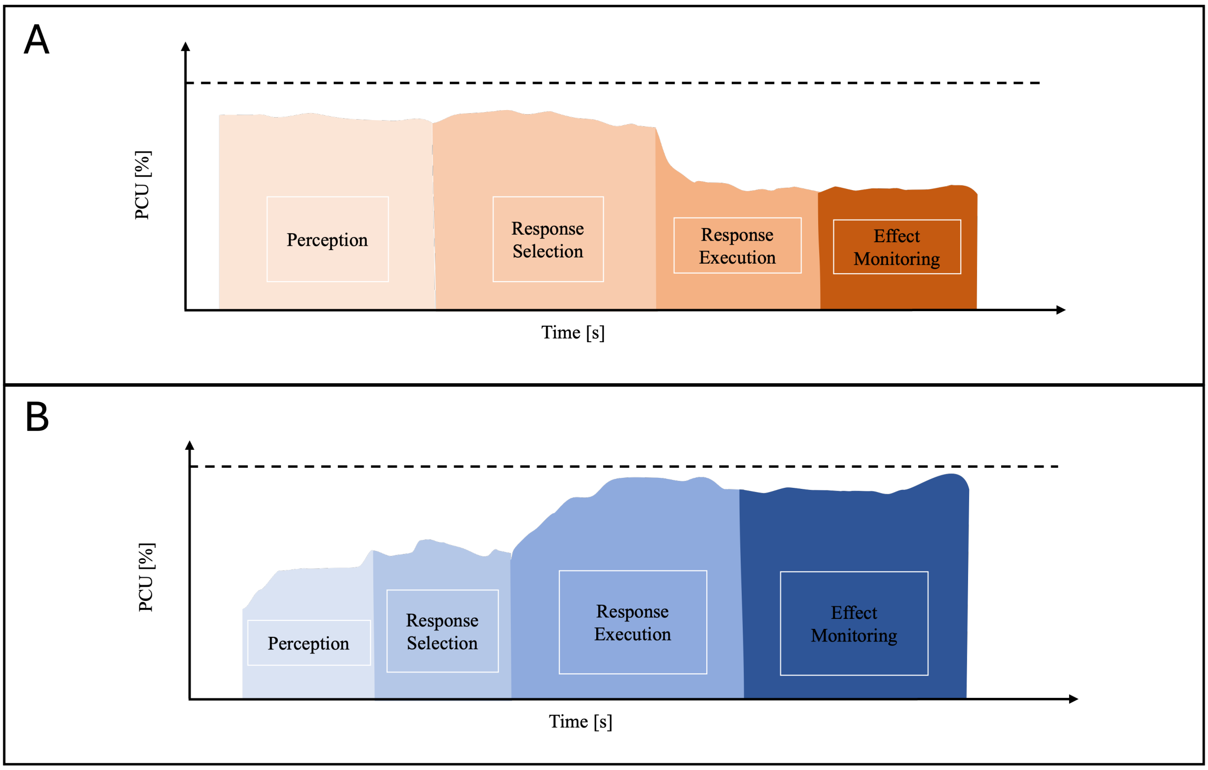

Figure 10: Time$_{Task}$ x PCU$_{Task}$ processing areas divided into different processing stages for tasks with different characteristics. (A) Cognitive task, where perception and response selection processing stages impose the highest capacity demands. (B) Motor task, where response execution and effect monitoring impose the highest capacity demands.

Naturally, this is a very simplified and generalized illustration that cannot accurately represent the actual requirements of specific tasks of a more motor or cognitive nature. Each task will in turn, have its own specific requirements in the different processing stages and within the time and PCU dimensions. Therefore, for each task used within a multi-tasking paradigm, a separate logical analysis of these requirements must be performed.

Based on the discussion above, it becomes evident that the nature and extent of task interference are contingent upon the precise alignment between processing requirements at various stages and the appropriate allocation of cognitive resources. Unlike earlier research that posited a binary classification of tasks as either entirely interference-free (i.e., automated) or not, it is now apparent that task combinations can yield highly variable outcomes. This leads to the central hypothesis that processing activities and resource allocations are dynamically adjusted to meet the specific demands of the current processing stages. These processes can be thoroughly analyzed only if the content and timing information requirements of the tasks are experimentally controlled, and measurement approaches are developed to capture data in a time-discrete manner. These aspects will be further detailed and refined in the subsequent subchapters.

3.1 Motor and Cognitive Task Interference

An extensive body of research focuses on the interference between cognitive and motor tasks. The cognitive tasks examined in this field predominantly involve auditory, verbal, and visuo-spatial working memory assessments (such as digit recall, n-back, and digit span), alongside tasks measuring inhibition (e.g., Stroop) and verbal fluency (Schott & Klotzbier, 2018). Motor responses, such as pressing a button in response to a visual stimulus, also rely on cognitive evaluations, highlighting the interdependence of both processes.

Research by Miller and Cohen (2001) and Cisek and Kalaska (2010) demonstrates that the brain frequently processes cognitive functions like decision-making and planning in parallel with motor functions. This was also shown by Künstler et al. (2018), who investigated how performing a motor task impacts different aspects of visual processing. Under dual-task conditions, both visual processing speed and visual short-term memory storage capacity are reduced, while the perceptual threshold remains unaffected. This gives evidence that cognitive resources are shared between motor and visual tasks, leading to performance declines in visual processing in this experiment. Passingham and Toni (2001) demonstrate that distinct neural networks underlie motor and cognitive functions, while simultaneously interacting with each other. Notably, the ventral and dorsal streams of visual information processing exemplify this division of labor, though they may overlap in specific contexts (Goodale & Milner, 1992; Mishkin et al., 1983; Müsseler & Rieger, 2017).

As discussed in previous chapters, a task may require both cognitive and motor representations to achieve a specific goal. However, the degree to which cognitive or motor components are necessary can vary significantly depending on the task. In a simple-choice reaction time experiment, the cognitive aspect - processing the stimulus to decide whether to press a button - may be more crucial and influence task speed more than the execution of the motor command that results in the button press. In contrast, for a task involving the use of a joystick to track a line on a screen, while visual stimulus processing to determine the joystick's movement is important, it is likely that the precise execution of motor commands, which demands a high level of fine control, plays a more significant role in determining performance.

Fitt's (1954) Law demonstrates that factors such as speed, accuracy, and load significantly influence movement planning. Shadmehr and Krakauer (2008) suggest that three core challenges for movement planning must be addressed: system identification, where sensory outcomes of motor actions are predicted; state estimation, which involves integrating these predictions with actual sensory feedback to determine the current state of the body and its environment; and optimal control, which requires adjusting sensorimotor feedback loops to enhance movement quality. The overarching objective in solving these problems is to execute movements that achieve a desired, rewarding outcome.

Response execution involves a closed-loop system (Bernstein, 1967), where a movement is continuously re-evaluated and adjusted in real time. This, however, assumes that the movement is ongoing continuously and not yet completed (as in the case of pressing a button in response to a stimulus). From this, it could be inferred that motor tasks, or more specifically, those with a pronounced motor component, inherently have a continuous nature, thus allowing for closed-loop control. Information drawn in the effect monitoring phase is directly fed to the stimulus perception phase in the following loop running through the four processing stages. In contrast, this suggests that cognitive tasks tend to be more discrete. Information gathered during the effect monitoring phase can be discarded when the next stimulus appears.

It seems apparent, that motor and cognitive tasks are linked by their localization in neuronal processing. For the work presented here, a distinction is only made in terms of temporal interdependence and the different weighting of the requirements for specific processing stages. Cognitive tasks are further assumed to have a more discrete temporal character and require the highest capacity in perception and response selection processing stages. Motor tasks are assumed to be rather continuous in nature and pose the highest processing demands during response execution and effect monitoring stages. In line with this profiling of the two types of tasks, representatives of these two task classes are selected in the following experiments.

To be able to decode the structurally apparent connection further in its temporal processing, stimuli are needed that can influence the processing of motor or cognitive tasks at certain times during task processing. So-called probe tasks (Posner & Boies, 1971) are used for this purpose. However, since these are mostly presented as beep sounds, which presumably do not interact specifically with cognitive or motor tasks, the idea is to provide these probe stimuli with specific semantics to do so. Action observation studies have shown a linkage between motor and semantic processing areas during the processing of action-related verbs (De Marco et al., 2018; Neely et al., 2018; Rüschemeyer et al., 2007), so it seems reasonable to argue that semantic auditory stimuli with motor or cognitive characteristics are also able to specifically influence these types of tasks.

3.2 Semantic Probe Tasks and Their Induced Interference

Given that auditory stimuli are frequently included in multitasking studies, it is important to examine the extent to which separate or shared resources are utilized for processing different auditory stimuli, as well as to determine which other stimulus or task types are likely to experience interference due to a neuronal activity overlap. Especially since this type of stimulus is often used to specifically interfere with another primary task and thus provide information on whether the processing of the probe stimulus leads to a competition for capacity at specific processing stages.

In early structural bottleneck theories (Broadbent, 1958), it was assumed that semantic processing of incoming stimuli takes place at a later stage and therefore only for stimuli considered relevant to elicit a response. Even though Cherry's cocktail party phenomenon (Cherry, 1953) and Treisman's attenuator model (Treisman, 1964) have already shown that a certain amount of semantic processing of stimuli occurs immediately, even if these stimuli do not always elicit a response. By assuming that different brain regions access separate resources and the differentiation of physical and semantic characteristics of incoming auditory stimuli might be of importance, studies that investigate the localization of processing stimuli with specific semantics can be interpreted and provide insight into the question of specific task interference.

Kohler et al. (2002) could show that monkeys have neurons in their premotor cortex that can be classified as bimodal audiovisual mirror neurons that discharge to observed and executed actions but also to heard noises related to the action. It has been sufficiently proven that a mirror neuron system also exists in humans (Rizzolatti & Craighero, 2004). Mirror neurons also show a somatotopic organization (Buccino et al., 2001), just like the primary somatosensory (S1) and the primary motor (M1) cortex, which are responsible for the processing of sensory information and the conscious control of muscles (Kaas, 1997; Penfield & Boldrey, 1937; Schieber, 2001). In human cognition, the comprehension of action extends beyond direct observation or auditory cues, encompassing the interpretation of action-descriptive language. This observation hints at the potential reliance of action understanding on the observation-execution matching system, even in linguistic contexts, such as the description of actions through language (Tettamanti et al., 2005).

Rizzolatti and Arbib (1998) suggest that the development of human language may even have been dependent on the presence of mirror neurons. They propose that language initially emerged through a form of pantomimic proto-language, relying on gestures where action comprehension was crucial. Over time, this gestural communication is thought to have evolved into the auditory-based speech system, leaving a functional link between action observation and language areas in the central nervous system (Arbib, 2003). In multitasking studies, where stimuli are often presented in written form or auditorily, the nature of the stimulus and its semantic processing may lead to additional interference with motor tasks due to the involvement of overlapping brain structures. It is also worth considering whether this interference could extend to auditory stimuli with specific cognitive semantics like numbers, potentially causing interference with cognitive tasks, like arithmetic as well.
Rüschemeyer et al. (2007) visually provided participants with simple motor verbs (e.g., grasping) and abstract verbs (e.g., thinking) while in an fMRI scanner. They observed greater activation in areas such as the posterior premotor cortex, M1, S1, and the secondary somatosensory cortex (S2) when participants processed motor verbs compared to those with abstract meanings. Another study by Pulvermüller et al. (2005) applied transcranial magnetic stimulation (TMS) to motor areas in the left language-dominant hemisphere to right-handed subjects. They found that the stimulation influenced the processing of action-

related words. Stimulation of arm areas facilitates the processing of arm-related words and stimulation of leg areas facilitates the processing of leg-related words by resulting in faster responses with the corresponding hand or foot to auditorily presented words with corresponding semantics. The findings of both studies indicate a functional relationship between the processing of words with clear limb-associated motor semantics and the actual execution of corresponding movements, as the execution can be directly influenced. This connection is likely due to the involvement of motor regions, which are responsible for both movement planning and execution, being activated during the semantic processing of these words.

In Rüschemeyer's study, participants were solely tasked with viewing words while their brain activity was measured (Rüschemeyer et al., 2007). In Pulvermüller's study, the TMS was also directly related to the classification of the respective words (Pulvermüller et al., 2005). The question that now arises, however, is to what extent (involuntary) semantic processing of probe stimuli can intervene with the execution of a primary task. In a study by Neely et al. (2018), subjects performed an oddball task while listening to auditory interference signals such as a sinusoidal tone, a "stop" command, or a "press" command. The results showed that reaction times to the oddball task were significantly longer for the "stop" signal, suggesting that semantic processing of the command inhibited the execution of the button press. In contrast, the "press" signal did not affect reaction times when compared to a baseline, indicating no additional effects. This can be interpreted as inhibition of the primary task through targeted semantic stimuli but not a boost in performance through specific stimuli.

Currently, there are no studies showing that the specific semantics of auditory stimuli directly result in faster reaction times during primary task execution in single-choice reaction time tasks. However, it is well-established that congruent auditory stimuli - those semantically related to the task - can improve reaction times (i.e., right-hand key press to the word “right” and left-hand key press to the word “left” in an S-R-mapping). In contrast, incongruent stimuli tend to slow reaction times, indicating that semantic alignment between auditory stimuli and task demands facilitates more efficient cognitive processing (Laurienti et al., 2004; Simon et al., 1967). This suggests that congruent semantic information across modalities facilitates quicker cognitive processing and more efficient multisensory integration.

The processing of stimuli with motor-related semantics, whether auditory or visual, can directly impact the corresponding motor commands and the execution of movements, owing to the activation of overlapping brain regions. This prompts the question of whether similar interactions might occur in tasks of more cognitive nature. Given the wide range of cognitively focused tasks, it is useful to examine one that is associated with well-defined brain areas. Arithmetic task processing offers a suitable case for such observation.

The prefrontal cortex is important for online calculations generating immediate results, while the angular gyrus is responsible for mentally representing numbers as quantities and exact calculations (Dehaene, 2011, pp. 174 - 185). A case study involving a patient with a progressive neurodegenerative disease further revealed evidence of two distinct central processing pathways for numerical processing. Although the patient could recognize and comprehend numbers, they were unable to read or write numerical words, supporting the existence of a multi-route model for numerical processing (Cipolotti & Butterworth, 1995). Campbell (2009) postulates a dual-route model for naming Arabic digits. He suggests a semantic route that involves the quantity representation and an asemantic route responsible for direct retrieval from a visual input to a phonological output during number recognition tasks. The semantic route serves tasks like magnitude comparison, while the asemantic route is more engaged during arithmetic tasks. This interpretation aligns with the distinct brain regions identified. It is plausible that the prefrontal cortex is associated with the semantic processing pathway, while the angular gyrus may correspond to the asemantic processing route. Other brain regions that were identified for arithmetic processing are the intraparietal sulcus (IPS) for processing numerical magnitude and performing arithmetic operations, and the posterior superior temporal gyrus (pSTG) involved in processing numerical words and symbolic representations (Dehaene et al., 2004). Dehaene and Cohen (1997) found that rote verbal knowledge of arithmetic accountable for recalling memorized facts and number words, is associated with areas of the brain involved in language processing. Again, this raises the possibility that number words (spoken or read) may affect performance in simple, verbally memorized arithmetic tasks. Another experiment by Damian (2004) has demonstrated that processing speed varies between digits (e.g., 1, 2, 3) and number words (e.g., one, two, three). The format in which numerical information is presented (digit vs. word) influences both decision-making speed and accuracy. Reading number words likely involves more extensive semantic processing, which accounts for the additional time required (Noël et al., 1997).

Campbell and Metcalfe (2008) demonstrated that digit naming time when performed in the context of number-fact retrieval (multiplication of one-digit numbers) as a primary task was about 15 ms slower than in the context of a task requiring semantic processing (magnitude comparison). This context effect is larger on digit naming as the secondary task in comparison to naming incidental features, such as font color of the visually presented numbers (Campbell, 2009). This indicates that this effect is specific to number processing.

A study by Reynvoet et al. (2002) demonstrated a semantic priming effect. Speed and accuracy of number naming are increased when participants receive a semantically related number (3) preceding the to-be-named target number (4). The authors suggest that semantic relatedness between prime and target numbers enhances cognitive processing by activating semantic networks where numbers are connected based on their meaning and relationships.

Number processing further underlies cognitive control as was found in a task-switching paradigm (Schliephake et al., 2022). The two tasks involved either determining the parity of a number or comparing its magnitude to a specified reference. In switch trials, congruency effects between the tasks are anticipated to be larger due to persisting activation of the task set, causing an increased interference. Congruency in a switch trial would imply that in the current trial, where the number 4 is assessed for parity, a key press is required. In the previous trial, if the task was to press a button for numbers smaller than 8 (magnitude condition), both the current and previous tasks would necessitate a key press for the number displayed, thus making the tasks congruent. Contrary to these expectations, congruity effects were reduced in switch trials compared to repetition trials. An influence of cognitive control as an inhibition of previously abandoned tasks on basic number processing is suggested by the authors, so that less influence on current number processing demands is exerted.

In summary, an interfering auditory stimulus is significant in multitasking experiments not only because its physical properties divert attention from the primary task, but also because its specific semantics can directly affect task performance. The stimulus may either engage the same cognitive resource as the primary task, causing interference (as explained by the multiple resource model), or it may enhance arousal through a priming effect, potentially increasing performance by freeing up additional capacity for the task (as explained by capacity theory).

An auditory probe task in the times regime model would be assumed to require relatively little capacity in comparison to more complex cognitive or motor tasks during all processing steps (see Figure 11). That is, if the probe task is designed as a simple choice response task to an auditory presented stimulus. It can be expected that perception and response selection pose the highest capacity demands, where response selection does not necessarily only imply the selection of a button push as an adequate response, but also the semantic processing to some extent. The response execution itself, as it is a simple button press, would pose the least capacity demands. Effect monitoring could pose low to medium demands depending on the feedback about and the consequences for a missed response to a stimulus.

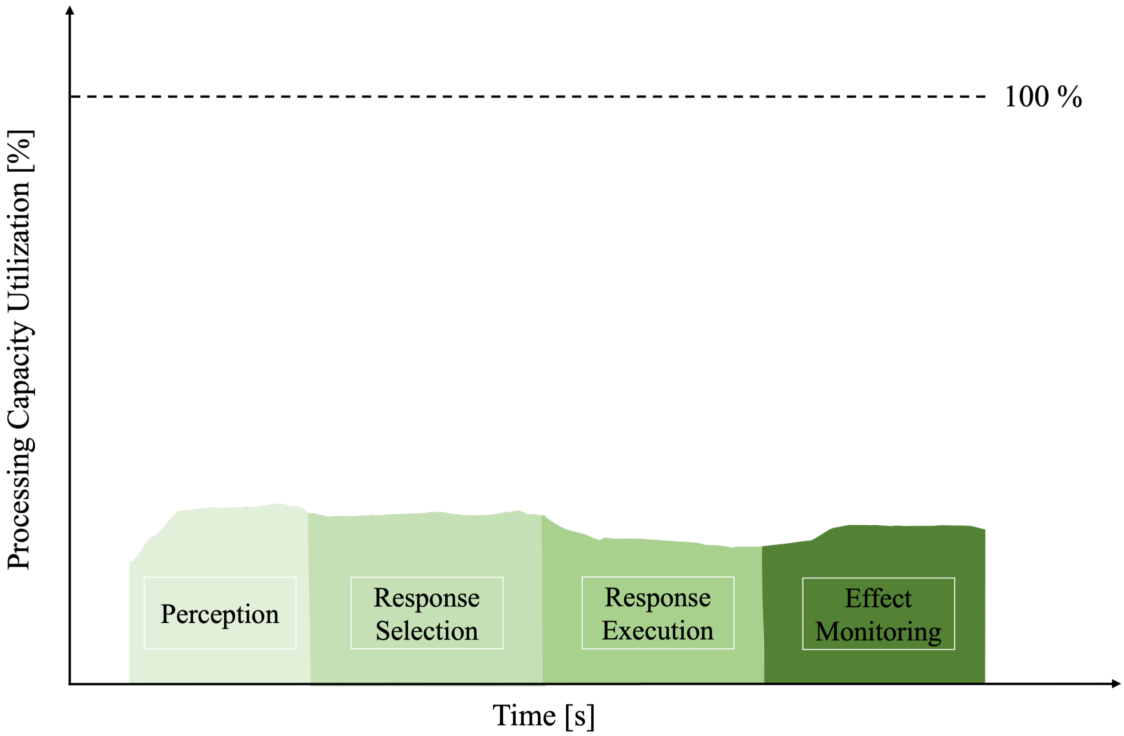

Figure 11: Capacity requirements for a semantic auditory single-choice probe task as represented by the area $Time_{Task}$ x PCU_{Task} during all four processing stages.

Three fundamentally different types of tasks have now been identified that are suitable for testing the time regimes model presented here. Cognitive tasks that are primarily discrete in nature and have particularly high capacity requirements for perception and response selection. Motor tasks that have continuous characteristics due to their high internal dependency and are particularly demanding in response execution and effect monitoring. And semantic auditory probe tasks, which are well-suited to cause specific interference with either the motor or the cognitive task. Based on the model and these three task types, three main research interests can now be identified for this thesis. These are explained in the next chapter.

4 Research Interest

Assuming that task type modulates the demands on processing capacity across different processing phases, this work examines the immediate adaptability of capacity allocation through three experimental investigations. The subsequent research objectives aim to thereby deepen comprehension of multi-tasking, with a specific focus on employing the developed time regimes model to categorize findings at a content level through tasks that allow for self-organized multi-tasking.

In the context of the time regimes model, the primary objective is to assess whether differentiating task performance based on time and errors can provide valuable insights into resource allocation. This involves further categorizing tasks into either a flexible or fixed time regime, enabling direct assumptions about how task execution can vary in time, which is expected to influence the error dimension when organizing multiple task executions. Broadly, the model posits that each task requires prior careful assessment of the requirements in both the PCU dimension (evaluating the significance of avoiding performance errors) and the time dimension (determining if the time allocation is fixed or flexible). This assessment is reiterated within each task for its distinct processing stages, where the capacity requirements of each phase for successful task execution must also be appraised. To substantiate the model in terms of content, three explicit research lines are outlined, aiming to derive conclusions regarding the model's conceptual framework.

To this end, the first line of research (R1) initially investigates the classical performance decrement, corresponding to the effect typically observed in cognitive psychology studies. Owing to the discrete nature of the tasks, the performance difference between single- and dual-task conditions can largely be assessed only at the level of task blocks or trials. However, since the individual processing stages within a task may impose distinct demands on processing capacity, it remains uncertain whether this approach adequately captures the flexibility of capacity allocation with sufficient precision. Consequently, the examination of the performance decrement effect serves as a control and reference point for the more detailed event-related analysis developed in the third line of research (R3). Performance parameters in time and PCU dimensions should always be superior (take less time for task execution and produce fewer errors) when performed under single-task conditions in comparison to multi-task (dual- and triple-task) conditions.

The second line of research (R2) investigates the semantic effect of the probe stimuli on a content-based interference. This analysis aims to determine whether, as postulated by the MRM, only modality-specific factors are relevant for task interference—based on the assumption of distinct resource pools—or whether additional task characteristics may also give rise to specific interference effects. Probe stimuli with motor semantics, such as action words or directions, should specifically impede performance in the motor force tracking task. The influence of semantic factors on cognitive tasks, such as the impact of spoken numbers, is hypothesized to specifically impair performance in tasks involving mental arithmetic.

To date, none of the existing models (bottleneck, capacity theory, MRM) sufficiently account for the precise mechanisms or generalizable principles underlying the allocation of processing capacity. Increasing the temporal resolution of interference analyses across different task combinations may provide valuable insights into this issue. To gain a closer insight into self-organized task interleaving by combining motor, cognitive, and probe tasks with one another, the third research line (R3) will focus on event-specific performance decrement during the combination of said tasks in dual- and triple-tasking scenarios. Performance measures along the time and accuracy dimension will not only be measured for a block-wise execution, but also during specific events of task combination during multi-task execution. These results will then be used to be classified based on the time regimes model and thus lead to new insights regarding the temporal entanglement in self-organized multi-tasking scenarios.

The next chapter will focus on the operationalization, namely the detailed description of the motor, cognitive and probe task and its quantification for the performance characteristics in the dimensions of time and accuracy.

5 Operationalization

The four phases of task processing that have been identified are: stimulus perception, response selection, response execution, and effect monitoring. For any executed task, all four phases are undergone, and the required attention for each phase in a task can be described by an area of $Time_{Task}$ x PCU_{Task}. Nearly every response to a stimulus requires some kind of motor action, but to which extent movement control determines the outcome of the response varies. Therefore, different tasks can be designed in derivation of the importance of different phases for task execution success. To generate new insights into the research objectives, tasks with different characteristics are designed. Tasks with different requirements in terms of PCU, for the identified phases of information processing, need to be thought of.

One main difference in task characteristics has proven to be the discrete or continuous task nature, as it requires different capacity allocation during the processing stages. To test the processing for a continuous task at the stimulus perception and response selection stage, a task should be chosen where those two phases pose the main challenge for successful task execution. Therefore, a mental cognitive subtraction task was chosen. The other continuous task should then rely mostly on response selection, execution, and effect monitoring. For this, a motor force tracking task was chosen. With this, time-based assumptions predicting a performance decrement can be examined. To be able to test for content-based task interference at discrete times during either the cognitive or the motor task, a probe reaction time task was designed, where the stimulus perception was the most crucial phase determining task success.

All three experiments presented here rely on the same task fundamentals, explained in this chapter. Note that for each task, the context is first explained, followed directly by the general description of task execution during the experiments. Also, the pre-processing of the recorded data per task in the experiments is given immediately, to impose a deeper understanding of each task characteristics. Variations of the tasks and other methods will be specifically described in the experiment where they are applied. Subsequently to each task description, an explanation of task combinations for the multiple task conditions will follow.

5.1 Experimental Setup

Throughout all three experiments, participants sat in a leg press (KÜNZLER Sport, Pluederhausen, Germany), focusing the projection plane on the wall that was between 3.2 and 3.5 m ahead of them, depending on the individually adjusted position of the chair from the leg press (see Figure 12). Visual stimuli (i.e., either ongoing motor performance feedback or the numbers for the cognitive task) were presented using a projector (EPSON EH-TW740, Meerbusch, Germany) onto an area measuring 1.2 x 1.9 m. Applied forces on the leg press during the motor task were measured using a 5 kN strain gauge (ME Systeme, Henningsdorf, Germany) and a Biovision (Wehrheim, Germany) amplifier. For recording manual responses to the cognitive task, a computer mouse was used (Dell Technologies, Round Rock, USA) and connected via USB port. For measuring manual responses to the probe reaction time task, a force sensing resistor (FSR X 400, Interlink Electronics, Los Angeles, USA) was used, measuring applied pressures between 0.3 and 50 N. The sensor was placed at the tip of the left index finger of the participant using double-sided adhesive tape. Spoken responses were recorded by a condenser microphone (LD SYSTEMS Adam Hall GmbH, Neu-Anspach, Germany) with a phantom power supply from an IMG Stageline EMA-1. The incoming signal was amplified by a low-noise amplifier (IMG Stageline MPA-102, Monocar International, Bremen, Germany). Using the data acquisition toolbox in MATLAB R2019b (MathWorks Inc., Massachusetts, USA) to control the NI-USB-6210 data acquisition box (National Instruments, Austin, USA), all response signals were recorded with a sampling rate of 1000 Hz, except for the computer mouse. The sampling rate for this is derived from the monitor refresh rate at which the visual stimuli were displayed, which is 60 Hz.

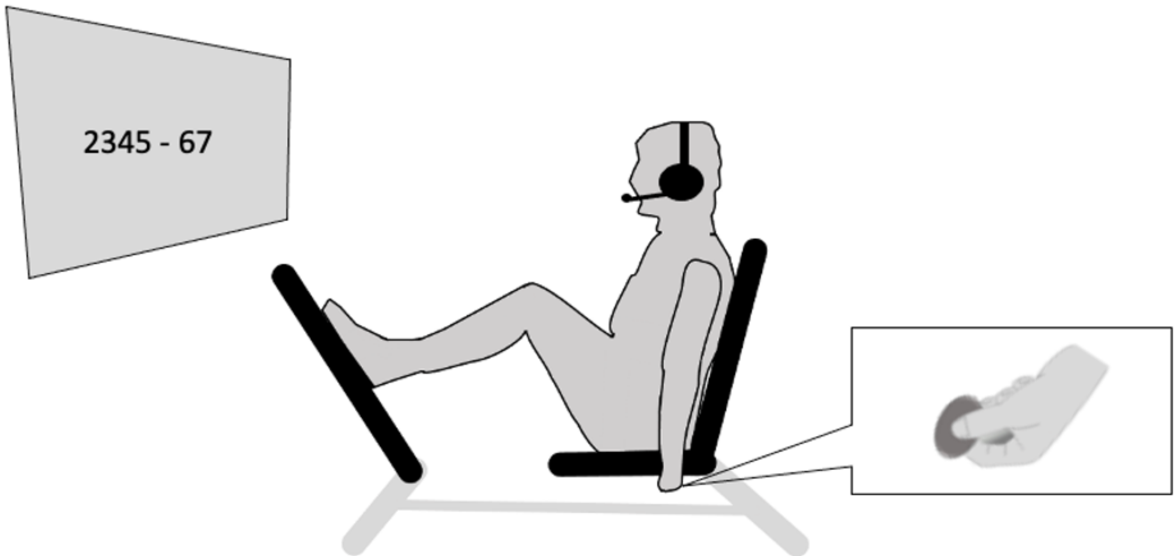

Figure 12: Experimental setup for experiments 2 & 3. For the first experiment, the microphone was replaced with a computer mouse that was held in the right hand

The entire experiments were programmed with the open source software Psychtoolbox running in MATLAB R2019b (MathWorks Inc, Massachusetts). Using an ASIO sound card (ASIO4ALL v2, Steinberg Media Technologies, Hamburg, Germany), probe stimuli were presented through noise-reducing over-ear monitoring headphones (Sennheiser HD 200 PRO, Wedemark-Wennebostel, Germany).

5.2 Probe Reaction Time Task

Probe reaction time paradigms have been widely used to examine attentional processes under multiple task conditions for a long time in cognitive psychology. In early experiments by Welch (1898), auditory probe stimuli were used to measure attentional demands of a primary task by interference with the secondary probe reaction time task. For primary motor tasks, Ells (1969) found that the probe had little effect on its performance, but that the reaction time to the probe stimuli reflected the central processing demands for the motor task very sensitively by being prolonged.

Therefore, it can be expected that the probe task as used in the experiments of this work might not lead to a measurable performance decrement in the other two tasks but will result in increased probe reaction times under multiple task execution in comparison to single task execution.

Auditory probe stimuli have the highest capacity utilization in the stimulus perception phase. If a stimulus is perceived, the execution of the simple manual reaction should not require much capacity due to its low complexity in comparison to other tasks requiring more complex motor skill solutions. Thus, those auditory stimuli can be used to probe time-based interference as capacity utilization for another ongoing task. Especially, in probe stimuli that are presented at critical moments of processing other tasks, reaction times to stimuli should be prolonged (McLeod, 1978). Furthermore, probe stimuli that are semantically coupled with the tasks can be expected to show larger interference with task performance than uncoupled probes in accordance with the context-based interference (Wickens, 2002). It has been proven that by observing others' actions, the mirror-neuron system is activated, which builds the observation-execution system that plays a role in understanding others' actions (Gallese et al., 1996; Grafton et al., 1996; Rizzolatti, Fadiga, Gallese, et al., 1996; Rizzolatti, Fadiga, Matelli, et al., 1996). As already reported in chapter 3.2, Tettamanti et al. (2005) tested whether this system also responds to spoken action-related sentences. They found that for sentences describing actions with the mouth, hands, or legs, the left fronto-parieto-temporal network, including the Broca area, was activated. These are areas of the premotor cortex, where the described actions are also motorically coded, due to the somatotopic organization of the premotor-parietal circuits (Buccino et al., 2001). The authors see this as evidence that action-related sentences also lead to an activation of the observation-execution system. This could mean that this activation, through listening to probe stimuli with specific action-related semantics, could also lead to a specific capacity utilization. Taking this consideration into account, an auditory probe reaction time task was designed, with varying semantic content of the probe stimuli. The stimuli were either unspecific (beep sound) or aimed at specifically interfering with the performance in the motor or the cognitive task. The latter stimuli will be introduced individually for each experiment at the appropriate text passage.

Auditory stimuli were presented over headphones. Participants then had to perceive the stimulus, select a manual button press as a response, and execute this button press. Task performance can then be measured as response time or as response error, if no button press was executed within two seconds after stimulus presentation. Subsequently, the button press would be monitored in an effect monitoring phase (see Figure 13).

Task requirements throughout the four processing stages are as follows. *Perception*: For this task, the perception stage has the highest requirement in PCU, because stimuli were presented at discrete times during multiple task execution and did not require continuous attention. Thereby, attention must always be redirected if a stimulus emerges. *Response Selection*: This stage does not require much capacity due to the single-choice character of the task. Each auditorily presented probe stimulus requires a button press. *Response Execution*: Executing a manual button press does not pose high motor requirements. Thereby, this stage also takes up a relatively small area in the TR model. *Effect Monitoring*: Monitoring if the button was pressed should also not impose high requirements. Since no external feedback was given on whether a reaction was registered, effect monitoring was not explicitly evoked.

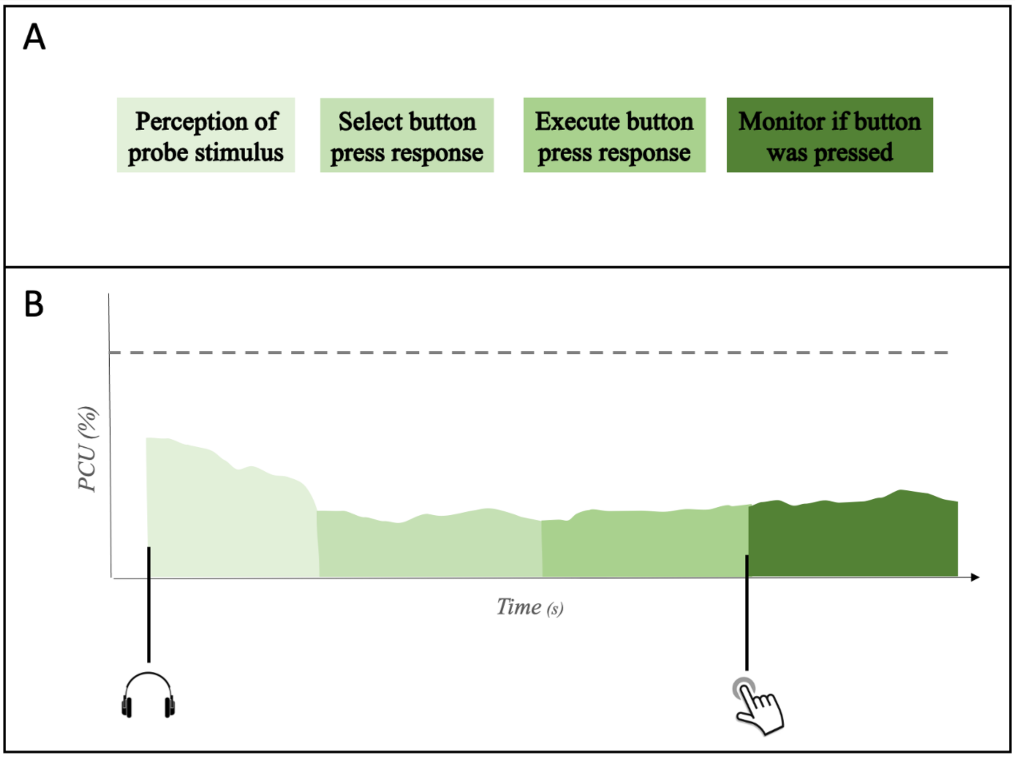

Figure 13: Processing stages (A) and their requirements in $Time_{PRT}$ x PCU_{PRT} dimensions (B) for the probe reaction time task. The headphone resembles the stimulus presentation; the hand resembles the recorded response via FSR.

The probe reaction time task (PRT) is a task with mixed time regimes. Because response times to stimuli allow for an elongation, it allows for flexibility. From a rigorous psychological perspective, the response to the probe stimulus should ideally occur within one second, as delays beyond this point would disqualify the task as a reaction time measure. However, this criterion was not strictly adhered to in the present experiments (as further explained in chapter 5.2.3). Still, if a button press response was not given within two seconds after stimulus presentation, an erroneous response was registered. This results in a somewhat fixed time regime characterization, given that this two-second threshold was the available time frame given for response execution.

While every experiment had different semantics for the stimuli, the response registration remained the same throughout all three experiments. To measure responses to the stimuli, a force sensing resistor (FSR) was used that was applied to the participant's left hand index finger using double-sided cohesive tape. As soon as a stimulus was presented over the headphones, participants had to apply force to the sensor in a pinch grip.

5.2.1 Probe Reaction Time Task throughout a Task Block

To adequately represent the tasks used in the experiments, a figure type is introduced below (Figure 14), which will later allow the multi-task conditions to be visually understood in terms of their temporal interdependence. The x-axis represents the time over a task block. The x-y plane holds the colored processing phases characterized for each task individually. The x-z-plane holds the presented stimuli starting at z=0 and registered responses at z=max. The y-z plane represents the visual stimulus presented ahead of the participants on a screen wall.

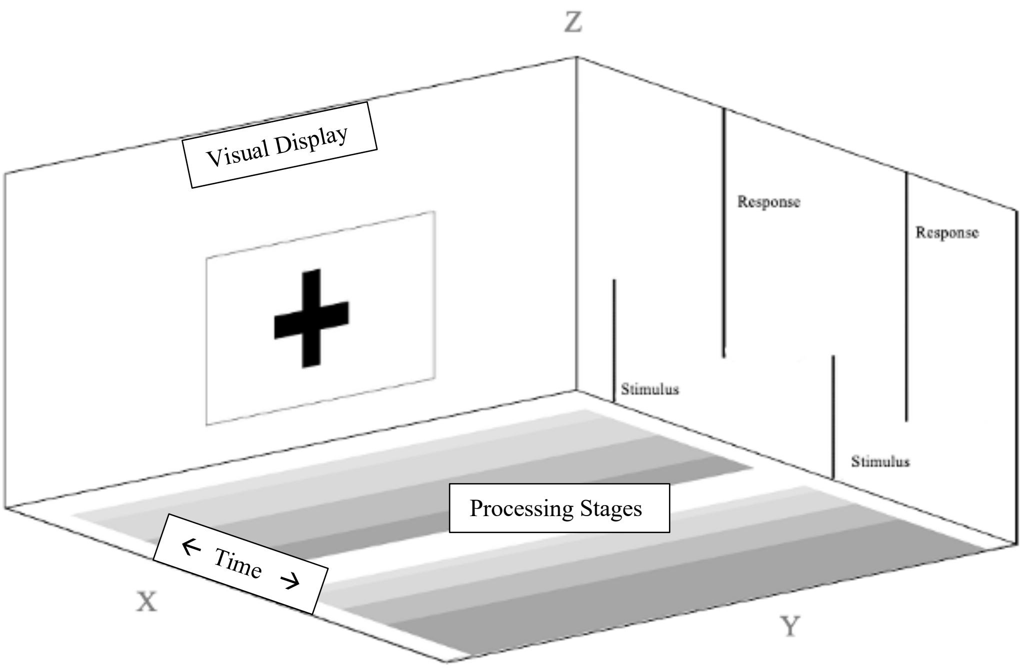

Figure 14: Figure type used for visualizing the tasks used in the experiments throughout a task block. Tasks are shown over the duration of a task block (x-axis) with respective stimulus and response onsets(x-z-plane), processing stages (x-y-plane), and display of information on the screen ahead of participants(y-z-plane).

The auditory probe reaction time task has a discrete nature, with stimulus presentation at different times throughout a task block of 60 seconds with varying inter-stimulus intervals to prevent predictability (see Figure 15). For the first habituation trials for this task, a total of 6 stimuli were presented throughout a block. The inter-stimulus intervals were pseudorandomized. For the task blocks that included the probe task in a multiple task setting, the inter stimulus interval was oriented at calculation times (see chapter 5.3) for the cognitive task to enable a stricter control of task interference.

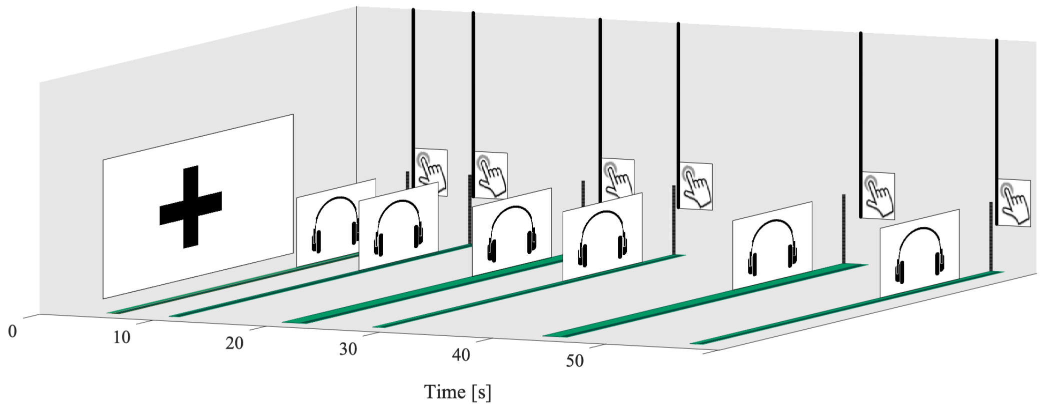

Figure 15: Probe reaction time task during a 60 second task block as single task execution. Headphone represents the presentation of an auditory stimulus; hand symbol represents a registered manual response to the FSR. Here, a fixation cross was presented in the PRT task over 60 seconds.

Performance for PRT in the time dimension of the TR model can be measured as reaction time. That is the time from probe stimulus onset to the recorded manual response. Performance in the PCU dimension can be measured as the reaction error. Errors can be made by not responding to a stimulus with a manual response before the next probe stimulus is presented, or within the defined time frame of two seconds. Another form of error can occur, if participants give a response without the presentation of a stimulus (false alarm).

5.2.2 Data Pre-Processing for the Reaction Time Task

The FSR sensor measures a voltage signal that has the characteristic of an inverted exponential function. For smaller applied forces, the FSR functions as a force sensor. When applying high forces, the curve approaches saturation, resulting in only minimal or no resistance changes, resembling a pressure sensor (Interlink Electronics, 2023). In the three experiments, the FSR is used for the detection of a reaction onset after probe stimulus presentation. The latter takes place at lower forces. Since the absolute applied force (i.e., the maximum of the curve, or area under the curve) is not of interest, the characteristics of the sensor are suitable. To develop an algorithm for the onset detection, transformation of the voltage signal into force data in Newton was applied. For this, a calibration was done for every FSR sensor used in the experiments. Only malfunctioning sensors were replaced. Transformation functions were always exponential functions following the characteristics of the sensor. They were fit to the data using a least squared error method. An example for one of the transformation functions is given below:

$$N = 0.306 * e^{0.321*V} + 0.001 * e^{2.194*V}$$

The amplitude of the original signal was thereby altered, resulting in steeper peaks but not in a timely shift of the peaks, which was a prerequisite to enable valid onset detection of the participant's reaction to probe stimuli (see Figure 16).

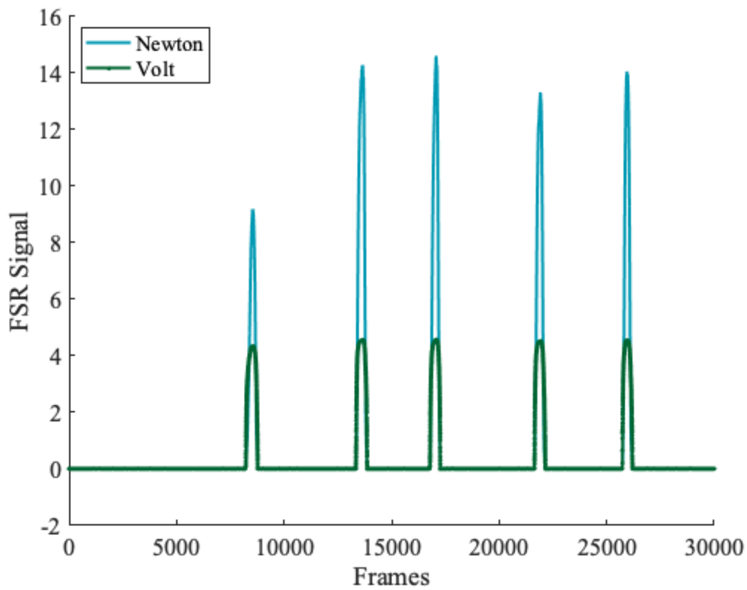

Figure 16: Signal of the force-sensing-resistor over a trial block of 30 seconds collected at 1000 Hz. Raw data was collected in Volt (green) and then transformed into Newton (blue) through a calibration process.

Raw data in Newton from the FSR consists of signal and noise. To enable the calculation of reaction times, the noise should be eliminated to enable accurate onset detection. For this, filtering methods can remove the noise if the frequency spectrum of both noise and signal can be distinguished. According to Winter et al. (1974), human motion can only take place at low frequencies. In their study measuring movement kinematics, they used a low pass filter with a 6 Hz cutoff frequency as proposed by Gold and Radar (1969, pp. 48–97). To gain a deeper insight into the signal, a Fast Fourier Transformation (FFT) can be calculated to dissect the signal into underlying frequencies. To enable the differentiation between signal and noise frequencies, a FFT is calculated for the first 200 frames (Figure 17 A) of a measurement, when the sensor was at rest. Another FFT is performed over the whole 3000 frames (Figure 17 B), where a total of five reaction onsets had occured.

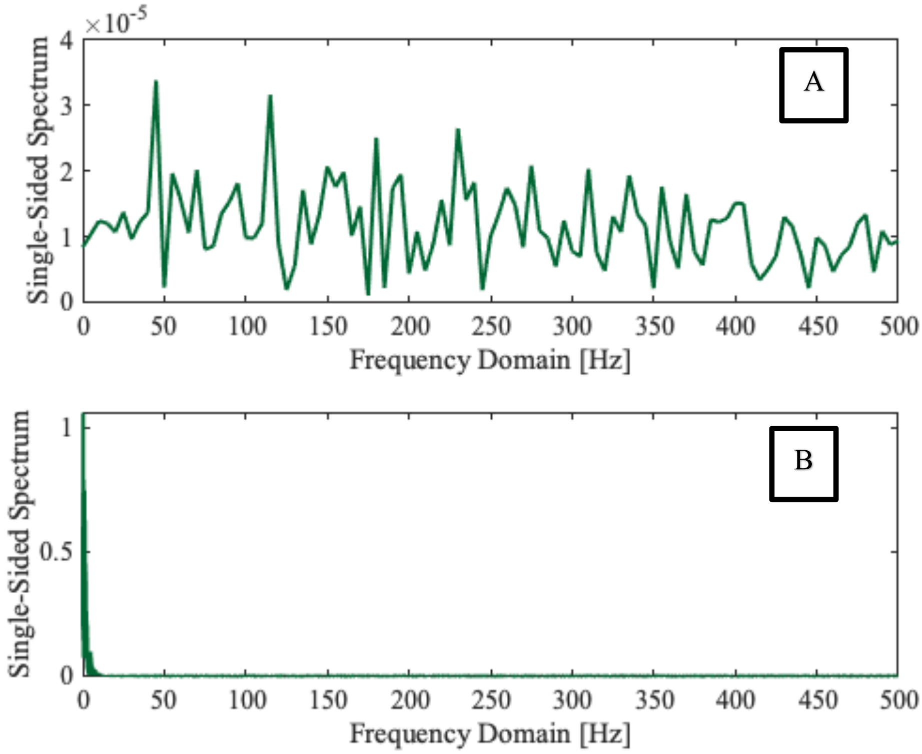

Figure 17: Fast Fourier Transformation for the FSR signal.
A: FFT for the first 200 frames where the sensor was at rest.
B: FFT for the whole measurement of 3000 frames with a total of 5 onsets.
Note that the y-axis limits differ to properly depict the single-sided spectrum.

For the first 200 frames without any reaction onset, the frequency domain is equally distributed from 0 to 500 Hz. Regarding the signal over 3000 frames, where the five onsets depicted in Figure 16 are included, the single-sided spectrum is much higher at lower frequencies. This was not visible when the sensor was at rest, leading to the conclusion, that those low frequencies must be signal rather than noise. To investigate those low frequent signal, Figure 18 depicts the frequency domain for the 3000 frames only from 0 to 10 Hz.

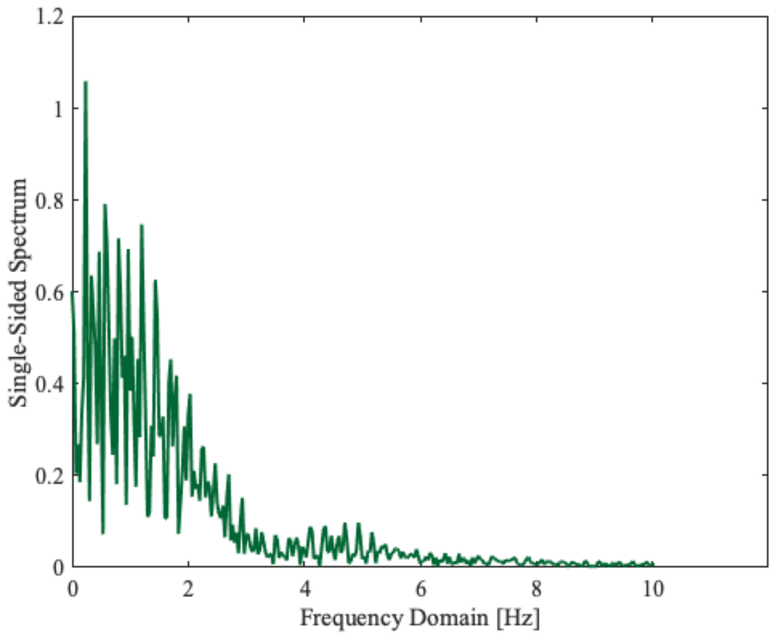

Figure 18: Fast Fourier Transformation (FFT) for the FSR signal over 30000 frames. Depicted is only the frequency domain between 0-10 Hz.

It is visible that most of the signal spectrum lies beneath 10 Hz, which is also in accordance with Winter et al. (1974). In the next step, filtering frequencies of six, eight, and ten Hertz have been applied to the data to test for any deviance in signal quality.

Visually, no differences were observed. No temporal shift was detected along the x-axis during the participants' reaction onset, which is a necessary condition for the valid calculation of reaction times. Moreover, lowpass cutoff frequencies of 2 and 30 Hertz were tested (see GitHub Repository[1]) to further verify that a cutoff frequency of approximately 10 Hz was appropriate for the data. While the 30 Hz cutoff did not alter signal characteristics, the 2 Hz cutoff resulted in a reduced slope of the peak and the signal dropped below zero. This is not a valid measurement, as the data had already been converted to Newton values, and negative force values are not plausible since a value of zero represents the absolute minimum. This analysis led to the conclusion that a cutoff frequency greater thatn 2 Hz up to 10 Hz is appropriate for the FSR signal. Consequently, a 10 Hz cutoff was consistently applied to all FSR data using a second-order Butterworth filter. This frequency was selected despite the potential suitability of lower frequencies, such as 6 and 8 Hz. However, since reactions to stimuli are characterized by very quick movements, a higher cutoff in comparison to kinematic data as in Winter at al. (1974) seems adequate in order to not filter any response behavior signal from the data.

---

[1] https://github.com/JelenaMueller/DynamicCapacityAllocation/tree/main/MethodologicalValidation/FilterAndSmoothFsrData

5.2.3 Determination of Reaction Onset and Calculation of Reaction Times

Data from the FSR signal has been transformed from V to N and filtered at a lowpass cutoff frequency of 10 Hz. Now, participant's reaction onset must be determined to enable the calculation of reaction times. Therefore, an algorithm must be developed that validly detects this onset in the pre-processed FSR signal. Mattes (2001) proposed a static threshold of 0.5 N as a criterion for this. However, if this threshold is used for the three experiments, onset detection is delayed, as it would detect an onset, when the slope of the signal had already increased (see Figure 19). Thus, a data-driven approach might be more suitable to reliably detect valid reaction onsets for the calculation of reaction times. For this, a similar approach as described by Brand et al. (2022) is used and depicted in Figure 19. To detect all reaction onsets throughout one task block, the local maxima of the FSR signal are first sought. These are the peaks, when the reaction is already ongoing. From there, the algorithm runs back along the x-axis to detect the onset in the signal.

In this context, the interval extends from a current local maximum to a point located halfway between this maximum and the preceding local maximum (or zero in the case of the first peak). In this interval, the slope is calculated, with the aim of identifying an onset where the slope approaches zero in one frame and subsequently increases in the next frame. This pattern suggests that the rest state (characterized by a zero slope) is disrupted, indicated by a high slope as pressure is applied to the sensor. The analysis entails identifying the last frame within the interval where the slope is less than 0.001 N/ms. For this particular frame, the applied force on the sensor (y-value) is evaluated. If this force is less than the static threshold defined by Mattes (2001), the frame is recorded as the reaction onset. If the force exceeds the threshold, the interval is adjusted to end at the frame immediately preceding this detected low-slope frame. This procedure is iteratively applied until a frame is identified where both a sufficiently small slope and the corresponding force value meet the criteria for reaction onset.

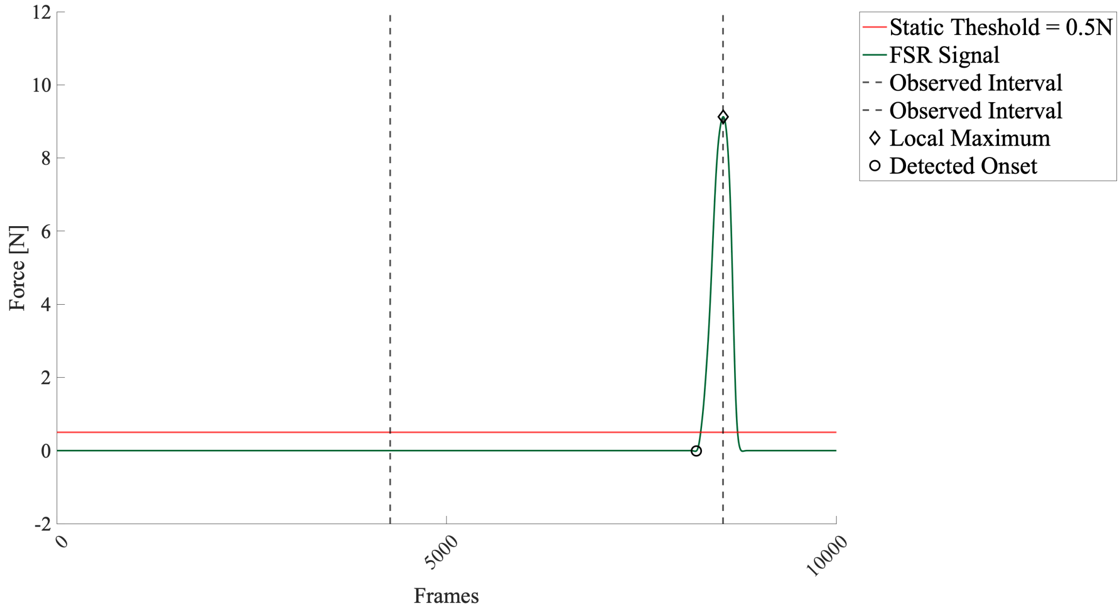

Figure 19: Depiction of the algorithm used for onset detection of FSR signal. First response to stimulus is displayed from a pilot trial. First peak is found as local maximum to determine the interval for the algorithm. An iterative procedure identifies the frame, where the slope is sufficiently small (prior to a sudden slope increase) and the corresponding force value lies below the static threshold of 0.5 N.

To determine reaction time as a dependent variable from the reaction onsets, an algorithm analyzes all probe stimulus onsets to determine whether a reaction onset occurs in the interval between a given stimulus onset and the subsequent one. The reaction time is then calculated by subtracting the probe stimulus onset time from the reaction onset time.

To determine which reaction times will be taken into further statistical analysis, the mean reaction time was calculated over all three experiments and conditions. The mean of all registered sensor onsets following one probe stimulus onset and previous to the subsequent stimulus occurrence was 614.515 (± 402.142) ms. As can be derived from the overall distribution of reaction times in Figure 20. Response times above 2000 ms are included in the unimodal distribution with the typical skewed profile describing response behavior, even though they would not be classified as such in psychological research. Therefore, reaction times up to 2000 ms still reflect valid measurements of the time needed for task processing[2]. All reaction times slower than 2000 ms were excluded resulting in the loss of 1.58 % of reaction time data.

[2] Although the term “reaction time” should thereby be updated to "response time," the original term will be used hereafter to ensure clearer distinction from the response behavior associated with the other tasks presented later in this work.

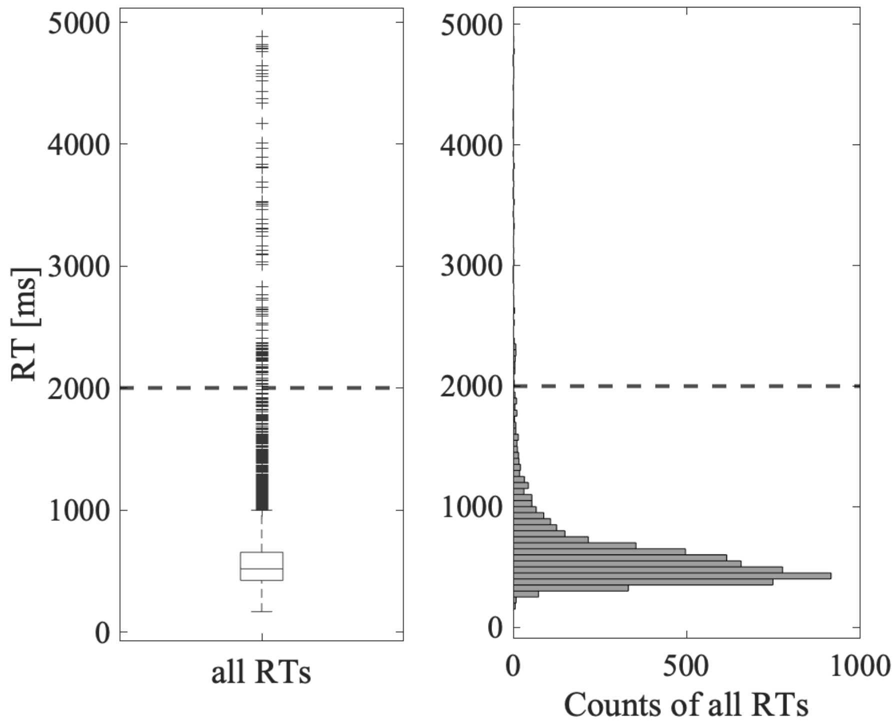

Figure 20: Boxplot (left) and histogram (right) of reaction times (RT) throughout all three experiments. The dashed line represents the cut-off of 2000 ms above which reactions were excluded from further analysis.

5.3 Cognitive Task

The cognitive task was designed as a complex continuous task that foremost strains the cognitive structures. The task is required to have critical processing demands at the stimulus processing and response selection stage for successful task execution. Therefore, a mental arithmetic task was chosen, where two-digit numbers had to be subtracted from four-digit numbers. The task was divided into two phases: a calculation and a result comparison phase. While in the calculation phase, solution latency alone (time needed from task presentation to a mentally calculated result) would provide a measure of arithmetic efficiency, it provides no information about the demands of each component process that is involved in calculations (Kaye et al., 1989). Therefore, by presenting solutions to the calculation tasks digit-by-digit after the participants had finished calculating a result and asking them to compare their result with the shown solution, an extended memorization of the calculated result is required. This allows for a more fine-grained estimation of processing demands of 1) calculation encoding and computation and 2) stated answer encoding, comparison, decision, and response stages.

Each subtraction task was divided into two distinct phases, as illustrated in Figure 21. During the calculation phase (highlighted in yellow), participants were required to mentally compute the solution to a visually presented subtraction problem. Once they arrived at a solution, they responded either by clicking a computer mouse with their right hand (experiment 1) or by verbally producing a 'pa' sound into a headmounted microphone (experiments 2 and 3). Following the response, the task disappeared from the screen, and the result comparison phase (highlighted in orange) commenced. In this phase, a computer-generated result was displayed digit-by-digit, starting from the unit position (rightmost) and proceeding to the thousands position (leftmost), with each digit appearing at 500 ms intervals. The full result was visible after 1500 ms and remained on the screen for an additional 1000 ms.

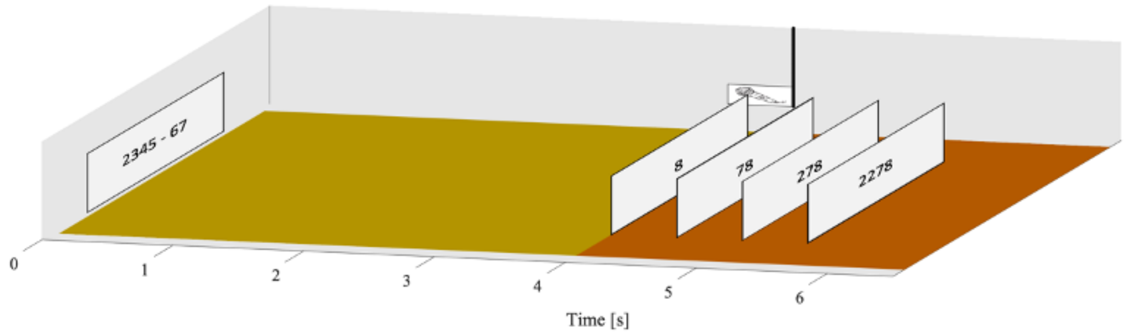

Figure 21: Exemplary subtraction task for the calculation task. In the calculation phase (yellow), the subtraction task was visually presented. After a verbal response from the participant after four seconds, the result comparison phase (orange) started, where a new digit of a presented solution was presented every 500 ms. The complete presented result remained on screen for another 1000 ms before the next task began.

In 50% of the trials, the presented result was correct, while in the remaining 50%, one of the four digits was altered by an algorithm, replacing it with a random number between one and nine. Participants were instructed to compare the displayed result with their own calculated solution and to respond as quickly as possible upon detecting an incorrect digit. The result comparison phase ended either when the participant clicked the mouse (experiment 1) or vocalized the ‘pa’ sound into the microphone (experiment 2) upon detecting an erroneous digit, or after 2500 ms regardless of the response (experiment 3). In cases where no response was provided, the phase concluded after 2500 ms in all experiments, followed by the presentation of the next subtraction task in the calculation phase.

Given the distinct characteristics of the task’s two phases, the time regime model must be applied separately to each phase. The calculation phase operates under a flexible time regime, as participants can extend this phase at their discretion. The phase concludes when the participants provide their response. Consequently, these are the resulting task requirements throughout the four processing stages for the calculation phase:

Perception: Due to the continuous character of the task over a 60 second task block, the perception of the visually presented stimuli does not have very high requirements because stimulus presentation is expected. *Response Selection:* This stage has the highest PCU as the actual calculation process is undergone. The digits must be mentally subtracted; a result must be formed in mind and the response indicating that the calculation is finished must be prepared. *Response Execution:* Once the result was calculated, the actual response (either verbally or through a mouse click) was carried out. Since these responses do not have high requirements on the motor system, the response execution stage does not take up a large area in the time regimes model. *Effect Monitoring*: The effect monitoring has no high requirements. Monitoring if the response was recognized and thus if the result comparison phase has been started is necessary, but since there is immediate feedback when the second phase starts, the monitoring should not require much attention.

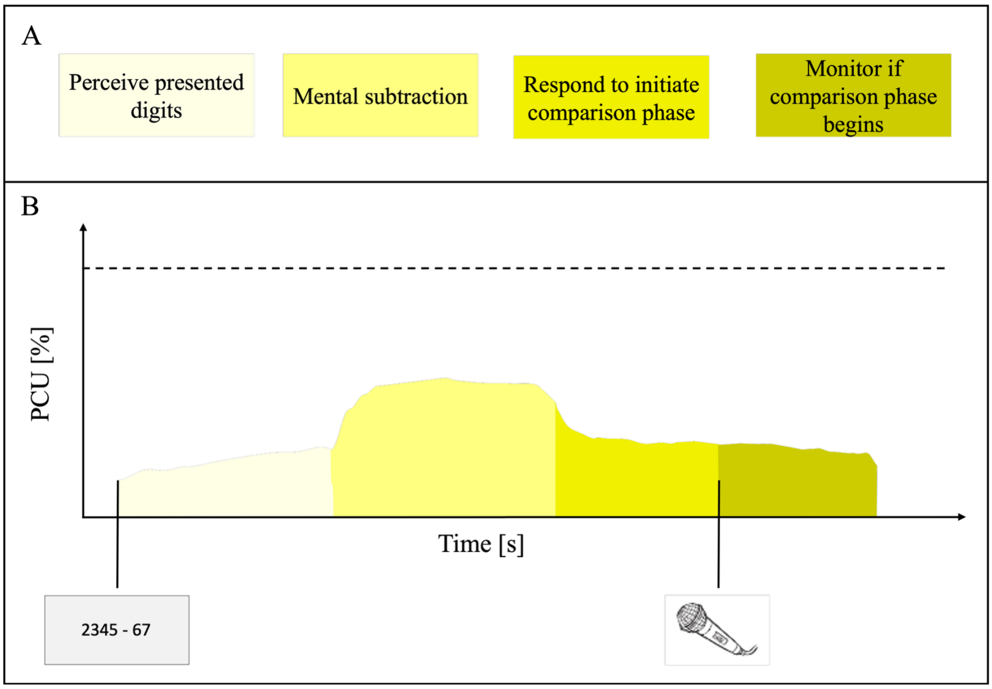

Figure 22: Calculation phase of the mental subtraction task.

(A) Different processing stages for the calculation phase of CLC.

(B) Modeled requirements of each stage in PCU and the Time dimension.

Performance can only be measured in the time dimension of the TR model and is measured as calculation time. This is the time passed from the visual stimulus onset of a new subtraction task to the measured vocal response (or clicked mouse in experiment 1).

The result comparison phase has a fixed time regime because the maximum time spent in this phase is limited to 2500 ms. When this time has passed, the next subtraction task will automatically appear in the calculation phase. With the presentation of each new digit, all stages must be undergone anew. Thereby, each stimulus (a new digit presented on screen) has a fixed time regime of 500 ms itself until the next digit is visible. After the presentation of the fourth digit, another 500ms remain for the fixed time regime of the result comparison phase.

Capacity requirements during the result comparison phase throughout the four processing steps are as follows. *Perception:* The presentation of a new digit is expected, and therefore, stimulus perception requirements are minor. *Response Selection:* In this stage, the presented stimulus must be mentally compared to the result that is kept in mind. By comparing the digits, the participant must decide whether the presented digit belongs to the correct result, which would then result in the inhibition of programming a response to the shown number. In case of an incorrect digit that does not belong to the correct result of the subtraction task, a response to that digit must be programmed. *Response Execution:* Executing or inhibiting the response might now pose higher requirements in comparison to PRT and the calculation phase because information processing theories suggest that with a higher number of alternatives, response times increase. And in this case, the alternatives are to either give a response or to inhibit it. *Effect Monitoring:* In this stage, participants need to monitor if their decision to respond or inhibit a response was correct. In case of an inhibition, no feedback is given. If a response to a wrong digit is given, feedback of the registered response was provided only in the first experiment by ending the result comparison phase and continuing with the next subtraction task. In the other experiments, there was also no feedback for a given response during the effect monitoring phase. Thus, response monitoring or inhibition effects were not possible through external feedback. Monitoring through internal feedback would be conceivable and could pose medium capacity requirements.

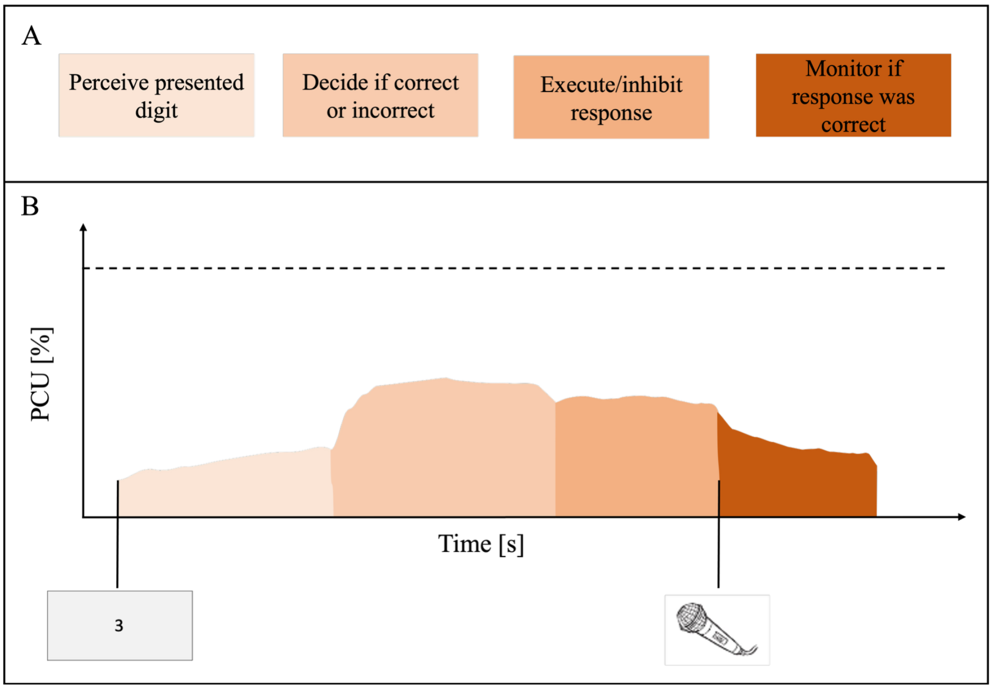

Figure 23: Result comparison phase of the mental subtraction task.
(A)Processing stages for result comparison phase of CLC.
(B) Modeled requirements of each stage in PCU and the Time dimension.

Performance in the result comparison phase is only measured in the PCU dimension. Since the time regime for responding to an erroneous digit is strictly fixed and no response must be given if the result is presented correctly. Therefore, the time dimension for this phase of the task is negligible in terms of a performance measure because a late response automatically results in an error. An error occurs when a response to a correctly presented result is given, or when no response to an incorrectly presented result is given. A response to an incorrectly presented number must always be given after this number becomes visible and within the time remaining in the comparison phase. The time interval between the presentation of the incorrect number and the corresponding response is irrelevant.

5.3.1 Calculation Task throughout a Task Block

With the beginning of a task block, the first calculation task is presented visually on screen, as depicted in Figure 24 in the y-z plane. For 60 seconds (represented on the x-axis), the instructions were to calculate as many tasks with as few errors as possible. A verbal (or manual in experiment 1) response to the calculation phase was always necessary to start the result comparison phase (see x-z-plane with microphone depicted left from response onset). A second response (microphone right from response onset depiction) was only necessary for subtraction tasks depicting an erroneous digit in the comparison phase.

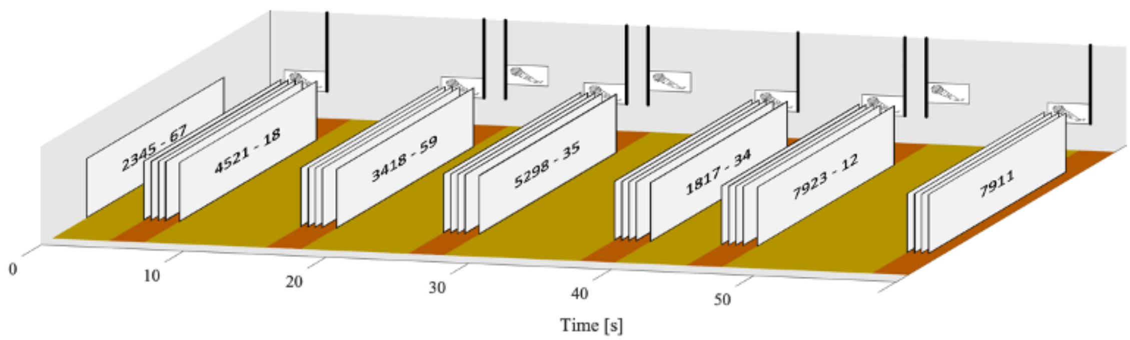

Figure 24: Task block of 60 seconds for the calculation task under single task execution. A total of six tasks is presented here with calculation phase (yellow) and result comparison phase (orange).

Performance can be measured as calculation time in the calculation phase for the time dimension and calculation error in the result comparison phase for the PCU dimension in the time regimes model. For the content-based interference, semantic probe stimuli that hold numbers should specifically interfere with the calculation task in both phases, leading to increased calculation times and errors.

5.3.2 Data Pre-Processing for the Cognitive Task

The detected mouse clicks from experiment 1 did not require pre-processing, as they were received as times in seconds (with two decimal places) throughout the experiment and saved as such. The data from the microphone also went through no other offline pre-processing, as the online voltage signal from the microphone was used to detect a verbal answer onset through a static threshold that was determined individually for every participant. The threshold was adjusted so that onsets were detected at the participant's preferred speaking loudness, without detecting any other noises, like breathing. Detected onsets were saved and used for the determination of calculation times for the calculation phase and the determination of error rates for the result comparison phase.

The detected microphone onsets were saved from the online data during the experiment. Data was reviewed with an algorithm to ensure that trials where the microphone was triggered accidentally by participants through loud noises like clearing their throat or coughing were excluded from further statistical analysis. It was not possible to completely differentiate between the given answers ('pa' sounds) and any noise due to their resemblance in the audio signal. By writing protocols during data capture, the resulting errors were corrected to the best knowledge. For the determination of calculation time, the first detected mouse or microphone onset after a new calculation task presentation was sought. For this, a dead zone of 200 ms was established around task stimulus onset because it can be expected that the calculation of a result to the shown task takes longer than the span one would expect from single-choice reaction times. Therefore, during this dead zone, no answer onset of the participants would be expected, and if an onset was saved in the data during this zone, this complete subtraction task was deleted from further analysis.

To assess valid subtraction tasks for the calculation of the error rates, another dead zone was established after the response onset in the calculation phase. This onset would automatically trigger the beginning of the result comparison phase, where the calculation errors would be derived from. For this phase, it also seems likely that any onset detected earlier than 200 ms after entering the result comparison phase must be due to other recorded noises made by the participants that were not meant to execute the task. Therefore, these trials were also excluded from further data analysis. For the remaining tasks, an algorithm determined if the participants reacted to an incorrectly shown result with a mouse click or 'pa' sound and to a correctly shown result with no such reaction. This trial-wise binary data (correct answer or error) can then be used to calculate the blockwise error rate in percent over each task block of 60 seconds.

To determine which values of the dependent variables, calculation time, and calculation error should be included in the further statistical analysis and which can be identified as invalid data or outliers, the distribution over all three experiments for the calculation time and the calculation error is displayed in Figure 25 and Figure 26.

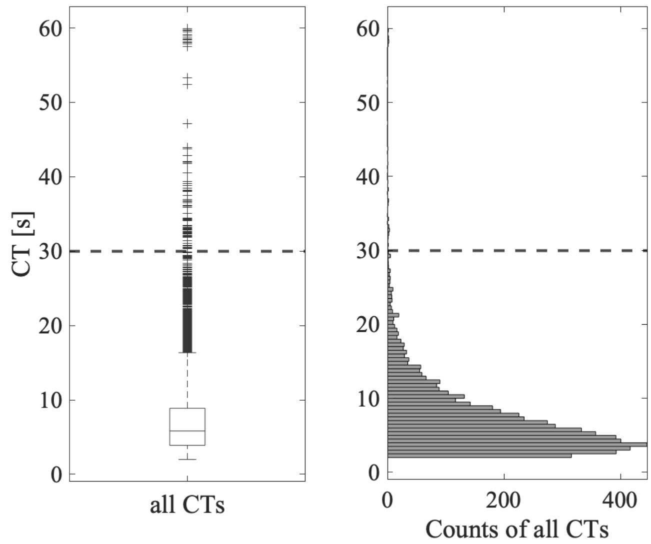

Figure 25: Boxplot and histogram of the distribution of calculation time (CT) for the cognitive task over all 3 experiments.

All calculation times shorter than 2 seconds were excluded as a quicker calculation does not seem plausible but can be attributed to a false onset detection by the microphone. It can be expected that answering the calculation task will take longer than answering the probe reaction time task, for which the upper cutoff was determined to be 2 seconds. Even for subtractions without any burrows, task processing and motor answer execution should take longer than answering a single-choice-reaction time probe task. For the upper cutoff, calculation times beyond 30 seconds were excluded from further analysis, resulting in a data loss of 0.448 %. Even though this criterion might seem very lose, for the Triple Task condition it was sometimes the case that only two calculation tasks would be executed during one block which would indicate that the calculation time per task was 30 seconds (under the assumption that the time taken for both tasks is equally long).

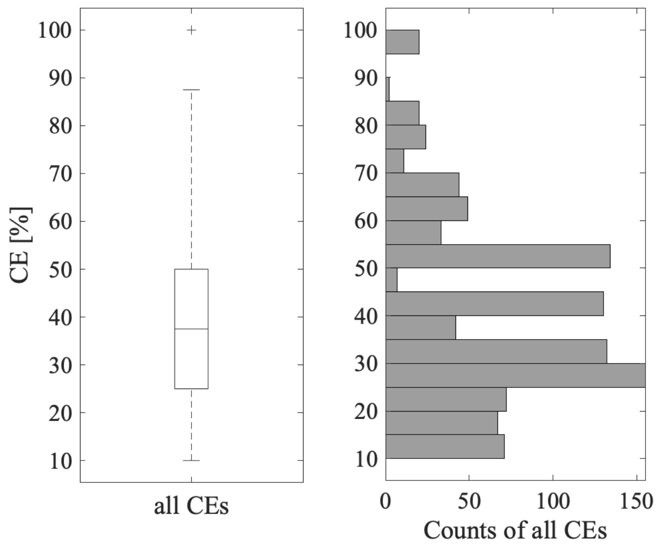

Figure 26: Boxplot and histogram for the distribution of calculation error (CE) in the cognitive task over all three experiments.

The distribution of calculation errors in percentage shows a very wide range from 10 to 100 %. An error rate of 50 % would reflect guessing probability. Anything above this rate would point towards systematic erroneous answers. One observation that should be taken into account is that participants often reacted too late when they intended to give a 'pa' sound in the result comparison phase. This was caused because this phase had a temporal limit to force participants to quickly decide if an error occurred or not. What could become visible is therefore not only the guessing probability of 50 %, but also the probability of making the right motor reaction to the task at the right time. This could mean that for trials, where participants did the right calculation and an error was displayed, that they would have reported, but their vocal reaction was too late, this could result in additional errors. A possible explanation for very high error rates above 50 % could also be that the task was systematically misunderstood and the answering mechanism was swapped. This could be the case for the first blocks of each experiment, where participants were still practicing the task. It does not explain, why error rates remained high throughout the whole experiment. Error rates around 50 % pose another problem, as they could indicate that participants were not executing the task by calculating, but by simply guessing the answers. This, however, seems very unlikely because participants were never alone during the experiment, and their results were being monitored closely to ensure correct task execution. Therefore, the most obvious explanation here is that this distribution is caused, because error rates were calculated blockwise, and if only a few tasks (i.e., two or three calculation tasks during a block) were executed, it is likely that higher error rates can emerge. To further ensure that participants executed the calculation task and understood the instructions, the error rates were also calculated for every participant throughout the whole experiment (see Figure 27).

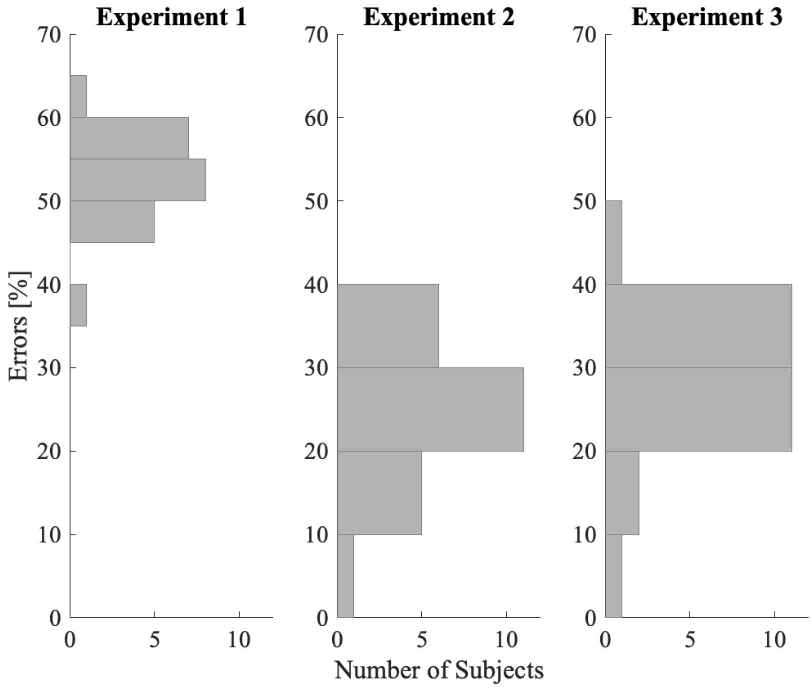

Figure 27: Calculation error in percentage over all conditions per participant for experiments 1-3.

Here, it becomes visible that the calculation error from experiment 1 should be excluded from any further statistical analysis, as the distribution would imply that participants gave systematically wrong answers. For experiments 2 and 3, however, the distributions show that calculation errors lie below the 50 % guessing probability. This implies that participants understood the task and executed it throughout the experiment. Therefore, all data displayed in Figure 27 from experiments 2 and 3 is kept for further statistical analysis.

In each of the experiments 2 and 3, there was one participant with an error rate below 10%. Here, it would be interesting to know if these participants were especially slow in their calculations to avoid errors. Therefore, the mean number of calculated tasks is determined. In experiment 2, the mean number of calculated tasks per participant throughout the whole experiment was 661.000 (± 204.080) tasks. The participant with an error percentage of 9.819 % accomplished 662 tasks, which is above the average of the sample. In experiment 3, participants calculated a mean of 57.039 (± 10.615) tasks. The participant with an error rate of 8.333 % completed 56 tasks, which lies slightly below the average of the respective sample. Additionally, there was no correlation between the number of calculated tasks and the error rate for neither experiment 2 ($r = -0.001$, $p = 0.999$) nor experiment 3 ($r = -0.232$, $p = 0.254$) in a Pearson correlation.

5.4 Motor Task

The motor task was designed to be a continuous task, where the movement control is decisive for performance. Ulrich et al. (2006) have shown that response execution can also be part of bottleneck processing. Thus, the task was designed in a manner where response execution and effect monitoring are the most critical phases for task execution.

Here, a visually presented compensatory force tracking task on a leg press was chosen, which required continuous comparison of the produced forces with the target force. The forces that had to be produced by participants were determined as force profiles described by curves that were created before every experiment using each participant's individual maximum force. The curves had different characteristics for every experiment, but had in common that they were always centered around 10% of each participant's maximum force. A precise explanation of curve characteristics will be given in the chapters for every experiment.

Participants had no explicit knowledge of these underlying curves but saw a visual representation of the difference between their produced force and a value that should be achieved at a specific time. A dashed line represented the set point, i.e., the vertical position that corresponds to zero difference between template and applied force and was visually fixed on the screen. Hence, the vertical position of a dynamically moving bar represented the deviation of the applied force from the required force. If the bar was below the dashed line, the momentary force was smaller than required, and vice versa. Thus, participants had to adapt the applied force to follow the instruction, i.e., keeping the arrows in the middle of the bar as steady as possible at the level of the dashed line.

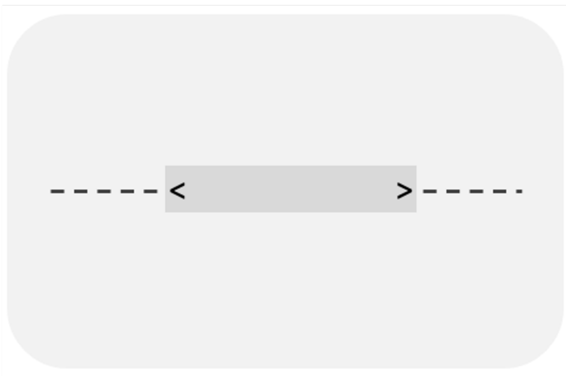

Figure 28: Visual presentation for the motor force tracking task. Dashed line = vertical position corresponding to zero difference between applied force and template. Bar = represents the deviation from the template.

For the dual and triple task conditions, the calculation task could also be represented within the moving bar, allowing participants to focus on one area visually on the screen during multiple task execution. The task has a continuous nature with inner dependency of processing stages, where the force comparison is always taking place during a task block of 60 seconds. The task has no explicit time limit other than block duration, as was the case for the result comparison phase in the calculation task. However, it does not allow for much timely flexibility due to its continuous character. Therefore, the force tracking task can be classified as a hybrid between flexible and fixed time regimes, allowing for some time slack that might be measurable as a time lag in the response to the visual presentation. Since this slack is limited because a continuous update of visually presented stimuli and produced forces is necessary, performance should also be measurable in the PCU dimension.

Requirements during the processing stages are as follows (see Figure 29). *Perception*: The position of the bar in relation to the arrows indicating the force requirement at a certain time must be perceived. Since the task is continuous, the visual stimuli are expected and have only moderate processing requirements. However, the identification if the bar position is below or above the force requirements is not trivial. *Response Selection:* During this phase the selection of a de- or increase in force must be computed regarding the underlying perception of the bar position. This leads to processing requirements similar to those in the preceding stage. *Response Execution:* The adaptations in force must be executed as programmed. Since this poses a complex motor action, this phase has the highest attentional requirements. *Effect Monitoring*: Here, the effect caused by the force adaptation must be monitored. This also poses high requirements because the monitored effect directly influences the next perception phase.

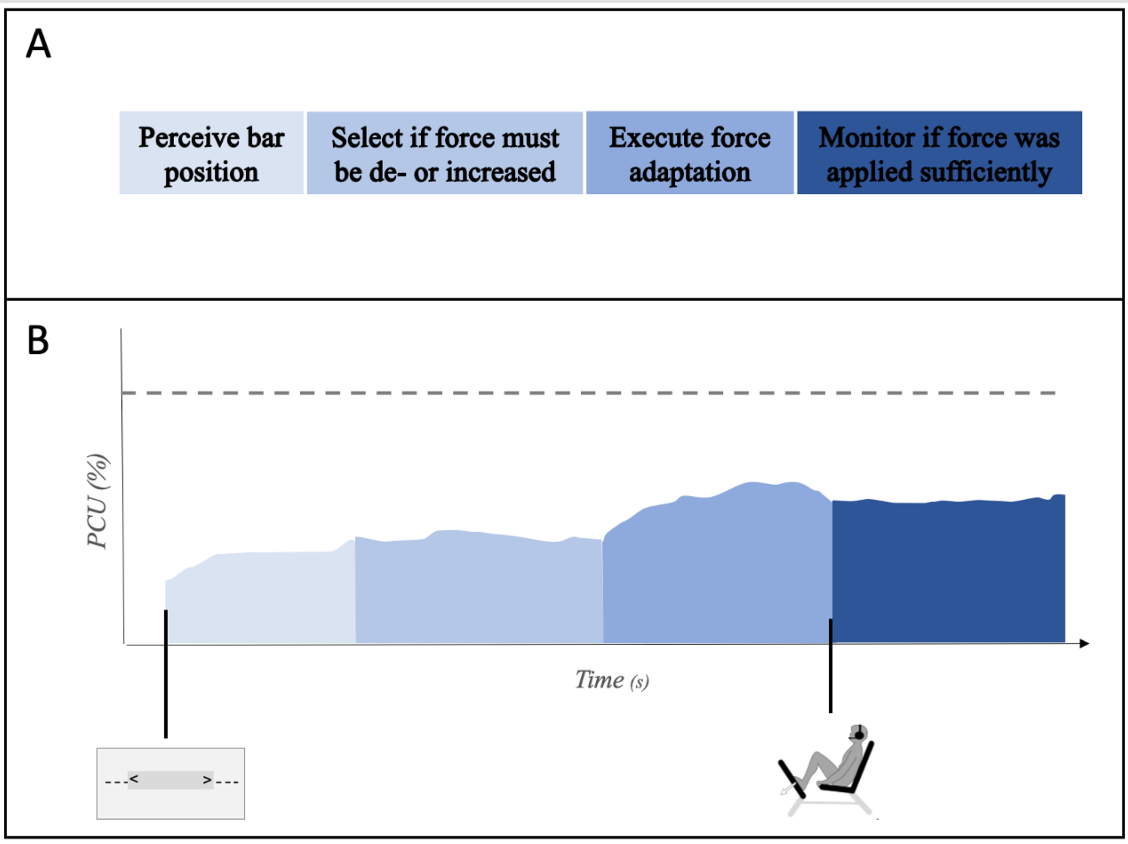

Figure 29: Combination of S-R-Model with the TR model for the motor force tracking task.

5.4.1 Force Tracking Task throughout a Task Block

During a 60-second task block, the described stages in Figure 29 have to be executed continuously, as represented by a blue coloring in the x-y plane over the complete course of the block in Figure 30. In said figure, an exemplary force production by a participant is depicted in the x-z plane. Visual feedback through bar positions was also provided continuously, but is drawn into the figure only during six discrete times for an exemplary demonstration.

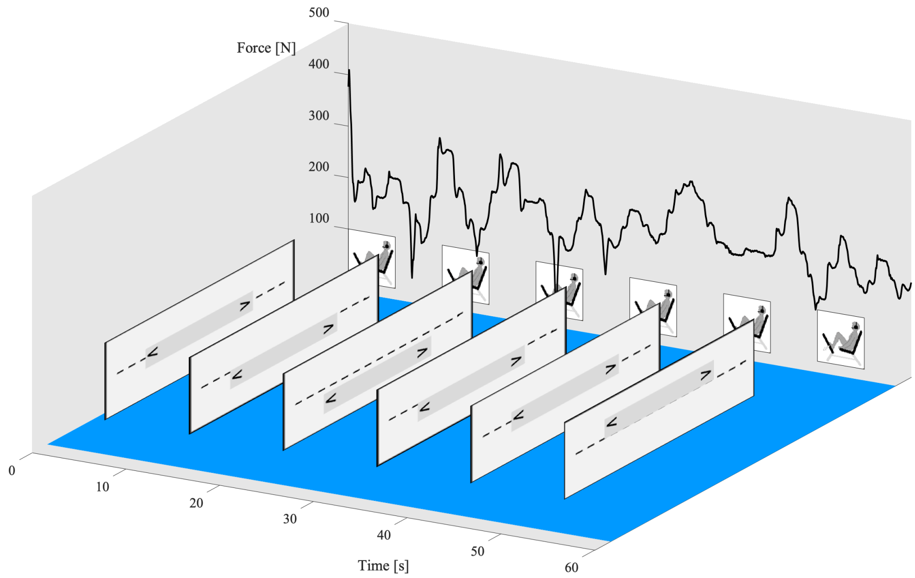

Figure 30: Motor force tracking task throughout a single task execution block of 60 seconds. The black line in the x-z-plane represents the produced forces by the participants resulting from the visual presentation on screen.

Performance can be measured as the time lag of the produced force curves by the participant compared to the underlying set curves for the time dimension and as the root mean square error for the PCU dimension in the time regimes model.

Regarding the content-based interference, probe stimuli referencing to force intensities or directions should specifically interfere with motor task performance.

5.4.2 Data Preprocessing for the Motor Task

The forces produced by the participants were measured as a voltage signal by the strain gauge. An online processing was done with the data, transforming it from Volt to Newton using a linear function that was determined in a calibration before the experiments. One of those functions was:

$$N = 666.8120 \times V + 2918.687$$

Online data was also filtered with a Butterworth lowpass filter of second order with a cutoff frequency of 10 Hz to ensure a smooth visual task representation. For experiment 1, this already pre-processed data was saved. For experiments 2 and 3, the raw data as voltage signal was saved to enable further pre-processing if necessary. To determine if the filtering frequency is suitable or should be adapted, an FFT is executed with the data from the strain gauge (Figure 31). The assumption of Winter (2003) that low cutoff frequencies are suitable for human kinematic data can also be applied here. Because the data from each measure of the sensor at rest before every force tracking task to determine the rest value of the sensor was not saved, contrasting the signal at rest with the signal with applied forces is not possible. Throughout task execution force was constantly applied by participants. Therefore, the FFT can only be calculated for the whole signal over a task block of 60 seconds.

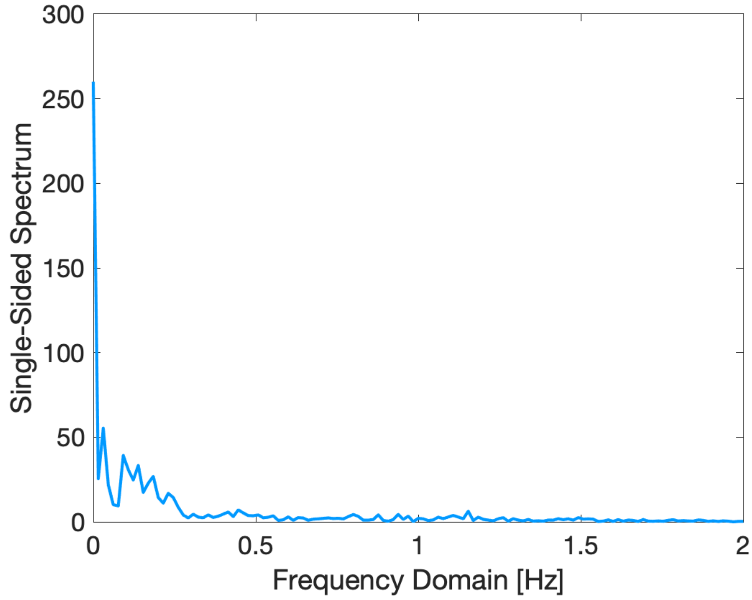

Figure 31: Single-sided spectrum of the signal from the strain gauge for the motor task depicted only over a frequency domain from 0 to 2 Hz.

The figure depicts that a high signal spectrum is taking place at very low-frequency below 2 Hz, as can be expected because the characteristics of the underlying curves induce low frequent movements. A cutoff frequency of 10 Hz or below is therefore recommended. Since the data was already filtered at 10 Hz in the online processing, this approach is maintained. After filtering the data with a second-order Butterworth lowpass filter with a cutoff frequency at 10 Hz, it should be ensured that no time shift along the x-axis was caused, resulting in a shift of local extrema, to ensure a valid determination of the motor error in the next step. To visualize this, the filtered and unfiltered data are plotted together in Figure 32.

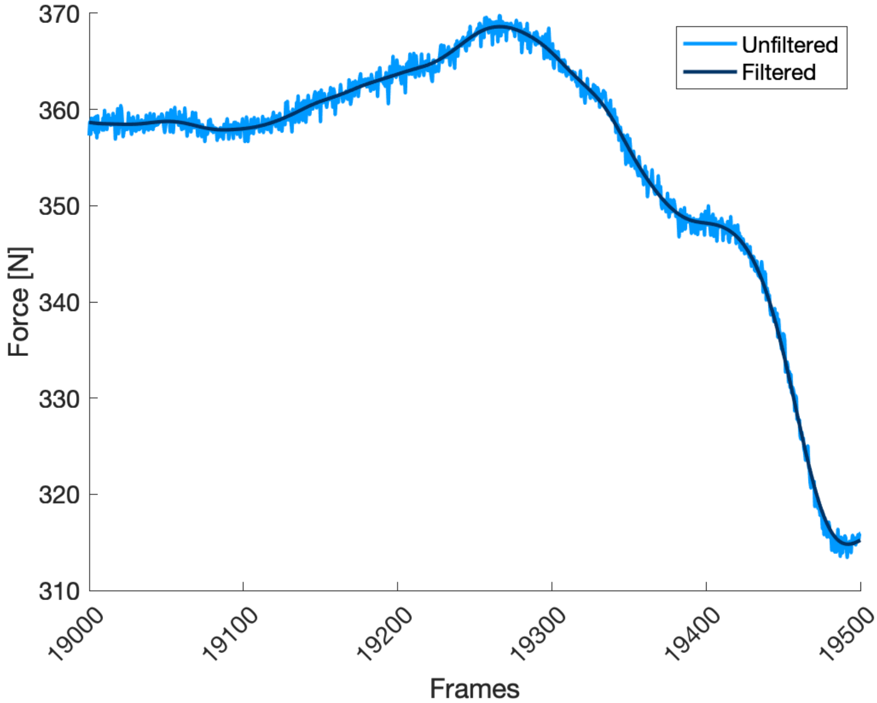

Figure 32: Raw and filtered data from the strain gauge.

The figure shows that the filtering has effectively smoothed the signal without causing any time shifts. As a dependent variable for the motor error, the root mean square error normalized at each participant's 10 % maximum force was calculated:

$$RMSE_{norm} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(curve - force)^2}}{0.1 \times f_{max}}$$

5.4.3 Calculating Δt as a Dependent Variable for the Time Lag of Produced Forces

While the motor error holds the information about how much the produced forces deviate from the curve, it allows no specific assumptions about the time lag with which participants adapt their forces to the visual presentation of the force tracking task. However, a time lag can certainly be observed in the data (see Figure 33).

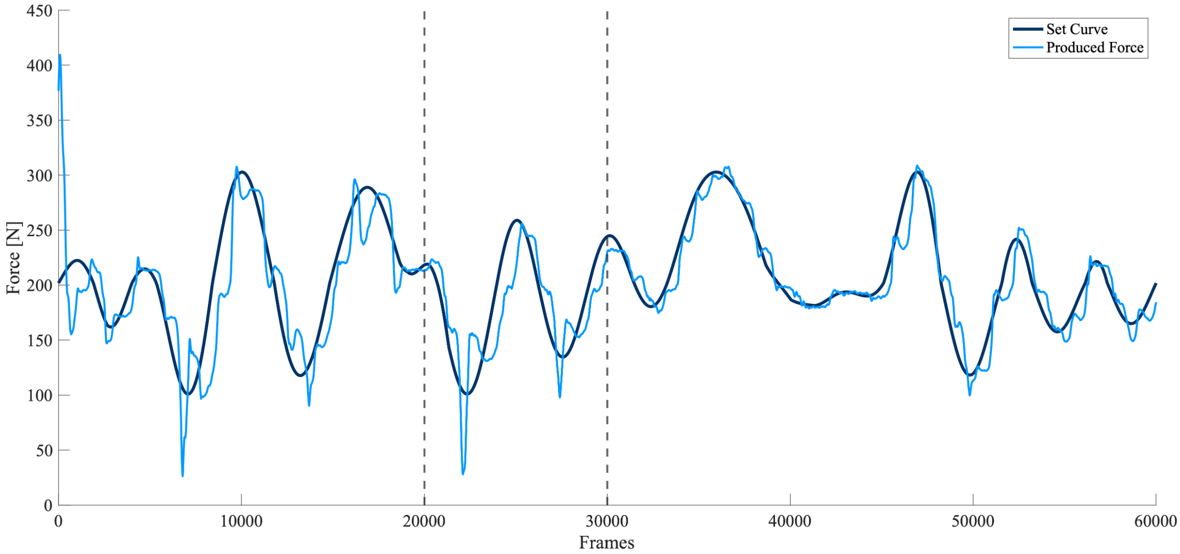

Figure 33: Set force curve and produced forces in an exemplary task block of experiment 3. Duration of the block was 60 seconds with a data acquisition frequency of 1000 Hz. The dashed lines represent a time window between 20000 and 30000 frames, which is used to test for different Δt.

An additional measure is calculated and used as dependent variable that can account for this observable time lag. For this, the filtered produced force of the participants is shifted on the time axis at different pre-defined windows (Δt) and then correlated with the set curve (see Figure 34). By calculating the correlation coefficient r^2, the best-fitting time shift can be sought.

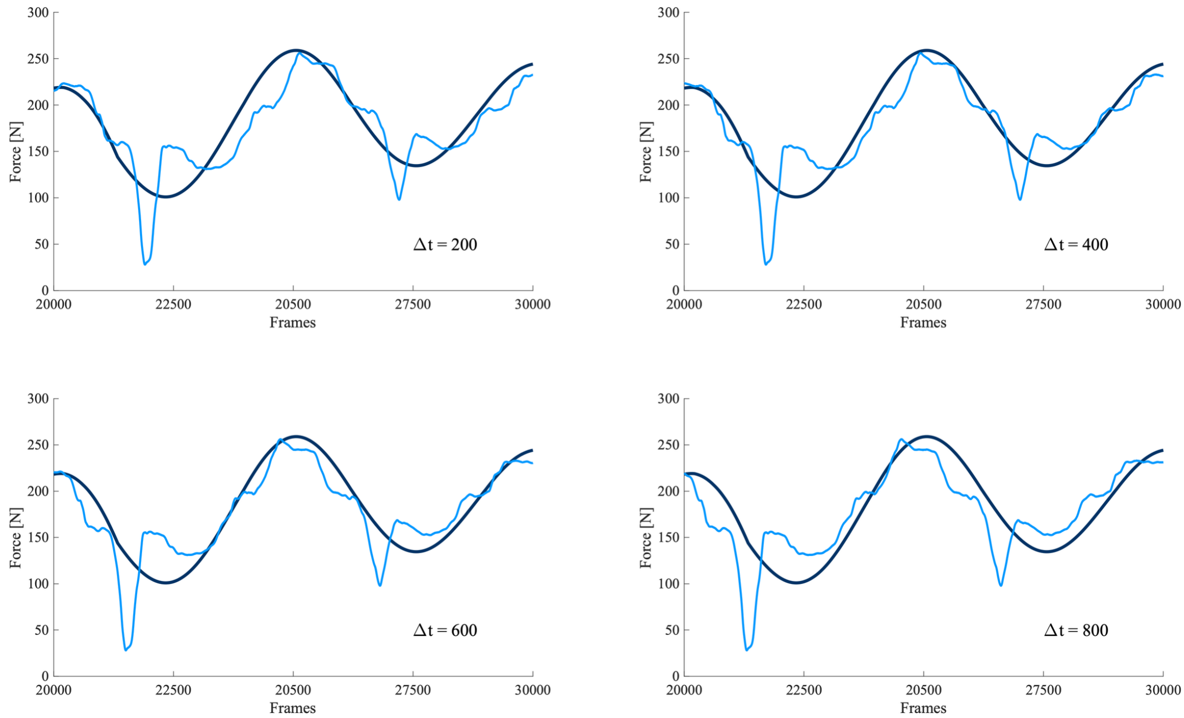

Figure 34: Data from Figure 33 in the time window between 20000 and 30000 frames. The produced force was shifted by the Δt depicted in each subplot.

For the data from the three experiments, produced forces were shifted from 0 to -1000 frames in steps of 10 frames of Δt on the x-axis. This results in 101 different Δt shifts. By correlating the shifted produced force with the set force curve, 101 r^2 can be calculated for each trial block. An example of the resulting r^2 values for one participant is shown in Figure 35.

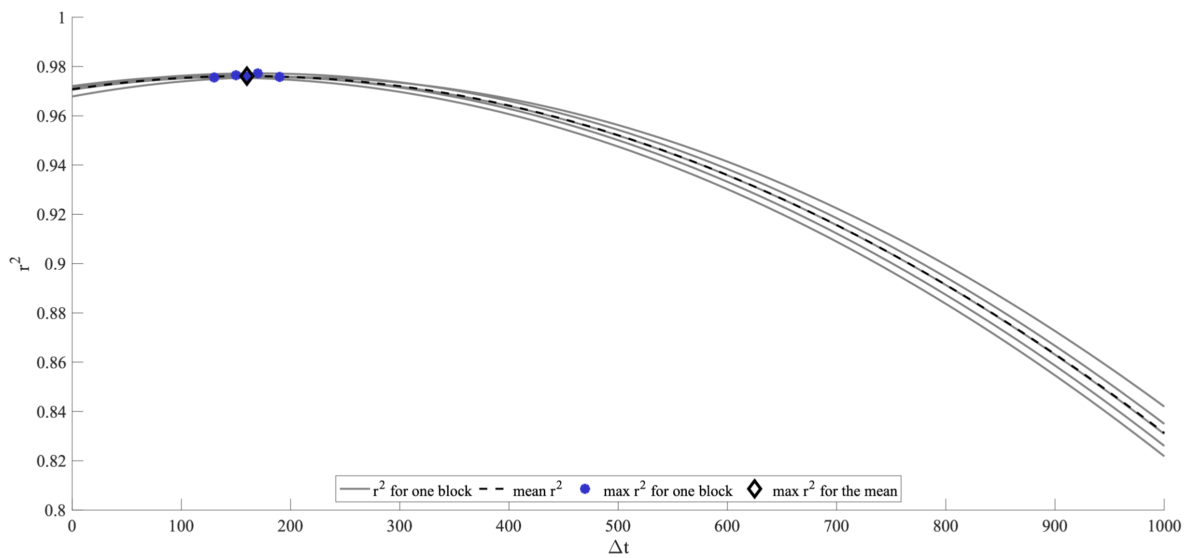

Figure 35: Exemplary r^2 at all Δt from 0 to 1000 frames for Experiment 1, Subject 1 under single-task condition.

The maximum of r^2 is then sought with its corresponding Δt and taken as performance measure for one block. To get Δt as the dependent variable for blockwise performance, the average r^2 values over Δt for all blocks in one condition per subject are calculated by transforming data into Fisher-Z-Values (Rasch et al., 2008, p. 128) to enable the calculation of the arithmetic mean. Data is then transformed back into r^2 values. Now again, the maximum r^2 is determined for the average r^2 over all blocks.

5.4.4 Data Included for Further Statistical Analysis

To determine which data for the motor error (ME) and Δt from the force tracking task should be included in the statistical analysis, distributions for the dependent variables are created for the data over all three experiments (see Figure 36).

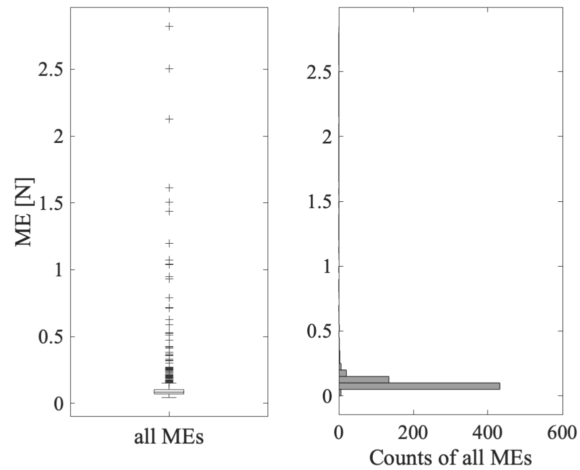

Figure 36: Motor Error (ME) in N calculated as RMSE over all three experiments represented in a boxplot (left) and histogram (right).

Only one trial had to be excluded with an RMSE of over 10N, which can be clearly identified as an outlier that was caused by technical problems. There was no other reason to exclude any further data.

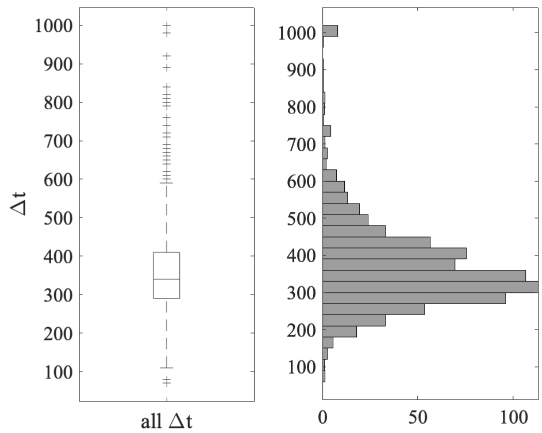

Figure 37: Motor time lag calculated as Δt over all three experiments represented in a boxplot (left) and a histogram (right).

For the Δt, a span ranging from 0 to 1000 frames was tested for the best fitting correlations. The data shows the profile of a normal distribution. No exclusion criterion was defined, as a timely shift of 1000 frames still seems to be valid, indicating that force production as an adaptation to the visually presented bar was shifted in time by one second. This seems to be a reasonable time when taking into consideration that force production of the legs poses a rather low frequent motion.

5.5 Remarks on the Inferential Statistical Procedures

The inferential statistical analyses for the three experiments are calculated using JASP 0.17.1. The α significance level of all location tests is set to 5%. Location tests are either paired samples Student's t-tests, repeated measures analysis of variance (ANOVA) or their non-parametrical equivalent.

Both parametric analyses require the dependent variable to be metric (interval or ratio scale level). If scale level is questionable, the non-parametric equivalent is calculated. This is explicitly communicated at the appropriate place. To test for normal distribution, i.e., bivariate normal distribution for the t-test, the Shapiro-Wilk test (SW test) is used. If the p-value for the SW test lies below 5%, normal distribution is violated. For the repeated-measures ANOVA, normal distribution of residuals is further visually assessed using Q-Q plots (see GitHub Repository[3]). Following Rasch et al. (2008, p. 49), the t-test and ANOVA are robust to the violation of normal distribution. Therefore, both statistical analyses can be used, nevertheless. Sphericity for the repeated-measures ANOVAs is tested using the Mauchly test, and if violated ($p < 0.05$), a Greenhouse-Geisser correction is used for the degrees of freedom.

For the post-hoc analyses of the repeated-measures ANOVAs the Tukey Honest Significant Difference (HSD) test is used to compare pairwise differences. To correct for α error cumulations, the Bonferroni correction is used for the p-values.

[3] https://github.com/JelenaMueller/DynamicCapacityAllocation/tree/main/Statistics

6 Experiment I – Pilot Study

The main purpose of this experiment was to replicate the classical performance decrement findings of other multi tasking studies for the time-based task interference assumption (R1). Also, the effect of the semantic content of the probe stimuli on the other two tasks is tested for the content-based task interference (R2). The chosen auditory probe stimuli were expected to specifically lead to a decrement in the performance of either the motor or the cognitive task, depending on certain semantic stimulus properties. In the first experiment, all three tasks were executed under single task conditions and then combined to triple-task execution.

The hypothesis for this experiment is, that performance in all three tasks decreases in triple, in comparison to single-task conditions, for the time-based interference assumption. Hypothesis for the content-based interference are that probe stimuli holding numbers lead to a greater performance decrement than other stimuli in the calculation task. Probe stimuli that hold intensities lead to a greater performance decrement than other stimuli in the force tracking task under triple task conditions. Even though in the simple reaction time task, no processing of the semantic content of the probe stimuli was necessary, a semantic upstream processing that would occur automatically could be possible (Jiang et al., 2016). This gives rise to the following two research hypotheses:

H1.R1: Performance decreases unter TT in comparison to ST.

H1.R2: The semantic content of probe stimuli leads to a specific performance decrement under TT, where numeric probe stimuli lead to a specific performance decrement in CLC and intensity probe stimuli lead to a specific performance decrement in FRC.

6.1 Methods Experiment 1

This experiment was conducted as a pilot study, so that no computation of required sample size was done a priori.

6.1.1 Participants

22 healthy subjects (18 female, 4 male) were tested on two consecutive days. Participants were university students of psychology or sports sciences of the lower semesters. All subjects gave their written informed consent to participate in the experiment. They reported to suffer under no neurological, psychiatric or orthopedic diseases. The eyesight was normal or could be corrected to normal. Participants were compensated with up to 3 hours of course credit. A defined amount of the latter is required in order to acquire the bachelors degree in the before mentioned fields.

6.1.2 Probe Reaction Time Task

In the probe reaction time task, participants had to react to the probe stimuli as quickly as possible by pinching index finger and thumb together. The FSR used to measure reaction times was attached to the participants left hand index finger.

The auditory probe stimuli were a beep-sound (beep), one pair of spoken numbers (num), two pairs of specific spoken words (int, dir) and one pair of unspecific spoken words (wor). Those four different probe types were chosen to either have none, or a specific interference with either the motor or the cognitive task. The pairs of probe stimuli and their expected interference are listed in Table 1.

The stimuli were recorded by an actor from the Hessian State Theater in Marburg to ensure that the intonation and pronunciation of the words remained consistent. The beep sound was created on a computer.

Table 1: Display of probetypes with used auditory probe stimuli and hypothesized influence on performance in type of task. Words were presented in German.

Type of Probe	Stimulus Pair	Stimulus Length (ms)	Targeted Performance
beep (beep)	beep	500	Unspecific
numbers (num)	„4“ and „6“	230, 553	Cognitive
intensities (int)	„stark“ and „schwach“ (“strong” and “weak”)	392, 323	Motor
directions (dir)	„rechts“ and „links“ (“right” and “left”)	403, 346	Motor
words (wor)	„Buch“ and „Schal“ (“book” and “scarf”)	230, 300	Unspecific

Independent from sound or actual semantics included in the probe, participants had to react as quickly as possible to either of the stimuli. Stimulus length is neglected, because participants did not need to fully semantically process the spoken words but could react as soon as they heard any sound over the headphones. Six probe stimuli were played at pseudorandomized times during single task execution. For triple task execution, probe stimuli were also played at pseudorandomized times taking into account the mean of the time needed for the calculation task under single task execution. This was carefully considered to be the best way of choosing the inter-stimulus interval for probe stimuli to ensure that for multiple task execution, a maximum of one auditory probe stimulus was played during each of the executed calculation tasks. One block of the single probe task lasted 30 seconds. During this time, data from the FSR was gathered, later allowing for the analysis of reaction times and reaction errors.

6.1.3 Cognitive Task

The cognitive task was the visually presented calculation task requiring subtraction of a two-digit from a four-digit number (as described in Chapter 5.3). Here, the task was answered using a computer mouse held in the right hand by the participants. The result comparison phase did not always last the full 2500 ms. If participants registered an error in the displayed result by executing a mouse click during result comparison, the second phase was instantly aborted and the next subtraction task was shown automatically initiating another calculation phase. This procedure was adopted from the motor cognitive dual tasking training study by Langhanns et al. (2022) where participants gave a verbal answer which resulted in the experimentator clicking the mouse to continue with the task. Due to the very time critical character of the underlying experiment, participants were instructed to execute the mouse click themselves. Thereby, no external factor (as the response time of the experimentator to the verbal cue of the participant) would interfere during multiple task execution.

Each task block had a duration of 60 seconds in which the participants were instructed to answer as many subtraction tasks as correctly as possible. For the single task conditions, the tasks and results were displayed at the center of the screen. In the triple task condition, the digits for the subtraction were shown at the center of the moving bar, to allow visual attention to capture both tasks at the same time, so that no task switching would be forced between the two tasks.

In this task, the timestamps for all clicks in phase one and phase two were recorded, as well as the information, if a click in the second phase was the right or wrong decision. Performance measure for the calculation task was the number of executed tasks per block (i.e. the time needed for every task in phase 1). The error percentage was excluded as dependent variable as described in chapter 5.3.2.

6.1.4 Motor Task

In this first experiment, the visually presented bar moved due to a regularly defined sinusoidal curve (see Figure 38) that was set around a mean value of 10% of each participants maximum force (F_{max}). The function describing the curve remained the same over all participants and trials and only differed in its mean y-intercept between participants due to their individual F_{max}. This type of regular curve was chosen as to not set a too high complexity in the motor task for the triple task design. Participants had no explicitly given knowledge about the appearance of the underlying curve. They were only given the information that the bar they could see on the screen would be moving and that it was their task to keep it as stable as possible around the solid line for a trial duration of 60 seconds.

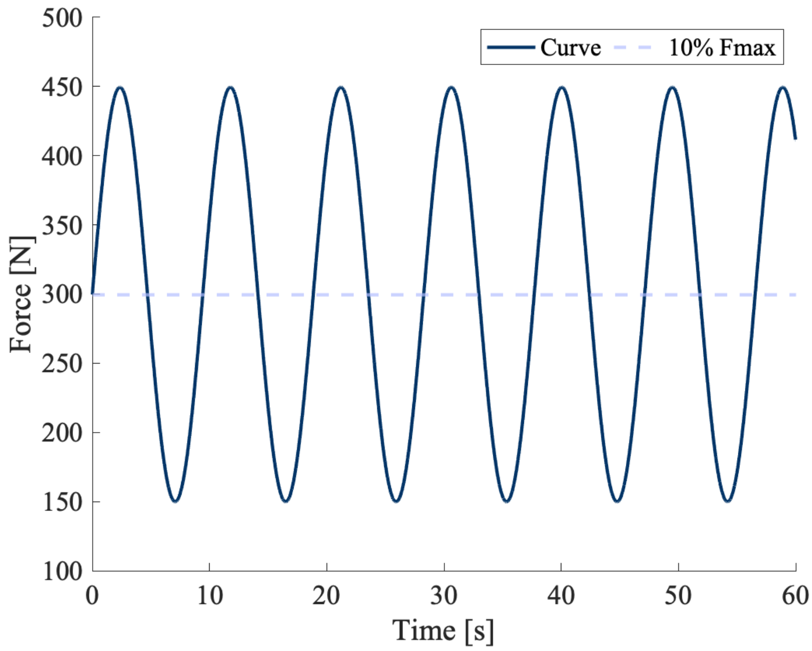

Figure 38:Exemplary force curve used in experiment 1 for a participant with F_{max} of 2996.70 N. Mean value of created curve lies at 10% F_{max}.

As raw data the produced forces of every participant per trial were collected and stored for further analysis. In later steps, this data was used to calculate the the time lag (Δt) and normalized root mean square error (RMSE) as performance measure for the motor task.

6.1.5 Triple Task Execution

The tasks were performed as single task conditions to record a baseline in performance measure in the time and PCU dimensions for each task. In the triple task condition, alle three tasks were combined during a 60 second task block. Figure 39 shows an example of such a triple-task procedure.

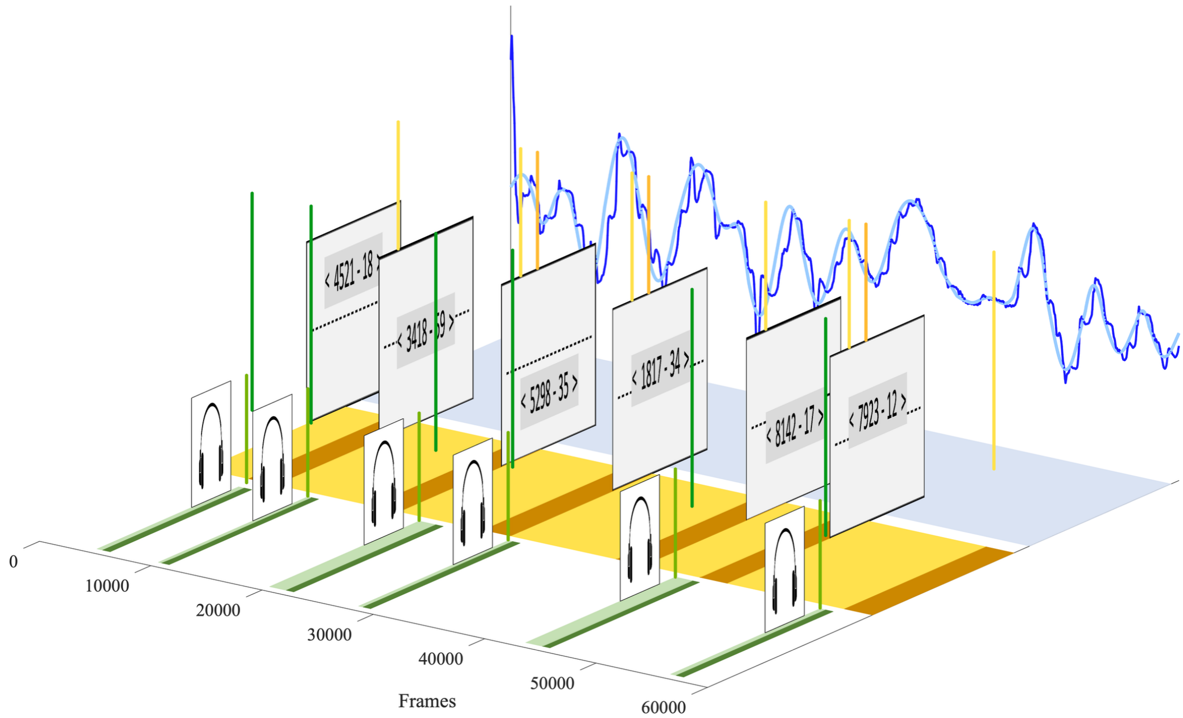

Figure 39: Triple task execution of the PRT, CLC and FRC over a task block of 60 seconds. PRT stimuli are placed at discrete times. CLC and FRC run continuously throughout the block, with CLC being divided into a calculation (yellow) and a result comparison (orange) phase. Visual digit presentation of result comparison phase in CLC is not depicted here to ensure better clarity of remaining task stimuli and responses. Captured stimulus and response behavior is depicted in the y-z-plane with their according time stamps on the x-axis.

In this example of a triple task block execution, six probe stimuli were given. Inter-stimulus intervals of the stimuli were calculated from the mean calculation time in CLC under single task conditions, as to target each calculation task during triple task execution with a probe stimulus.

6.2 Procedure Experiment 1

Day 1

On the first day subjects trained the cognitive as well as the motor task under single task conditions. They had to perform a minimum of five blocks for the cognitive, and two blocks of the motor task. If for the cognitive task during the last block at least 6 tasks were solved with an error rate below 50%, performance was rated as sufficient and no more single task blocks had to be executed. For the motor task the performance criterion was that the deviation must be lower than 50 % in the last executed block. Participants were all able to reach the required criteria in the minimum number of five blocks. The probe reaction time task was not trained on that day, as learning processes need not be expected here due to neurophysiological constraints of reaction time tasks.

Day 2

At the beginning of the second day, the calculation task was first executed as single task condition, to get a baseline for the comparison with the triple task. A total of five blocks with a duration of 60 seconds were accomplished. After this, the same procedure was repreated with the force tracking task. After that, the probe reaction time task with one stimulus pair was first carried out as single-task (30 seconds), followed by 60 seconds of the triple task with the same probe stimulus. This alternation of single task probe and triple task was repeated five times in order to include all five stimulus pairs. After that, this procedure was repeated one more time, resulting in ten single task probe and ten triple task blocks. The order of probe stimuli remained the same over all subjects.

6.3 Data Pre-processing Experiment 1

The data pre-processing was done as described in chapter 5. Note that the data for the error information in the calculation task was neglected. Only the times for clicks in the calculation and comparison phase are complete and valid and therefore used to derive calcuation times in the calculation phase of CLC. For every block, the mean calculation time is calculated. Calculation times under 2 seconds are not taken into account because they are not considered valid measures but rather appear to be caused by accidentally double-clicking the mouse.

6.3.1 Elimination of Sequential Effects

The statistical analysis testing for performance decrement and the semantic probetype effect can be done with the direct measures for the dependent variables. Those are:

Reaction Time [ms]
Reaction Error [%]
Calculation Time [s]
Motor Error [N]
Motor Δt [frames]

For all five dependent variables, higher values indicate a decrease in performance (i.e. slower reaction time, higher calcualtion error, etc.) than lower values. By the design of the first experiment that was done as a cross-sectional study, familarization effects, as well as fatigue cannot be fully excluded. Even though all participants have supposedly carried out a mental subtraction, force production of the legs and reactions to auditory stimuli before, the context and conditions under which the tasks are carried out in the experiment were new to participants. Therefore, familiarization and fatigue have to be considered and, if taken place, removed from the data in order to regard only effects that are caused by the different task load conditions and probe type semantics. To remove such effects, a regression analysis can be calculated. The regression then describes familarization by a function with a negative slope, resulting in smaller values of dependent variables with increasing trial numbers and therefore an increase in performance. Regression functions with a positive slope could indicate fatigue by resulting in higher values with increasing trial numbers and therefore worse performance. By calculating residuals from the regression function and using these for further statistical analysis, only the task load and probetype effects remain.

For the regression analysis, the best-fitting curve can be sought for every dependent variable and participant. A first assumption about curve characteristics should however be made, to allow specific testing for several function types. It can be presumed, that with an increasing number of trials, the performance for a measured dependent variable will also increase. This would correspond to a linear function. Newell and Rosenbloom (1981) introduced the power law of practice in alignment with the empirical findings of James (1950) that the acceleration of the learning performance decreases with an increase in practiced trials. This conforms to an exponential or quadratic function with a negative slope. When looking at the empirical data of the first experiment it becomes quite clear however, that linear regressions are sufficient to depict possible familiarization or fatigue (see Figure 40).

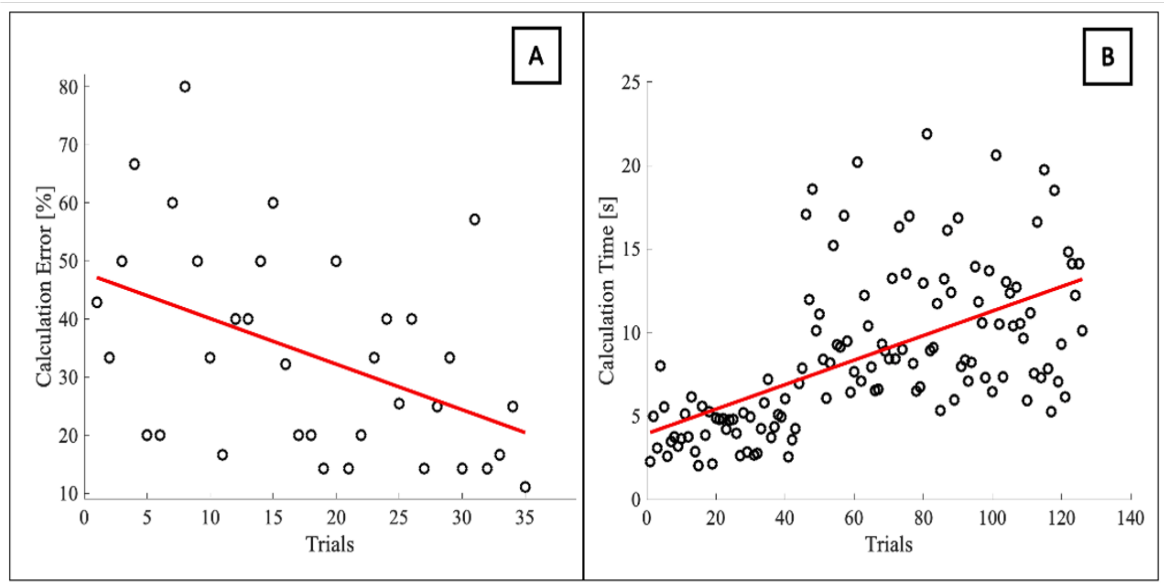

Figure 40: Examples for familiarization (A) and fatigue (B) in the dependent variables calculation time and calculation error.

A: Familiarization for calculation error of subject 1 in experiment 2

B: Fatigue for calculation time of subject 2 in experiment 1

Note that for the calculation error less data is available because the percentage of errors can only be calculated blockwise, while the calculation times are depicted for every trial of each block.

To ensure, that the regression analysis does not eliminate any task load effects from the data that could result from the experimental design, a condition-wise horizontal alignment was undertaken before calculating the regression and resulting residuals (Figure 41). The mean value per task load condition (in this case single and triple task) and participant was calculated and subtracted from the corresponding trial data. For this horizontally aligned data, the linear regression could then be calculated with the least squared error method. Residuals are calculated by building the distance from each trial data to the regression. The residuals are then subtracted from the trial data before horizontal alignment.

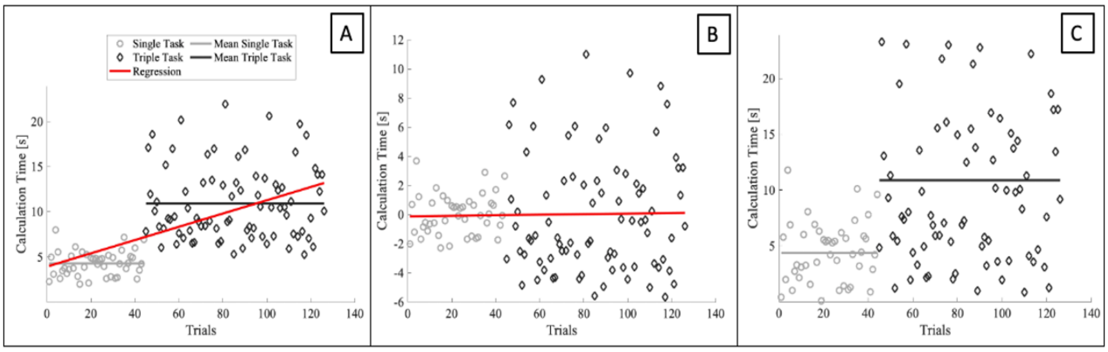

Figure 41: Condition wise horizontal alignment and calculation of residuals.
A: Calculation time from subject 1 in experiment 1 under single and triple task conditions with respective mean values and regression over the course of the experiment.
B: Data after horizontal alignment (condition wise mean was subtracted from respective data). Linear regression is fit through the data and residuals are calculated.
C: Calculated residuals from B were subtracted from data in A. The lines representing mean values for single and triple task hold the mean of the calculation time – residuals.

This method is used for all dependent variables in the first experiment. The data corrected by the horizontally aligned residuals to eliminate familiarization or fatigue effects without eliminating possible task load effects is then used for further statistical analysis.

6.4 Results Experiment 1

Statistical analyses are first calculated blockwise for the dependent variables to check for a general performance decrement, as found in cognitive psychological research. Analyzing probe semantics will then enable a closer look at the context-based interference.

6.4.1 Blockwise Performance Decrement

Because tasks were only executed under single- and triple task conditions, paired samples t-tests (one-sided) were calculated for every dependent variable. The prerequisite of bivariate normal distribution was tested with the Shapiro-Wilk (SW) test. However, the t-test is robust to a violation of the assumption of normal distribution for sufficiently large sample sizes (Bortz & Schuster, 2010). If no other prerequisites are violated, parametric tests are nevertheless calculated. For all calculated statistical tests, the alternative hypothesis specified, that performance under triple task conditions was decreased (resulting in higher values within each dependent variable), in comparison to the single task conditions.

Reaction Time

The mean reaction time under single task execution is 0.549 (±[4] 0.075) s. For the triple task execution, the mean reaction time is 0.745 (± 0.103) s. The SW test showed that bivariate normal distribution can be accepted ($p = 0.906$), therefore the dependent samples t-test is interpreted for the one-sided alternative hypothesis that $RT_{SingleTask} < RT_{TripleTask}$: $t(21) = -9.308$, $p_{one\text{-}sided} < 0.001$, $d = -1.985$. Reaction times are prolonged under the triple task condition in comparison to the probe reaction time task execution as single task condition.

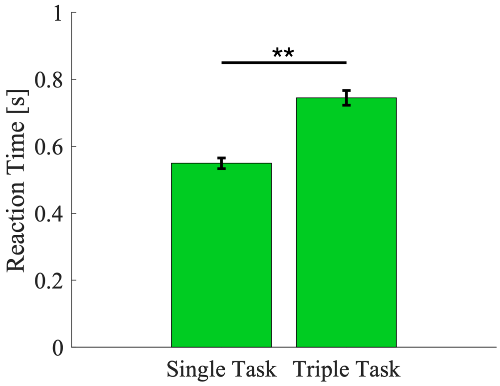

Figure 42: Reaction time for single and triple task conditions in experiment 1. Error bars represent standard error of the mean.

[4] Values in brackets with "±" always report standard deviation.

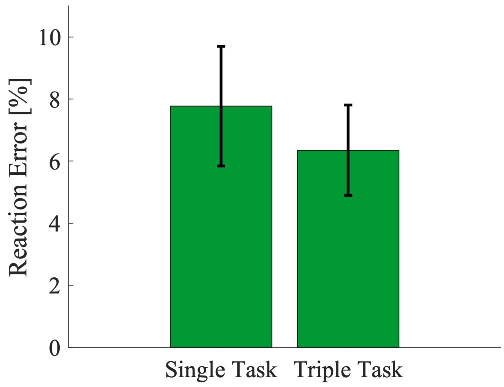

Figure 43: Reaction error in % for single and triple task conditions in experiment 1. Error bars represent standard error of the mean.

Reaction Error

The mean reaction error rate in percent is 7.771 (± 9.041) % for single task and 6.351 (± 6.829) % for triple task execution. Normal distribution is accepted using the SW test (p = 0.457). In the dependent samples t-test, no statistical significance was found for indicating that reaction errors are lower under single task in comparison to triple task conditions: $t(21) = 0.782$, $p_{one\text{-}sided} = 0.782$, $d = 0.167$.

Calculation time

Under single task execution, the mean calculation time is 6.330 (± 1.329) s; under triple task execution it is 7.444 (± 1.907) s. For the calculation time, bivariate normal distribution can be accepted as well ($p = 0.917$). The t-test statistics ($t(21) = -5.993$, $p_{one\text{-}sided} < 0.001$, $d = -1.278$) show, that calculation time under triple task condition is prolonged in comparison to single task execution.

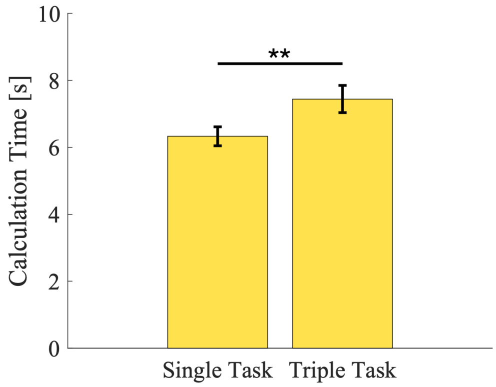

Figure 44: Calculation Time for the Cognitive Task under Single and Triple Task Conditions in Experiment 1. Error bars represent standard error of the mean.

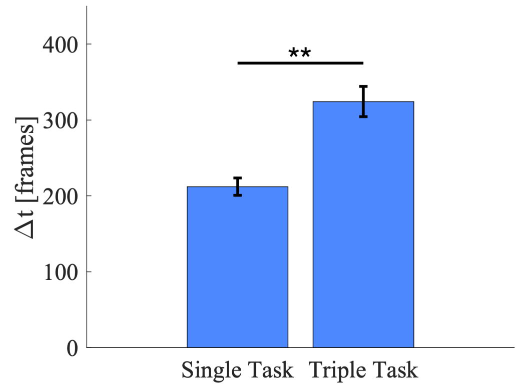

Figure 45: Motor Time Lag for the Motor Task under Single and Triple Task Conditions in Experiment 1. Error bars represent standard error of the mean.

Motor Time Lag

The mean motor time lag during single task execution is 212.115 (± 53.659) frames; for triple task execution it is 324.351 (± 93.481) frames. For Δt, the SW test showed a violation of the normal distribution (*p = 0.012*). T-test and Wilcoxon signed-rank test led to the same result pattern. Thus, the t-statistics are reported here for better comparability of results: $t(21) = -7.386$, $p_{one\text{-}sided} < 0.001$, $d = -1.575$. Thus, time lag under triple task was higher in comparison to single task.

Motor Error

The mean normalized RMSE for single task execution is 0.063 (± 0.013) N; and for triple task execution 0.096 (± 0.025) N. Normal distribution for the normalized RMSE is also given ($p = 0.228$). The t-test showed a statistically significantly higher RMSE under triple in comparison to single task conditions ($t(21) = -8.967$, $p_{one\text{-}sided} < 0.001$, $d = -1.912$).

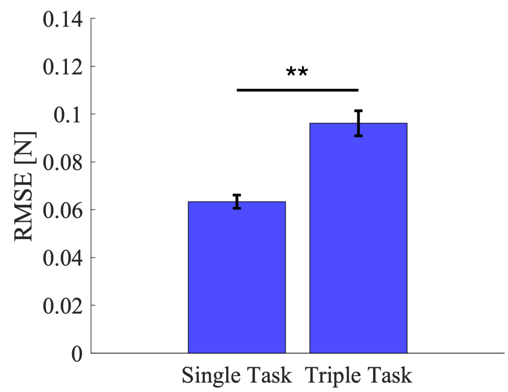

Figure 46: Normalized Root Mean Square Error for the Motor Task under Single and Triple Task Conditions in Experiment 1. Error bars represent standard deviation of the mean.

6.4.2 Semantic Probetype Effect

For the examination of the probetype effect, only data from triple task execution is analyzed. Normal distribution is tested using the Shapiro Wilk Test (S-W Test) and normal distribution of residuals is visually tested using Q-Q Plots. Repeated measures ANOVAs for the factor probetype (beep, num, int, dir, word) are calculated.

Reaction Time

Repeated-measures ANOVA for the factor probetype (beep x num x int x dir x word) included 19 participants. For three participants, reaction time data was missing for either one of the probetypes, either due to responses measured more than two seconds after stimulus presentation or not responding to a specific probetype at all. The Shapiro-Wilk test shows normal distribution for all factor levels ($p > 0.05$), except for the reaction times to the beep probetype (p = 0.006). Q-Q-Plots revealed that normal distribution of residuals can be accepted. Sphericity is tested using the Mauchly test (p = 0.305). A statistically significant main effect was discovered for the probetype in reaction times (*$F(4,72) = 34.974$, $p < 0.001$, $\eta^2_p = 0.660$*). Statistically significant post-hoc comparisons are listed in Table 2.

Table 2: Statistically significant post-hoc comparisons of the repeated-measures ANOVA for the factor probetype (beep x num x int x dir x word).

		t	**Cohen's d**	**p_{bonf}**
beep	num	-11.143	-2.844	<0.001
	int	-3.914	-0.999	0.002
	dir	-2.941	-0.751	0.044
num	int	7.228	1.845	<0.001
	dir	8.201	2.093	<0.001
	word	8.514	2.173	<0.001

The beep stimulus (0.625 ± 0.035 s) differs statistically from all other probetypes, except for the word (0.690 ± 0.025 s) stimuli (see Figure 47). Reactions to the beep stimulus were the fastest. The numeric probetype (0.971 ± 0.031 s) also differs statistically significantly from the other probetypes (int 0.7316 ± 0.035 s; dir 0.698 ± 0.022 s). Reaction times to numeric stimuli were the slowest.

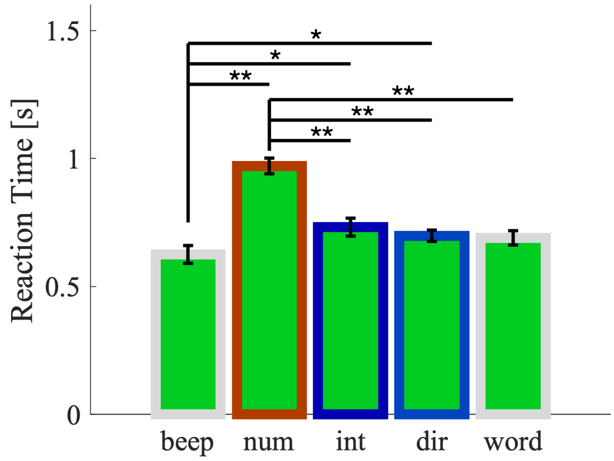

*Figure 47: Reaction times to probe stimuli with different semantic characteristics. Error bars represent standard deviation of the mean. *p ≤ 0.05; **p ≤ 0.01.*

Reaction Error

For the repeated-measures ANOVA for the factor probetype with the reaction error, data of 21 participants could be included in the analysis, due to one missing value for the word probetype.

The error rates in the PRT for each probetype were not normally distributed (*S-W test $p < 0.05$*). Q-Q-Plots also revealed that residuals are not normally distributed. Therefore, the Friedman-Test is used as non-parametric equivalent to a repeated-measures ANOVA for the factor probetype (beep x num x int x dir x word). No statistically significant effect for the factor probetype was found (*$\chi^2 = 4.404$, $p = 0.354$*).

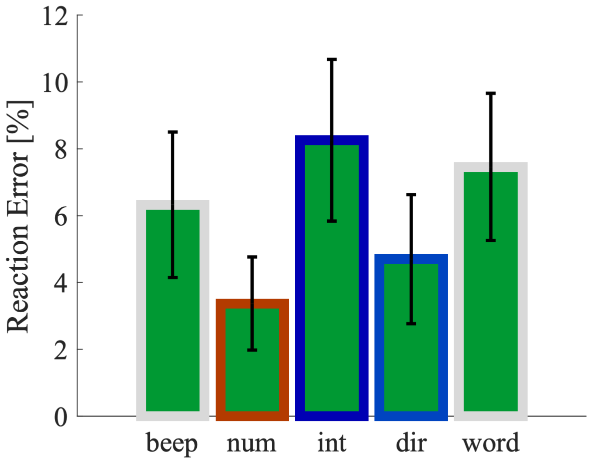

Figure 48: Reaction errors to probe stimuli with specific semantics. Error bars represent standard deviation of the mean.

Calculation Time

For the analysis of the probetype effect for the calculation time, a repeated-measures ANOVA was used, where 21 participants could be included. In the S-W Test, calculation times for all probetypes showed normal distribution. Q-Q-Plots showed that normal distribution for all probetypes can be accepted. Sphericity tested by the Mauchly test was accepted (p = 0.055). A statistically significant effect was found for the factor probetype ($F(4, 80) = 7.910, p < 0.001, \eta^2_p = 0.283$).

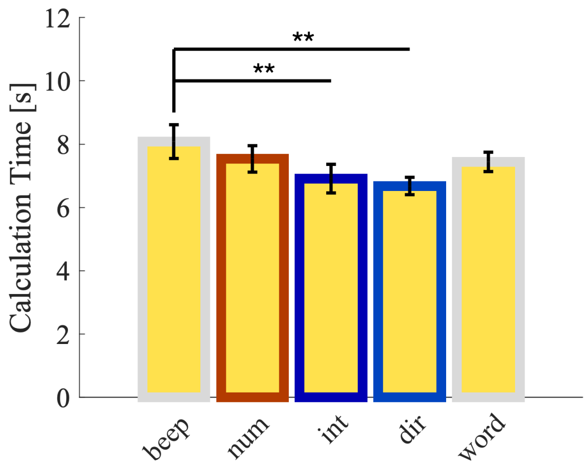

*Figure 49: Calculation times in s for the calculation phase of CLC during triple task execution with different semantic characteristics of probe stimuli in PRT. Error bars represent standard deviation of the mean. **p ≤ 0.01.*

Post-hoc comparisons revealed statistically significant differences between the beep and intensity probetype ($t = 4.593$, $p_{bonf} < 0.001$, $d = 0.760$) and between the beep and direction probetype (t = 4.919, $p_{bonf} < 0.001$, $d = 0.814$).

Motor Time Lag

One outlier was excluded from the statistical analysis here because Δt was above 650 ms for each probetype in the force tracking task. Mean Δt values for the probetypes after the exclusion of the outlier ranged from 304.087 ms (num) to 320.347 (beep) ms. After the exclusion, normal distribution was given ($p > 0.05$) for all probetypes. A repeated-measures ANOVA for the factor probetype was executed with the data for the motor time lag in the force tracking task. A statistically significant probetype effect was found ($F(4,80) = 19.623, p < 0.001, \eta^2_p = 0.495$). Post–hoc comparisons revealed that motor time lag for the probetype beep differs significantly from all other probetypes (*num:* $t = 7.789$, $p_{bonf} < 0.001$, $d = 0.266$; *int:* $t = 7.138$, $p_{bonf} < 0.001$, $d = 0.243$; *dir:* $t = 6.487$, $p_{bonf} < 0.001$, $d = 0.221$; *word:* $t = 5.835$, $p_{bonf} < 0.001$, $d = 0.199$). All other comparisons were not statistically significant ($p > 0.05$).

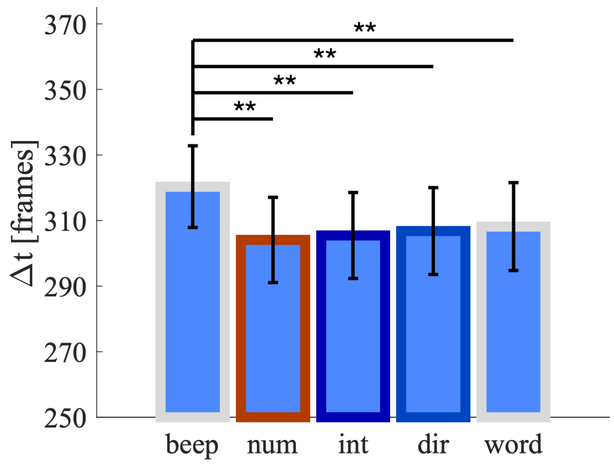

*Figure 50: Motor time lag (Δt) in frames, i.e., ms, for FRC during triple task execution with different semantic characteristics of probe stimuli in PRT. Error bars respresent standard deviation of the mean. **p ≤ 0.01.*

Motor Error

The same participant's data excluded from the motor time lag, was excluded from the analysis of the RMSE, too. Data after outlier exclusion showed normal distribution in the S-W test and residuals showed normal distribution in the Q-Q-Plots. A repeated-measures ANOVA was performed with the RMSE for the factor probetype and showed a significant effect ($F(4,80) = 198.880, p < 0.001, \eta^2_p = 0.909$).

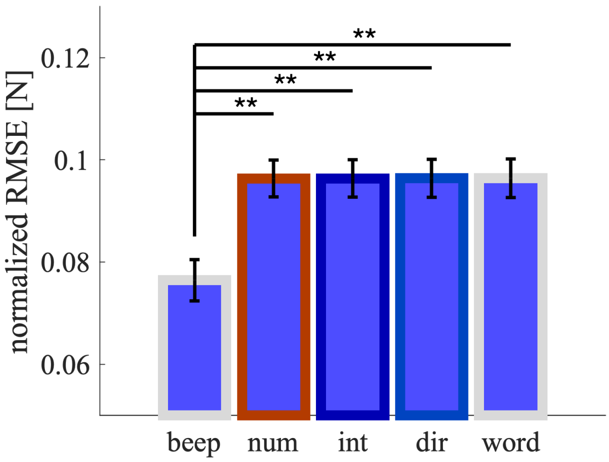

*Figure 51: Motor error (RMSE) in N for FRC during triple task execution with different semantic characteristics of probe stimuli in PRT. Error bars respresent standard deviation of the mean. **p ≤ 0.01.*

In the post-hoc comparison, RMSE in the beep probetype differed statistically from all other probetypes (*num:* $t = -22.273$, $p_{bonf} < 0.001$, $d = -1.129$; *int:* $t = -22.290$, $p_{bonf} < 0.001$, $d = -1\text{-}130$; *dir:* $t = -22.306$, $p_{bonf} < 0.001$, $d = -1.131$; *word:* $t = -22.322$, $p_{bonf} < 0.001$, $d = -1.313$). No other comparison was statistically significant.

6.5 Discussion Experiment 1

In this first study a motor-cognitive task study was conducted, adding a probe reaction time task for a triple task design. Performance measures were taken in all three tasks under single and triple task conditions for analyzing performance decrement and to test for a specific influence of the probetypes on either the motor or the cognitive task. For all three tasks a blockwise performance decrement was found for the reaction time (PRT), calculation time (CLC), motor time lag (FRC) and RMSE (FRC), indicating that performance under the triple task condition decreased in comparison to single task execution. Only for the reaction errors (PRT), no performance decrement became visible. This might be caused by the low task complexity. Since the PRT was a single-choice reaction time task, where no semantic processing was necessary, error rates could generally expected to be low.

However, reaction error was even lower under triple task, in comparison to single task execution. What needs to be considered is that for this single choice reaction time task, errors were caused by not responding to a stimulus within a maximum of two seconds. Taking this into consideration, error rates of 8 % under single task execution seem quite high. One assumption could be that those errors were caused by a lack of concentration or familiarization with stimuli, because the PRT single task blocks were executed directly before the triple task blocks only once for each stimulus type.

The probetype effect was further analyzed to test for semantic interference on a cognitive level with the auditory probe stimuli. For this, only data from the triple task execution was analyzed. Participants reacted the fastest to the beep stimulus in the probe reaction time task differing statistically from all other semantic tasks, except for the word stimuli. The numeric stimulus caused the longest reaction times, also differing from all other stimuli. This cannot be explained with the different stimulus lengths of the probe stimuli ("four" = 230 ms [shortest stimulus] and "six" = 553 ms [longest stimulus]) as they were pseudo-randomized over the whole experiment. Therefore the differences in length cancel each other out on average as a result. Also, participants did not have to wait until stimulus presentation was terminated, because no semantic processing was required in this single-choice reaction time task. The result could therefore rather be regarded in unison with the findings of Ells (1969) who reported that an additional probe task did not affect the performance of a primary task, but reflected the additional processing demands by being prolonged itself. Because the central processing demands under triple task execution might partly be attributed to the calculation task, this could indicate a lack of a specific resource for further numeric processing.

Regarding the calculation task, the beep stimulus led to the longest calculation times, differing statistically only from the intensity and direction probetypes. The beep stimulus does repeatedly differ from other semantic stimuli, even though for the calculation time it seems surprising, that it causes a decreased performance in the cognitive task. The numeric probetype did not result in higher calculation times, leading to the rejection of the specific probetype hypothesis in the calculation task. Again, this is consistent with the results by Ells (1969), reporting no performance decrement in the primary task (here calculation task), but only for the reaction times, that were specifically prolonged by numeric probetypes.

For the motor time lag and the motor error in the force tracking task, the beep probetype differed statistically from all other probetypes. Interestingly, Δt was increased for the beep probetype, while for the RMSE it was decreased in comparison to the other probetypes. This is remarkable, because a higher time lag should naturally also result in a higher RMSE because both variables are dependent. Here, the actual tracking speed decreases but leads to an increase in accuracy, nevertheless. No specific probetype could be found in the force tracking task, because the intensity probetype did not lead to higher Δt or RMSE measures. Still, the sole semantic in comparison to a beep sound itself seems to influence performance in the motor task under triple task execution, leading to a decreased time lag, but increased RMSE for semantic probe stimuli.

No expected specific probetype effect was found for the numeric probe in the calculation task and the intensity or direction stimuli in the force tracking task. But the numeric probe stimuli led to an increased reaction time in the probe reaction time task, pointing towards a specific interference of central processing demands for cognitive tasks. Another finding was that, while probe stimuli do not specifically influence performance in the motor or cognitive task, there is a general difference between semantic and beep stimuli leading to differences in performance in the other tasks.

One problem that arose during the experiment concerned the bimanual task execution. The calculation task is answered with a computer mouse in the right hand while the probe reaction time task is answered with the FSR in left hand. Many studies have discussed whether a task interference in bimanual dual-tasking takes place on the motor-programming or response-selection stage (Pashler, 2004, pp. 155–189). Some research provides evidence, that those two interference effects do not exclude each other but can both be affected during bimanual task execution. The study of Stanciu et al. (2017) stresses the importance of the cognitive demands that a task poses to be the most crucial thing for predicting interference effects: planning the movement of one hand effects simultaneous movement planning and execution of the other hand. This interference increases when cues are symbols (i.e., letters) in comparison to spatial cues, which is why the translation of the cue and the process of response selection are suggested to cause this interference rather than an ongoing cross-talk at the motor programming level (Stanciu et al., 2017). This could provide another explanation for the prolonged reaction times to numeric probe stimuli. Therefore, in the next experiment, another response mechanism will be chosen for the calculation task.

It was also a problem that if participants spotted an error in the result comparison phase and then clicked the computer mouse, the next task would directly appear, rather than waiting the 2500 ms that this phase would have lasted otherwise. This way, it was an advantage if a wrong result was shown. Also, participants could have used this as a strategy when they did not like to solve a more difficult task (more burrows) by skipping to the next task earlier. This behavior could have been controlled by checking the error rates of the calculation task, which could not be analyzed here.

Another point of criticism was that only single and triple task execution was executed in this experiment, while no dual task conditions were carried out. Thereby, the interactions of the tasks cannot clearly be separated from one another. For this reason, the next experiment will provide all possible task combinations as dual task execution, additionally.

One further aspect to consider is, that during triple task execution, two continuous and one discrete task are executed simultaneously. By regarding performance over a whole task block, fine-grained changes in central processing demands caused by differing task requirements are averaged out and thereby neglected. Therefore, in the next chapter, a new approach of analyzing data that was collected over a specific period of time, is presented.

6.6 Conclusion Experiment 1

Under single-task conditions, all three tasks are executed undisturbed. This results in a baseline of requirements in the dimensions of time and PCU (Figure 52), which can then be compared with the triple task execution.

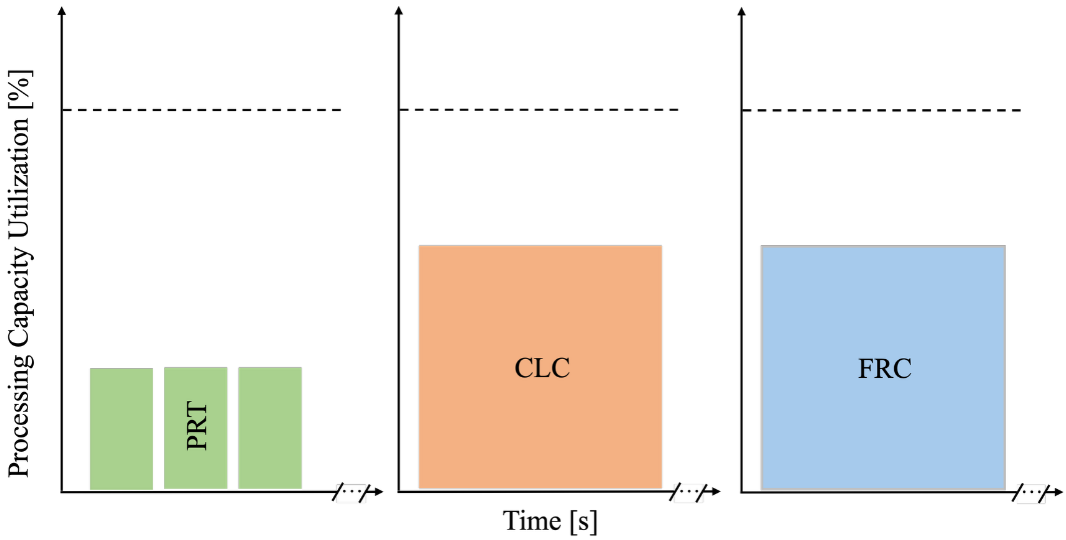

Figure 52: Outline of task requirements in the times regime model in the dimensions time and PCU for PRT, CLC, and FRC under single task execution.

During triple task execution, it was evident that all three tasks are extended in the time dimension if they are processed simultaneously within a task block (see Figure 53). The response errors do not appear to exceed the PCU limit, but the FRC requirements do. Unfortunately, no statement can yet be made for the PCU of the CLC due to the lack of calculation error data.

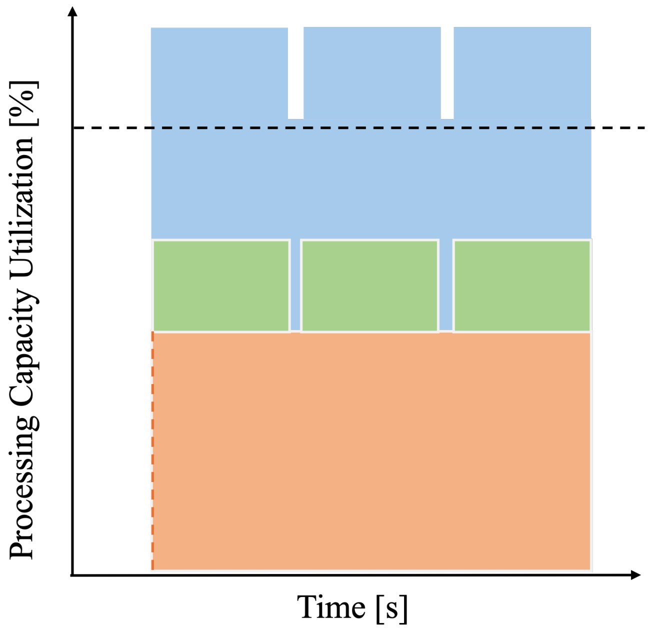

Figure 53: Outline of task organization and adaptations in time and PCU dimensions in the times regime model for PRT, CLC, and FRC under triple task execution.

To compensate for the missing area i.e. capacity resulting from the error-free PRT execution, either the motor or the cognitive task would then have to be adapted accordingly at these times. The extent to which this mechanism takes place can only be speculated at present, as the CLC data are not available in the PCU dimension and it is not possible to decipher the exact temporal distribution of capacity with the analyses performed to date. The next chapter will therefore focus on enabling a more precise temporal resolution of the triple task block so that further insights can be gained into the distribution of the required resources.

In addition, it was shown that probe stimuli with semantics seem to differ in their requirements in the Time x PCU area compared to a non-semantic stimulus. To which extent however cannot be determined yet, as the semantics resulted in different directions depending on the dependent variable under consideration.

7 Event-related Analysis of Experiment 1

In the first experiment, the numeric probe stimulus differed from the other probetypes by leading to increased reaction times. Neither the cognitive nor the motor task showed a specific probetype effect leading to an interference with the numeric stimulus in the calculation task and the intensity stimulus in the force tracking task. The mere semantic of the probe stimuli could however lead to a specific interference. For both the Δt and RMSE the beep stimulus led to a statistically different performance in comparison to all other probetypes. This inconsistent result pattern could point to a major concern regarding multiple-task studies.

Traditionally, in multi-tasking studies, task performance is calculated for a whole trial block, which in the first experiment lasted 60 seconds. Thereby performance is averaged over a large time span. Especially during triple task execution this could neglect the fine-grained changes in performance during undergoing certain discrete events caused through the combination of three different tasks with different time regimes. These discrete events would allow for a much more precise analysis of performance during multiple task execution than averaging performance over a full task block. To observe interference during multiple task execution through performance changes, an event-related analysis will be established as third research line (R3) to enable the determination of task performance during time discrete events.

7.1 Description of Events

In the first experiment, several events can be defined during triple task execution. As a first step, events are differentiated at the phases in the calculation task, namely either the calculation phase (Calc) or the result comparison phase (Comp). Since the calculation task has a continuous nature, either one of the phases is always addressed. This also applies to the force tracking task, which also has a continuous nature. The force component (Force) is therefore always part of any defined event. The only discrete task is the probe reaction time task. Here the events can either hold a probe stimulus and its reaction (Probe), or not (Only). These combinations lead to four different events (CalcOnlyForce, CompOnlyForce, CalcProbeForce, CompProbeForce). The CalcProbeForce event is shown in Figure 54. Here, the probe stimulus and its reaction take place in the calculation phase of the calculation task.

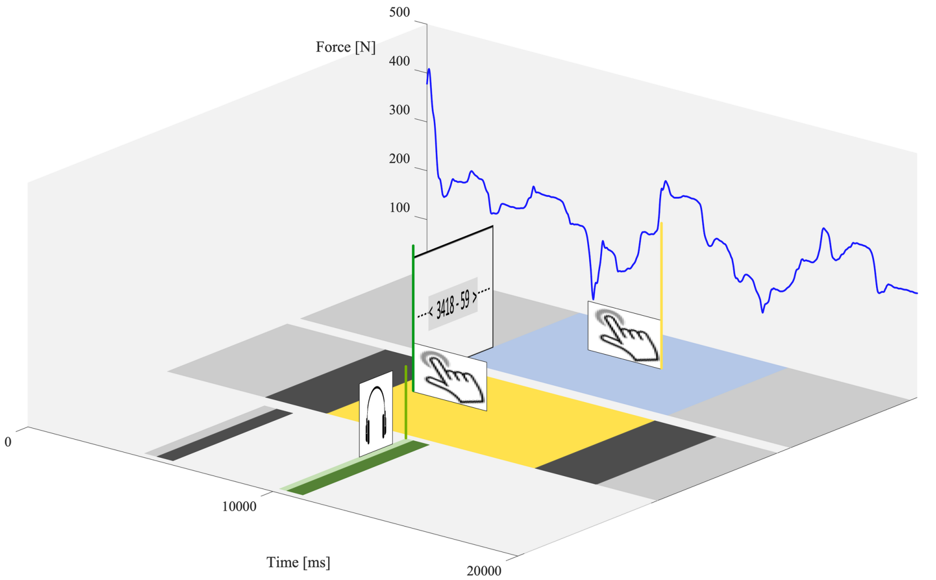

Figure 54: CalcProbeForce event in the first experiment for triple task execution. The probe stimulus and its response take place during the calculation phase of the calculation task, while the force tracking task is also executed.

In Figure 55, the CompProbeForce event is shown, where the probe stimulus and its reaction take place in the comparison phase of the calculation task.

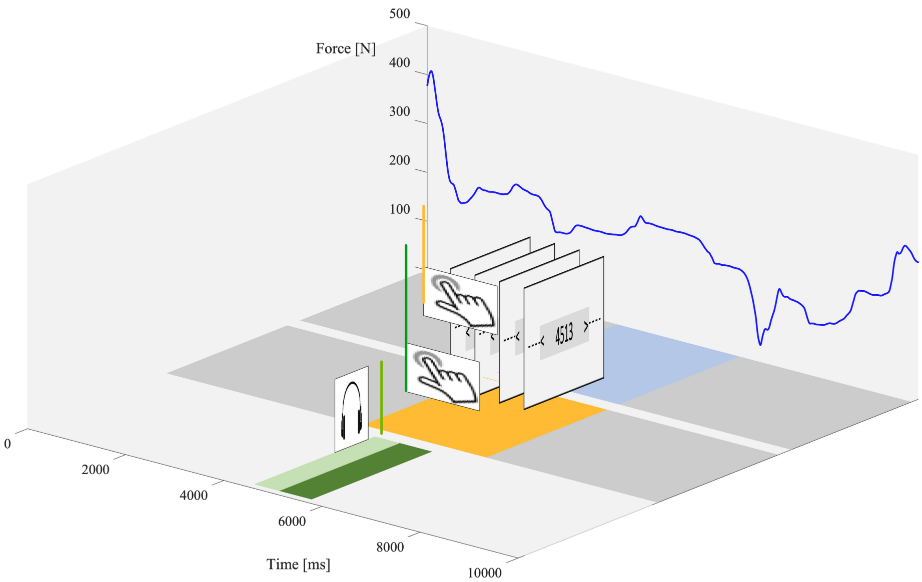

Figure 55: CompProbeForce event in the first experiment for triple task execution. The probe sitmulus and its response take place during the comparison phase of the calculation task, while the force tracking task is also executed.

If no probe stimulus and reaction underwent in the calculation or result comparison phase, the events are titled CalcOnlyForce or CompOnlyForce. If the probe stimulus and the reaction did not take place during that same phase of the calculation task, an algorithm checked, if more than 50 % of the time span from stimulus onset to reaction onset fell into the phase of the stimulus onset. If this was the case, only the previous calculation phase with the probe stimulus onset was labeled as a ‘Probe’ event. If more than 50 % of the time span fell into the next phase, where the response was executed, this phase was also labeled as a ‘Probe’ event.

7.2 Regions of Interest

Another fine-grained differentiation can further be executed accounting for the performance in the force tracking task around the probe stimuli. The results for the probetype analysis in the first experiment have shown, that performance in the force tracking task (Δt and RMSE) differs between different probetypes. Thus, it seems reasonable, that the probe stimulus, as well as the reaction to it can lead to discrete performance changes in the force tracking behavior. To further entangle task interference here, three different regions of interest (ROI) are defined (see Figure 56). The PreStim interval, is the interval before probe stimulus onset. This interval is foremost used for comparing the RMSE to the following two intervals after stimulus and response onset. The PostStim interval is the time span after the probe stimulus was played, but before a response to the stimulus was measured.

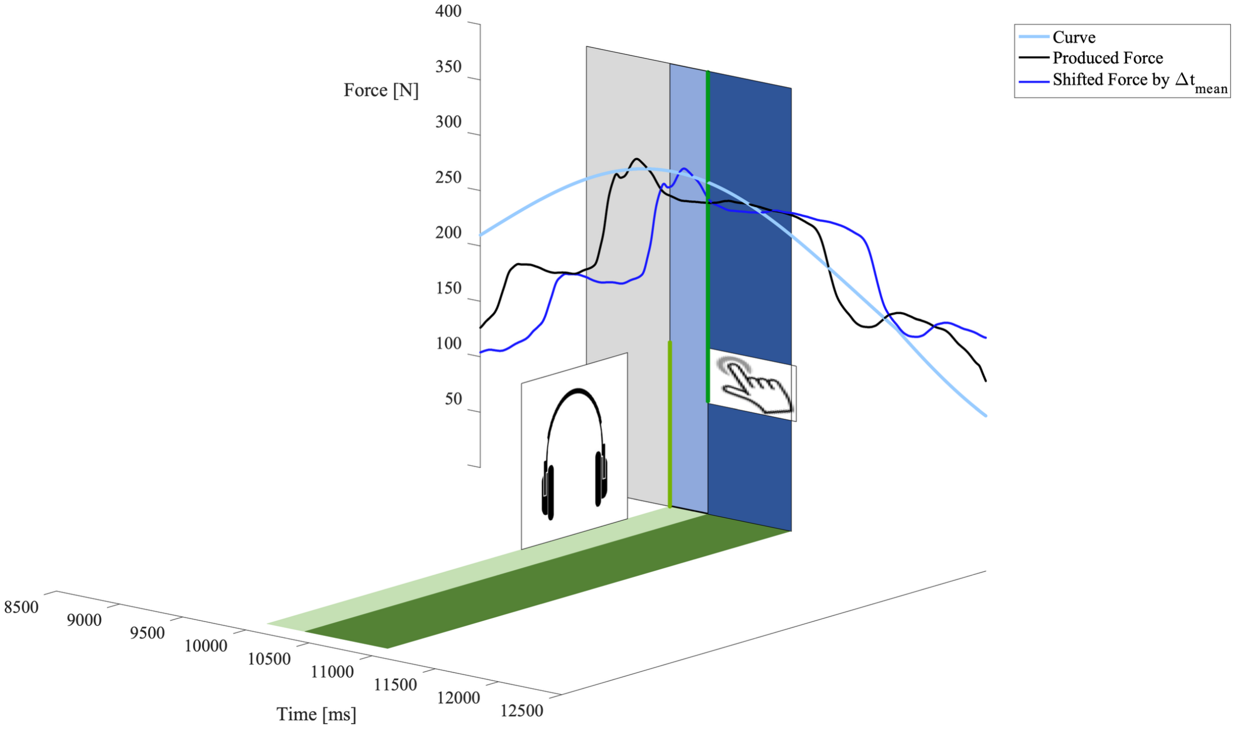

Figure 56: Three regions of interest during triple task execution (CLC not depicted here). Discrete probe stimuli enable differentiation of force tracking performance into three regions of interest: prior to probe stimulus (PreStim), after stimulus onset but before response registration (PostStim) and directly after response execution (PostReact) was measured. Δt_{mean} was calculated over all participants in the three experiments.

In the PostReact interval, the RMSE is analyzed for a short time interval after probe response execution. To calculate the RMSE as a performance measure during the three ROIs, the force curve is first shifted on the time axis by the overall Δt_{mean} = 376 ms (rounded by 375.582 ms), that was calculated over all participants in all three experiments. This was done, because with the existence of this motor time lag, the caused interference by a probe stimulus will only become visible after the motor time lag. For the PreStim and PostReact interval, a length of 661 ms (rounded by 660.800 ms) is used, as this is the RT_{mean} over all participants and experiments. This time span is suggested as the best estimator during which a task interference should be measurable because it is also the time that the stimulus perception, response selection and response execution take in the probe reaction time task.

During the PostStim interval, performance in the motor task can be analyzed during stimulus perception, response selection and response execution to the probe task. The PostStim interval allows for the determination of performance in the motor task during the effect monitoring phase of the probe reaction time task.

7.3 Results of the Event-Related-Analysis for the First Experiment

The event classification was done for the triple task blocks in the first experiment. If the analysis is done by also differentiating between the 5 different probetypes, only very few cases are detected for each event with corresponding probetype. To receive meaningful statistics with sufficient sample sizes that have a small variance, the probetype is neglected in the event-related analysis.

Reaction Time

For the analysis of the reaction time as dependent variable, two events could be tested under triple task execution: the CalcProbeForce and the CompProbeForce events. Two outliers could be detected in the boxplots. The produced reaction times were however within the 2 seconds boundaries and therefore considered as valid. To test differences for the reaction times between the two events, data of 21 participants could be included in the analysis. CompProbeForce data of one participant was missing, because there were no such events to be found throughout the whole experiment.

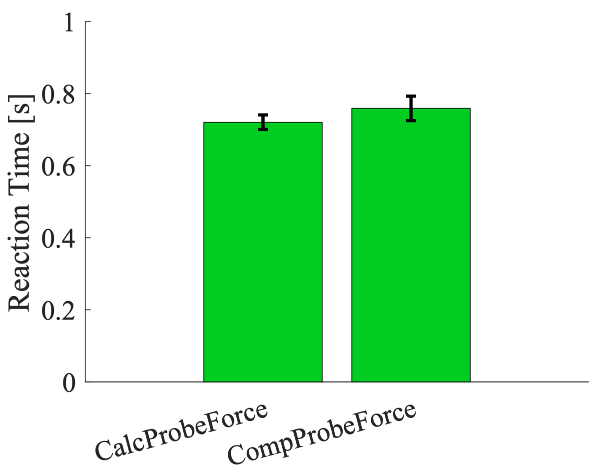

Figure 57: Reaction times for PRT during triple task execution in the two events CalcProbeForce and CompProbeForce differentiating between the calculation and the result comparison phase in the CLC task.

A two-sided hypothesis was tested, resulting in no differences between the two events (*t(20) = -1.207, p = 0.241, d = -0.263*).

Reaction Error

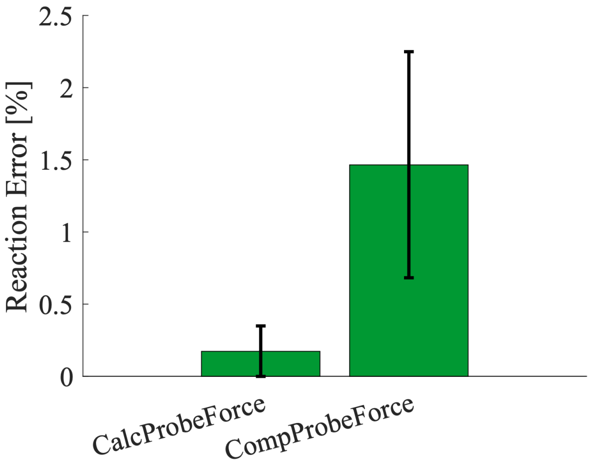

Figure 58: Reaction error for PRT under triple task execution for the differentiation in the two events CalcProbeForce and CompProbeForce.

For the reaction error, the same events could be tested. Normal distribution was violated for both events ($p < 0.001$); therefore the Wilcoxon signed-rank test was calculated. No statistically significant difference was detected for the two-sided hypothesis (W = *1.000, p = 0.106, r = -0.867*).

Calculation Time

Calculation times for the events CalcForce and CalcProbeForce are normally distributed (S-W test CalcForce: *p = 0.149*; CalcProbeForce: *p = 0.989*). A t-test for dependent samples is calculated with the one-sided hypothesis that calculation time is increased in the CalcProbeForce condition caused by the additional probetype. The result shows a statistically significant difference ($t(21) = -10.497$, $p < 0.001$, $d = -2.238$).

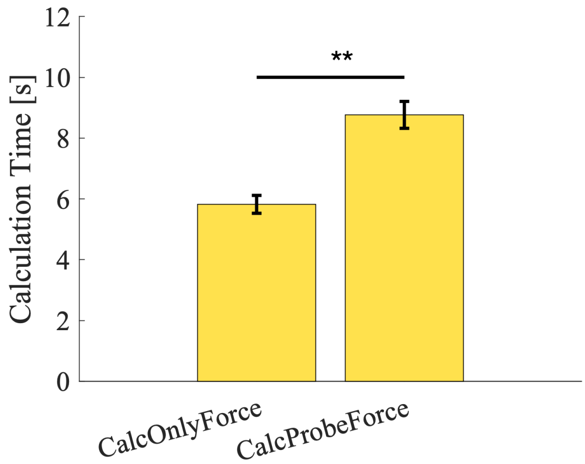

Figure 59: Calculation time in CLC for triple task execution and the events CalcOnlyForce and CalcProbeForce differentiating whether a probe stimulus occured during the calculation phase, or not.

Motor Error

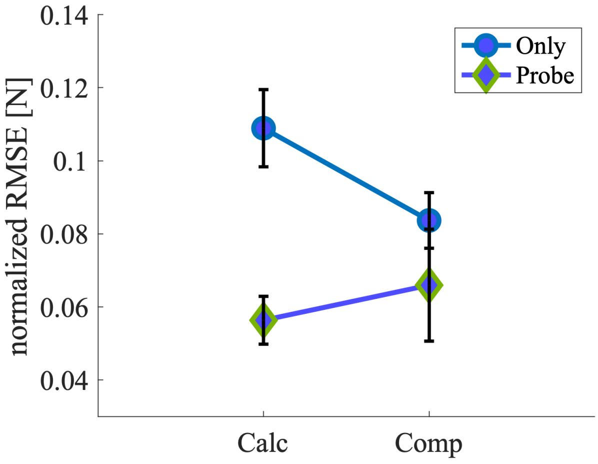

Figure 60: Motor error (normalized RMSE) in FRC during triple task execution as four different events. In CalcOnlyForce and CompOnlyForce no probe stimulus was presented during the calculation or the result comparison phase of the calculation task, whereas in CalcProbeForce and CompProbeForce a probe stimulus was registered.

For the RMSE, events during different phases of the calculation task could be differentiated (Calc/Comp), as well as if a probe stimulus was presented during that phase, or not (Only/Probe). The four events CalcOnlyForce, CompOnlyForce, CalcProbeForce and CompProbeForce are used for the analysis. Two different outliers were identified in the CompOnlyForce and the CompProbeForce events. Both had plausible RMSE values and were therefore included in the further analysis. A 2 x 2 repeated-measures ANOVA was calculated for the factors calculation phase (Calc x Comp) and the factor probe stimulus (Only x Probe). A significant main effect was found for the factor probe stimulus ($F(1,20) = 4.901$, $p = 0.039$, $\eta^2_p = 0.197$). No statistically significant main effect was found for the calculation phase ($F(1,20) = 3.234$, $p = 0.087$, $\eta^2_p = 0.139$), or for the interaction of both factors ($F(1,20) = 1.209$, $p = 0.285$, $\eta^2_p = 0.057$).

Regions of Interest

The motor error was further differentiated into three different ROIs in relation to the probe stimuli. Here, the events CalcProbeForce and ComProbeForce were each considered with a PreStim, PostStim and PostReact interval. A 2 x 3 repeated-measures ANOVA was therefore calculated for the factor calculation phase (CalcProbeForce x CompProbeForce) and the factor ROI (PreStim x PostStim x PostReact) for 21 participants.

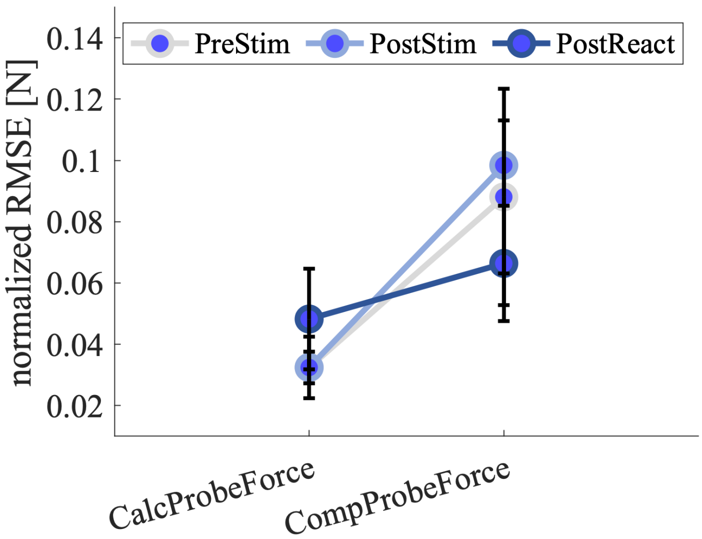

Figure 61: Motor error (normalized RMSE) in FRC under triple task execution for the CalcProbeForce and CompProbeForce events differentiated into three different ROI: PreStim, PostStim and PostReact interval.

Neither the main effect calculation phase ($F(1,20) = 4.162$, $p = 0.055$, $\eta^2_p = 0.172$), nor the main effect ROI ($F(1.160,23.203)$

$= 0.114$, $p = 0.776$, $\eta^2_p = 0.006$) reached statistical significance. The interaction of Event x ROI ($F(1.149,22.983) = 1.725$, $p = 0.204$, $\eta^2_p = 0.079$) was also not significant.

7.4 Conclusion for the Event-related Analysis of Experiment 1

To allow for a better time resolution of discrete effects for multiple task execution interference, an event-related analysis was conducted. For the dependent variables reaction time and reaction error, a differentiation between the two calculation phases (Calc/Comp) was done. No statistical difference was found between the CalcProbeForce and CompProbeForce events, neither for the reaction time, nor for the reaction error. A tendency towards higher reaction errors during the comparison phase of the calculation task could be visualized, however. It would certainly be profitable to shed light on the extent to which the reaction times and errors differ for the different semantic probetypes during the two events. Reaction times to numeric probe stimuli were prolonged in comparison to the others in the probetype effect analysis. Thus, it would be useful to allow for a differentiated statement to which extent the different calculation phases might affect this behavior as well. Unfortunately, there were not enough events, especially for CompProbeForce, to differentiate further between the different probetypes.

For the calculation time, the two events CalcOnlyForce and CalcProbeForce could be analyzed, differing in the presentation or absence of a probe stimulus. Here, calculation time differed between the two events, leading to a prolonged calculation time for the events, where a probe stimulus was presented. This was expected, because the stimulus presentation poses an additional task that needs attentional resources that may lack in the cognitive task, consequently. While the calculation time was prolonged for the CalcProbeForce event, it was comparable to single task execution (6.330 ± 1.329 s) of the calculation task in the CalcOnlyForce (5.822 ± 1.380 s) event. This can be interpreted as a very discrete change in performance when a probe stimulus is presented. Calculation time increases only during those events and then returns to single task execution level during the triple task execution where no probe stimulus is presented during a calculation phase. It seems that there is no further cognitive load for the mere expectancy of stimuli that might be presented. What is even more remarkable is, that the force tracking that is continuously executed during the triple task execution seems to not increase calculation time further.

For the motor error, effects on the tracking performance through the calculation phase (Calc/Comp) and through the probe stimulus availability (Only/Probe) were examined for the events CalcOnlyForce, CompOnlyForce, CalcProbeForce and CompProbeForce. Only the probe stimulus availability showed an effect on the motor performance. Surprisingly, RMSE was decreased for events where a probe stimulus was present in comparison to the absence of the auditory stimulus. This could indicate that for the motor task, the probe stimulus might lead to a performance boost. An even more small-layered differentiation could be made for the motor error. For the events that held a probe stimulus (CalcProbeForce and

CompProbeForce), three different ROIs could be determined to examine, whether performance in the motor task would differ before (PreStim) stimulus onset, for the time span between stimulus onset and response onset (PostStim) or for a close time span after response execution (PostReact). For those ROIs no difference in the RMSE could be detected. This does not seem surprising since the probe stimulus did not lead to a performance decrement under triple task execution in comparison to the absence of a stimulus (Only/Probe) but rather to a decreased motor tracking error. Therefore, no specific decrement in performance should be expected for the ROIs.

7.5 Summary for the Event-related Effects

The event-related analysis (ERA) was executed for all dependent variables of the first experiment, and additionally for the ROIs for the motor error. Results showed that the comparison of events where a probe stimulus was given (Probe) with events where no stimulus was presented (Only) led to changes in performance in the cognitive task. Prolonged calculation times could be examined when a probe stimulus was given in the cognitive calculation task.

This irregular pattern of result could be interpreted in a way that calculation and comparison phase do not differ in the amount of cognitive resources that they bind. Because there were differences in reaction time, nor motor error for the two calculation phases. The probe stimulus seems to influence performance in the cognitive task at discrete events. For the cognitive task it leads to a decrement in performance while for the motor task it seems to boost motor performance regarding the data descriptively. However, this effect was not statistically significant in inference statistics. These results can now be classified in the times regime model as follows (see Figure 62).

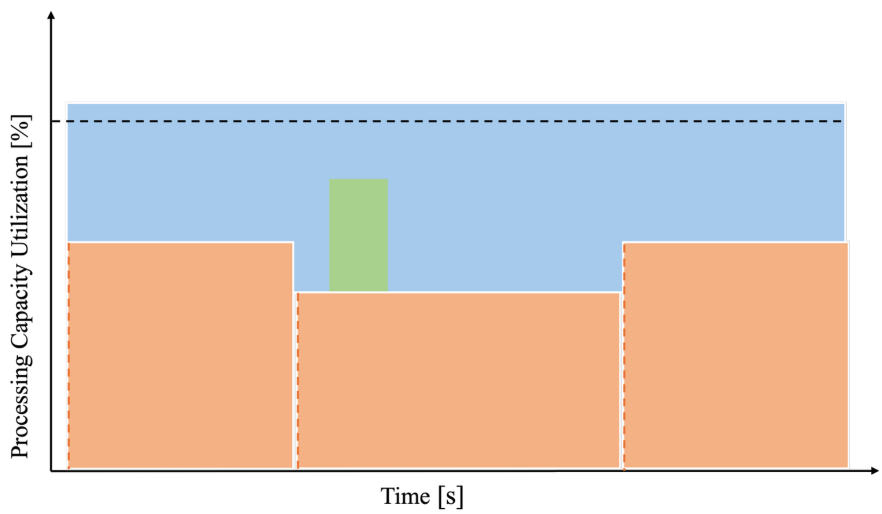

Figure 62: Sketch of capacity distribution during triple task execution in the time and PCU dimensions of the time regimes model for PRT, CLC, and FRC. Calculation time is specifically prolonged when a probe stimulus has to be responded to but decreases after response execution back to single task level.

It is known from the classical analysis of the performance decrement that Δt and RMSE increase for the FRC task. The reaction time in PRT also increases, although the reaction error does not increase. The calculation time becomes longer in the triple task condition. However, this is only the case when a probe stimulus is presented and must be answered. Once this discrete task is completed, the computing time for CLC drops back to the initial level in the single task. Unfortunately, it is still not possible to make any statements about the calculation error, as the relevant data could not be analyzed.

To conclude, the event-related analysis can be useful to depict fine performance changes to discrete events during continuous task execution blocks. Although only one effect of probe stimuli on the calculation time in CLC could be detected so far, this analysis should be integrated in the following experiments to detect further time-critical interactions of the tasks and a flexible distribution of the capacity.

8 Experiment II - Vocal Response to Cognitive Task

The purpose of this experiment was to introduce small alterations in all three tasks that may have ostensibly led to the absence of the hypothesized probetype effect in the cognitive and motor task in experiment 1. Thus, the calculation task was no longer answered using a computer mouse, because during the first experiment participants reported that they often mixed up which hand should be used to react to the probe stimuli and which to respond to the calculation task. Wickens and Liu (1988) also reported that responses to a discrete manual task with one hand could disrupt the continuous tracking response of the other hand. To circumvent this bimanual interference, a microphone was now used to record verbal answers instead. Another alternation was that the result comparison phase now always lasted 2.5 seconds regardless of whether the participants indicated a spotted error in the displayed result or not. This way, it was no longer advantageous for participants to indicate errors in the result comparison phase because this would mean that a new calculation task would appear immediately. Probe stimuli were altered to account for a possible spatial cognitive representation of the motor task. Thereby, the spoken numbers remained the same, the semantic probe pair aiming at the quantity of the motor task were restored ("stronger" and "weaker") and additionally, the spatial pair ("up" and "down") was added. Force curves were also altered to have a higher difficulty and less predictability.

To further determine how the allocation of cognitive processes affects task performance, the cognitive, motor and probe reaction time tasks were not only executed under single and triple task conditions, but also under all three possible combinations of dual task execution. In this context, dual task conditions can help uncover which task combinations may perform well simultaneously, and for which combinations performance decreases.

8.1 Methods Experiment 2

Based on the effect found in the event-related analysis in the computing time between CalcProbeForce and CalcOnlyForce, an a priori computation of required sample size was performed. Assuming a power of 0.7, a total sample size of 21 participants was required.

8.1.1 Participants

A total of 23 participants (11 female, 13 male), with a mean age of 23.7 (± 3.04) years conducted the experiment. Participants were undergraduate students and had to fulfill the same inclusion criteria as described in 6.1.1. Additionally, they should not have participated in the first experiment. Participants received course credit for their participation.

8.1.2 Probe Reaction Time Task

The auditory probe stimuli were the two spoken numbers from the first experiment and pairs of words expected to interfere with the motor task (see Table 3). Since the stimuli used in experiment 1 did not elicit any specific semantic effects in the cognitive or motor task, the semantics of the stimuli targeting the motor task in particular are adapted once again. Stimuli that previously contained intensities and directions are now replaced by two other semantic characteristics. The latter sought to either influence the spatial ("higher"&"lower") or the quantitative ("stronger"&"weaker") aspect of the task.

Table 3: Semantic probe stimuli for the probe reaction time task in experiment 2.

Probetype	Stimulus	Stimulus Length (ms)	Targeted Performance
Numeric (num)	Vier (four) Sechs (six)	552.0 ms 551.8 ms	cognitive
Spatial (spa)	Höher (higher) Tiefer (lower)	531.5 ms 524.1 ms	motor
Quantitative (qua)	Stärker (stronger) Schwächer (weaker)	781.4 ms 725.2 ms	motor

Buccino et al. (2001) could show that listening to action-related sentences increased activation in the motor system. It can therefore be assumed that the spatial information specifically addresses the visual part of the task, which consists of recording the position of the bar on the screen and deducing whether it needs to be moved *higher* or *lower*. This is then done by applying either higher or lower forces to the leg press by having the participants press *stronger* or *weaker*. The beep and unspecific word stimuli are omitted in this experiment. Probe responses were measured with the FSR in the same procedure used in experiment 1.

8.1.3 Cognitive Task

For this task, a condenser microphone was now attached to the participants heads. The result comparison phase always lasted 2.5 seconds so that an answer in the second phase would not be advantegous in the sense of the number of accomplished tasks during one block. Apart from these small changes, the task remained the same as in experiment 1.

8.1.4 Motor Task

For the force tracking task, the leg press from experiment 1 was used. The main aspect that was varied from the first experiment were the tracked curves. The underlying force curves were altered to be more complex and less predictable in comparison to experiment 1. Therefore for every block, a new curve was generated randomly from pre-defined parameters of sine and cosine curves. This was done in a manner, that between 20 and 40 seconds, the curve remained the same within each participant. The first and last 20 seconds, however, were altered for every block and within each participant. This way, the curves always held a habituated and two new segments to allow for the control of a sequence learning for the middle part.

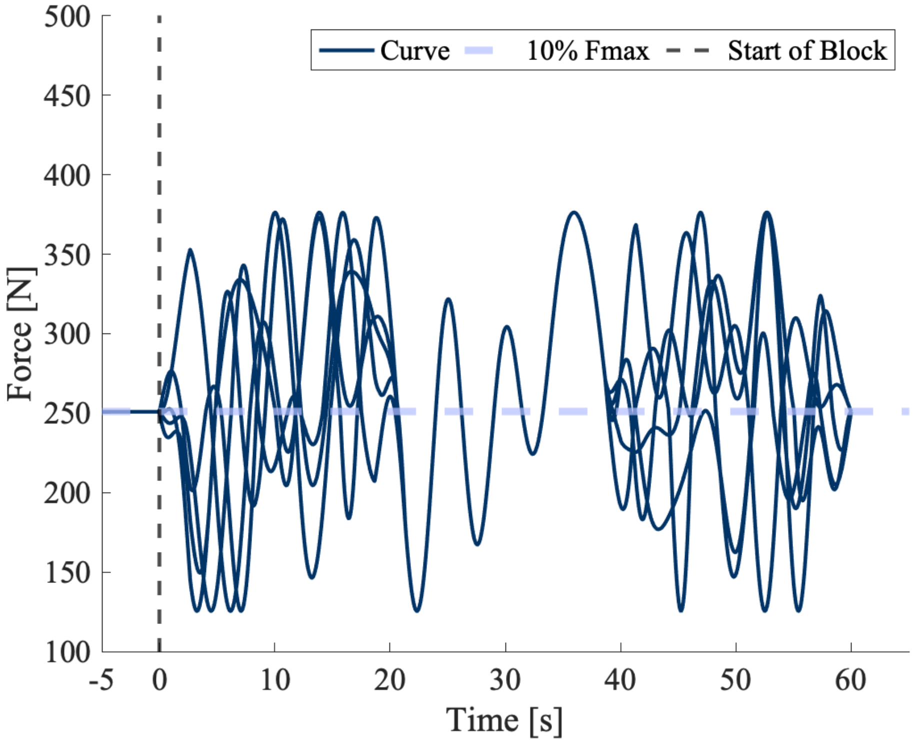

Figure 63: Exemplary force curves for the second experiment. First five generated curves of participant 1.

The first five seconds were always taken as an adjustment phase before the actual start of the block. Here, a countdown was displayed, and the set force (10% of each participant's individual maximum force) was indicated through arrows at the left and right side of the screen.

Participants were to adjust the bar that they could move with their leg force to the arrows within the countdown. Data from the adjustment phase were not used for further analysis, as the produced forces would over- or undershoot the set forces due to the initiation of force application.

8.2 Procedure Experiment 2

The experiment included three consecutive days of data acquisition.

Day 1

On the first day, participants' maximum force was measured at the leg press. For this, the maximum force had to be built up and shortly kept within an interval of 5 seconds. This procedure was repeated three times, and the maximum force was then determined as the maximum of the three trials.

Then, four cycles of different task combinations were executed, each consisting of 9 task blocks. One task block had a duration of 60 seconds. Every participant started with executing the motor (STm) and then the cognitive (STc) task under single task conditions once. After this, these two tasks were combined into a motor-cognitive (DTmc) dual task block. This dual task block was repeated five times total. After that, first the cognitive and then the motor tasks were again executed as single task conditions once. This first cycle of tasks was the same for all participants, as it served as a familiarization with the two more complex tasks.

For the three cycles that followed, a similar pattern of task configuration was chosen. In the middle of a cycle, there were now either the dual task conditions, combining the motor or the cognitive task with the probe reaction time task, or the triple task condition, combining all three tasks. The dual task condition would always be repeated five times in a row; the triple task condition would be repeated three times. Before and after those, the involved tasks from DT and TT conditions were all carried out as ST once. For each of the three rounds, a new probetype was chosen the whole round, remaining constant throughout the different conditions.

Day 2 & Day 3

On the second and third day, three cycles of different task combinations were executed, each consisting of 9 task blocks. The protocol remained similar to the last three cycles of the first day; simply pseudo-randomizing the DT and TT conditions and probetype combinations over all participants. An exemplary testing protocol is shown in Table 4.

Table 4: Exemplary protocol for experiment 2 for the first participant. A cycle consists of nine task blocks where all tasks are first executed under ST conditions. Respectively, five DT, or three TT task blocks are executed, followed by every task as ST again. For example, the DTmc cycle contains the following blocks: STm, STc, 5x DTmc, STc, STm. Where m = motor task FRC, c = cognitive task CLC, p = probe reaction time task PRT.

Day 1		**Day 2**		**Day 3**	
Cycle	**Probetype**	**Cycle**	**Probetype**	**Cycle**	**Probetype**
DTmc	-	-		-	
DTmp	num	DTcp	num	TT	num
DTcp	spa	TT	spa	DTmp	spa
TT	qua	DTmp	qua	DTcp	qua

8.3 Data Preprocessing Experiment 2

Data processing of the reaction time data from the FSR and the force data from the strain gauge used for the force tracking task remained the same as in the first experiment. The vocal data from the microphone that was used in this experiment to respond to the calculation task was already processed online. For this, an individual loudness threshold was determined for every participant before the start of the experiment. The initial threshold was set to 0.9 V and then adapted to each participant's preferred speaking loudness. In the experiment, a response to the calculation task was measured whenever the microphone input exceeded the threshold.

Data of two participants had to be excluded from further analysis, because the microphone data was missing due to technical problems, diminishing the sample size to 21 participants.

8.4 Results Experiment 2

Following the approach from the first experiment, the blockwise performance decrement will first be analyzed, comparing the single, dual, and triple task execution with each other to replicate the performance decrement research. Thereafter, the semantic probetype effect will be examined for all dependent variables under multiple task execution. Lastly, the event-related analysis will follow to gain a deeper understanding of the task interference at discrete times during continuous task execution.

8.4.1 Blockwise Performance Decrement

To test for the blockwise performance decrement, repeated measures analyses of variance (rmANOVA) with the factor task load, indicating which of the single-, triple-, and the three dual-task conditions were executed for each dependent variable, are calculated. For the rmANOVA only a two-sided hypothesis can be tested. In the post-hoc analyses, a one-sided hypothesis would be testable. However, for the different dual task combinations, no specific effect is hypothesized because it is unclear, what task combinations lead to less capacity utilization in comparison to the others.

Reaction Time

For the reaction time, a rmANOVA for the factor task load with the factor levels STp, DTmp, DTcp, and TT was calculated. Reaction time data of one participant was missing due to technical problems. Thus, only 20 participants are included in the analysis. Two outliers were detected that still produced values below two seconds and were therefore left in the analysis. Although the SW test for the reaction time under the four different conditions showed a violation of normal distribution for two factor levels (*ST: $p < 0.001$, DTmp: $p = 0.004$, DTcp:*

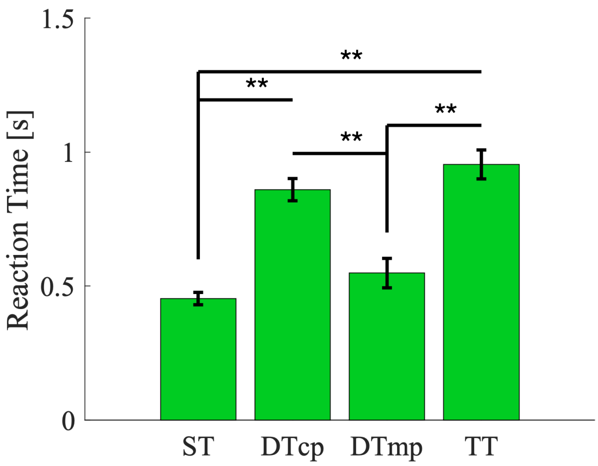

Figure 64: Mean reaction times in PRT for different task load conditions. Error bars represent standard deviation of the mean

$p < 0.001$, TT: $p = 0.374$), the Q-Q-plot indicated that normal distribution of residuals is reasonable and only violated through the outliers. Sphericity was also violated ($p = 0.010$). A significant main effect for the factor task load was found with $F(1.981, 37.643) = 30.588$, $p < 0.001$, $\eta^2_p = 0.617$. The post-hoc pairwise comparison revealed statistically significant differences between ST and DTcp, ST and TT, DTmp and DTcp, and between DTmp and TT (see Table 5). All other comparisons were not statistically significant.

Table 5: Parameters for post-hoc pairwise comparisons for statistically significant task loads.

		t-value	p_{bonf}	Cohen's d
ST	DTcp	-6.523	<0.001	-1.964
	TT	-8.173	<0.001	-2.460
DTmp	DTcp	-4.995	<0.001	-1.504
	TT	-6.646	<0.001	-2.001

Reaction Error

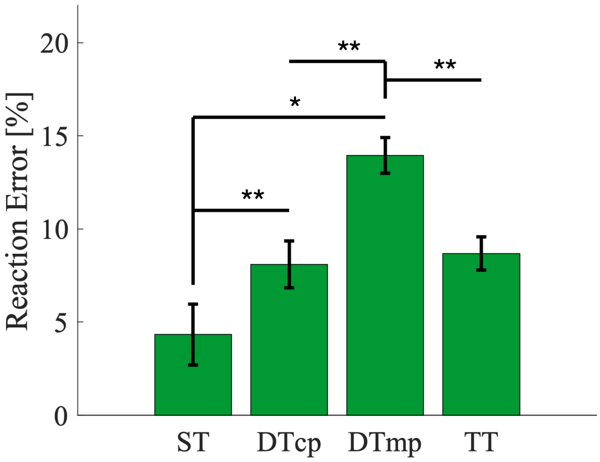

Figure 65: Mean reaction error in PRT for different task load conditions in experiment 2.

For the reaction error, the same factor levels as for reaction time are tested in a repeated-measures ANOVA. Normal distribution was violated only for single task execution and the dual task condition, combining the cognitive with the probe reaction time task. Q-Q plots revealed that normal distribution of the residuals could be accepted except for one to three outliers per factor level. Therefore, the ANOVA was calculated, nevertheless. A significant main effect for the factor task load (ST x DTmp x DTcp x TT) was found ($F(3,57) = 19.078$, $p < 0.001$, $\eta^2_p = 0.501$). All pairwise comparisons were statistically significant (see Table 6), except for the comparison of single task execution with the triple task and DTcp with the triple task.

Table 6: Statistical parameters of post-hoc pairwise comparison for reaction error in PRT under different task load conditions.

		t	p_{bonf}	Cohen's d
ST	DTmp	-7.634	<0.001	-1.762
	DTcp	-3.629	0.011	-0.690
DTmp	DTcp	5.166	<0.001	1.072
	TT	4.775	<0.001	0.964

Calculation Time

For the calculation time, the repeated-measures ANOVA with factor task load was conducted for the factor levels STc, DTmc, DTcp and TT. Normal distribution was only violated for DTcp ($p = 0.024$). The repeated-measures ANOVA showed no significant main effect for the factor task load ($F(1.984,57.650) = 2.589$, $p = 0.088$, $\eta^2_p = 0.115$). A post-hoc calculation of power was executed using GPower 3.1 (Faul et al., n.d.) to test whether the sample size was large enough to enable the detection of a task load effect on the calculation time. For this the correlation among repeated measures was first calculated by using a Pearson correlation for every possible combination of factor levels. Correlation coefficients were then transformed into z-values using a Fisher-z-transformation. The mean of these z-values was then calculated and retransformed into r^2, resulting in a value of 0.867. With this, a power of 0.961 was calculated for the rmANOVA, which is sufficiently high to keep the Nullhypothesis, that there is no difference between the task load conditions for the calculation time.

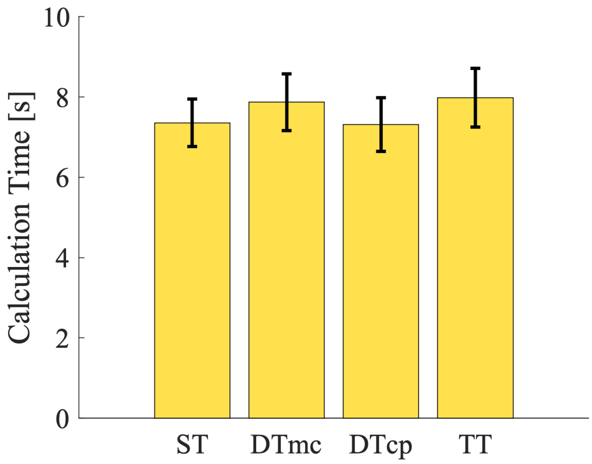

Figure 66: Mean calculation times for CLC under different task load conditions.

Calculation Error

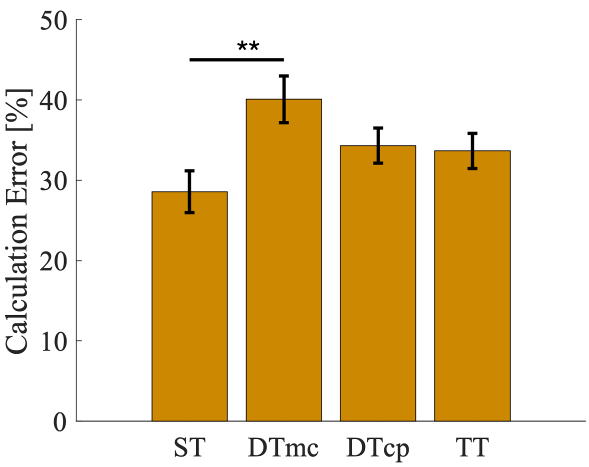

Figure 67: Mean calculation error of CLC during different task loads in experiment 2.

A repeated-measures ANOVA with factor task load was done with the factor levels STc, DTmc, DTcp and TT. Normal distribution was given over all task load conditions. The analysis revealed a significant main effect for task load ($F(3, 60) = 6.389$, $p < 0.001$, $\eta^2_p = 0.242$). In the Tukey HSD only the comparison of ST and DTmc showed a statistically significant difference ($t = -4.367$, $p_{bonf} < 0.001$, $d = -1.008$).

Motor Time Lag

For the repeated-measures ANOVA with Δt values for the factor task load with the factor levels STm, DTmc, DTmp and TT, normal distribution was given for all conditions ($p > 0.05$). Sphericity was however violated ($p = 0.015$). A significant main effect ($F(2.066, 41.328) = 5.468$, $p = 0.007$, $\eta^2_p = 0.215$) was revealed. Pairwise post-hoc comparisons with the Tukey HSD showed significant differences between ST and DTmc ($t = -3.404$, $p_{bonf} = 0.007$, $d = -0.0776$) and between ST and TT ($t = -3.502$, $p_{bonf} = 0.005$, $d = -0.798$).

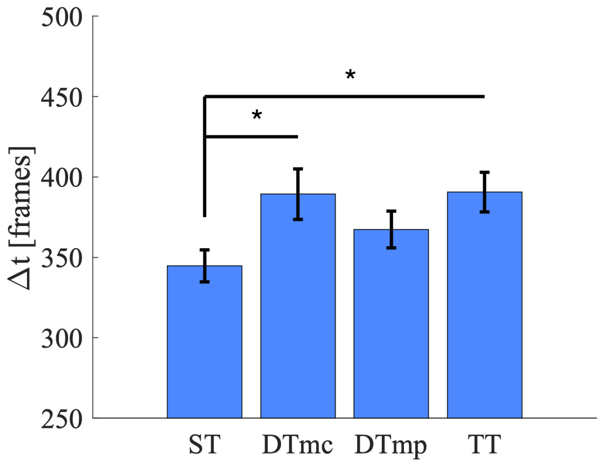

Figure 68: Mean Δt of FRC for different task loads during experiment 2.

Motor Error

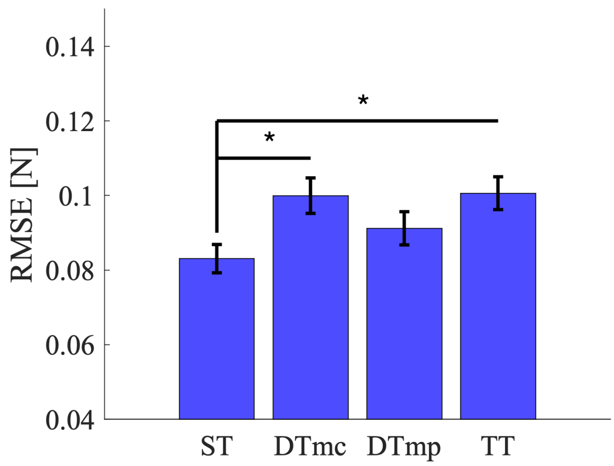

Figure 69: Mean motor error as normalized RMSE in FRC for different task loads in experiment 2.

For the motor error, operationalized as the RMSE normalized at each participant's 10% maximum force, normal distribution was violated for DTmp ($p = 0.024$) and TT ($p = 0.037$). Q-Q plots led to the validation of normal distribution for the residuals. Sphericity was also not given (p < 0.001). The repeated-measures ANOVA for the factor task load with the factor levels STm, DTmc, DTcp and TT was calculated resulting in a significant main effect for the task load ($F(1.715, 34.308) = 8.337$, $p = 0.002$, $\eta^2_p = 0.294$). For the Tukey HSD two comparisons were statistically significant (see Table 7).

Table 7: Statistical parameters for post-hoc pairwise comparison for normalized RMSE in FRC for statistically significant task load conditions in experiment 2.

		t	p_{bonf}	Cohen's d
ST	DTmp	-4.159	<0.001	-0.844
	TT	-4.324	<0.001	-0.877

8.4.2 Interim Summary for the Performance Decrement

For the reaction time, a statistically significant task load effect was found. Reaction time was slowest for TT execution (TT > DTcp > DTmp > ST). With this pattern of results, the classical performance decrement could be replicated for this dependent variable. Performance between ST and DTmp execution did not differ, leading to the assumption, that the combination of the FRC with the PRT task does not increase reaction times, while the combination with the CLC does. However, another pattern of result emerges when regarding the reaction error as the other dependent variable derived from PRT task execution. Here, errors were the highest for DTmp, which differed significantly from all other task load conditions. This is quite opposing to the results related in the cognitive task: While reaction time remains the same for DTmp execution in comparison to the PRT single task, reaction errors increase significantly under DTmp. This could be interpreted in a way that participants either react very quickly to the probe stimuli or not at all. Specifically in comparison to DTcp and TT, where we find a significant increase in reaction times in comparison to ST execution and only a moderate increase in the error rates.

For the calculation time in the CLC task, no task load effect was found. Calculation time was thus not influenced by multiple task execution in comparison to single task execution. No performance decrement could be replicated here. For the calculation error rate in %, a task load effect was found. The only difference however was found between ST and DTmc execution. The combination of CLC with the FRC task leads to an increased error rate here. Interestingly, this effect does not occur under TT execution.

The result pattern for the motor time lag and the motor error in the FRC task are similar. There is a task load effect, which is shown by the fact that there are differences between ST and DTmc and ST and TT execution. ST does not differ from DTmp for both dependent variables. This leads again to the assumption, that the combination of PRT and FRC tasks does not seem to lead to a performance interference and therefore decrement for the motor task.

8.4.3 Semantic Probetype Effect

To test for the probetype effect for the six dependent variables in the second experiment, the two dual task conditions, combining the probe reaction time task with either the motor or the cognitive task can be used (DTmp, DTcp), as well as the triple task. For every dependent variable, repeated-measures ANOVAs are calculated with the factors task load and probetype.

Reaction Time

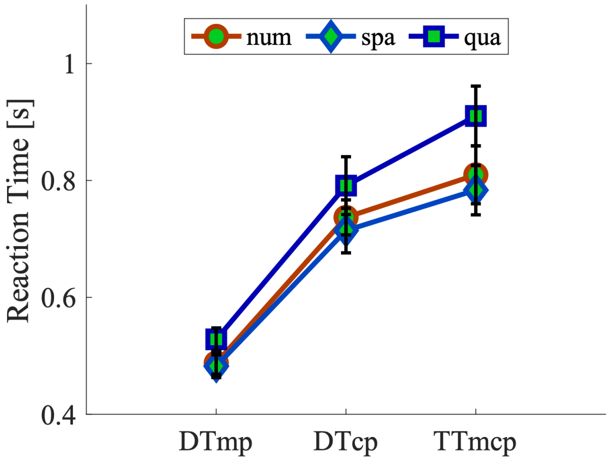

Figure 70: Mean reaction times for different task load and probetype conditions in experiment 2.

For the reaction time, the factor task load has three different factor levels (DTmp, DTcp, TT) as well as the factor probetype (num, spa, qua). 20 participants are included in the analysis because reaction time data of one participant was missing due to technical problems. Normal distribution is given, except for DTcp data for the qua ($p = 0.042$) and DTmp for the spa ($p = 0.022$) probetype. Q-Q-plots allow acceptance of normal distribution of residuals. There is a significant main effect for the factor task load ($F(2,38) = 87.112$, $p < 0.001$, $\eta^2_p = 0.821$) as well as for the factor probetype ($F(2,38) = 6.972$, $p = 0.003$, $\eta^2_p = 0.268$). The interaction task load x probetype was not significant ($F(2.515,47.791) = 0.276$ $p = 0.808$, $\eta^2_p = 0.014$). Post hoc comparisons for the task load showed that all factor levels differ statistically from each other (see Table 8). For the factor probetype, the num and spa probetypes did not differ statistically; the two other comparisons reached statistical significance (see also Table 8).

Table 8: Statistical parameters from post-hoc pairwise comparison for the factors task load and probetype in experiment 2.

Task Load					**Probetype**				
		Statistics					Statistics		
		t	p_{bonf}	d			t	p_{bonf}	d
DTmp	DTcp	-9.363	<.001	-1.430	num	spa	0.698	n.s.	0.093
	TT	-12.739	<.001	-1.946		qua	-2.828	0.022	-0.379
DTcp	TT	-3.376	0.005	-0.516	spa	qua	-3.526	0.003	-0.472

Reaction Error

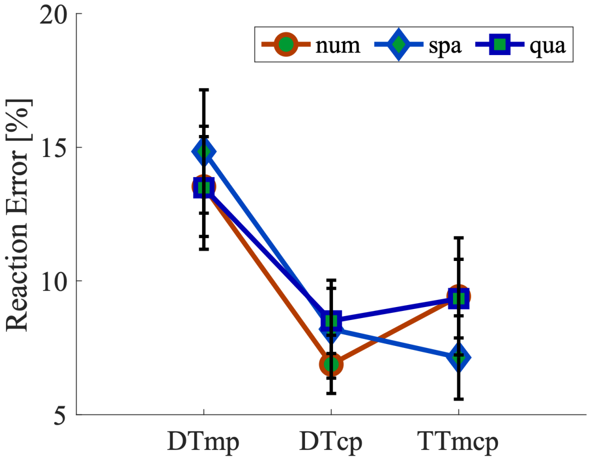

Figure 71: Mean reaction error for task load and probetype factors in PRT for experiment 2.

In the probetype analysis for the reaction error, data of 19 participants was included. Data of one participant was missing due to missing reaction time data. Another participant had to be excluded, because for some probetypes, the reaction error was beyond 50% which indicated that here, the participant did not properly execute the reaction time task, when missing more than half of the stimuli. For five of the nine combinations of factor levels (task load: DTcp, DTmp, TT; probetype: num, spa, qua), normal distribution was violated. Q-Q plots could reveal that normal distribution was violated only due to some outliers that still lay well within the valid error rates of below 50%. Thus, the repeated-measures ANOVA was nevertheless calculated. A significant main effect was only found for the factor task load ($F(2,36) = 16.646$, $p < 0.001$, $\eta^2_p = 0.480$), but not for the factor probetype ($F(2,36) = 0.106$, $p = 0.899$, $\eta^2_p = 0.006$), or the interaction ($F(4,72) = 0.340$, $p = 0.850$, $\eta^2_p = 0.019$). Because task load differences were already discussed with the performance decrement, the post-hoc comparison for this main effect is not again listed here.

Calculation Time

For the analysis of calculation time differentiated by different probetypes, the task load conditions DTcp and TT could be analyzed. Therefore, a 2 x 3 repeated-measures ANOVA for the factor task load (DTcp, TT) and probetype (num, spa, qua) was calculated. Data from two out of six combinations of factor levels was not normally distributed. Q-Q plots showed that normal distribution of residuals can be accepted. Neither one of the main effects task load ($F(1,20) = 1.794$, $p = 0.195$, $\eta^2_p = 0.082$) nor probetype ($F(2,40) = 0.105$, $p = 0.901$, $\eta^2_p = 0.005$), or the interaction ($F(2,40) = 0.062$, $p = 0.940$, $\eta^2_p = 0.003$) showed statistical significance.

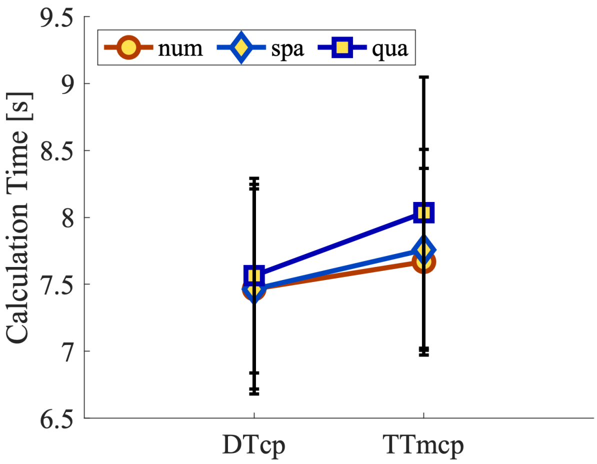

Figure 72: Mean calculation time for different conditions for the factors task load and probetype in experiment 2.

Calculation Error

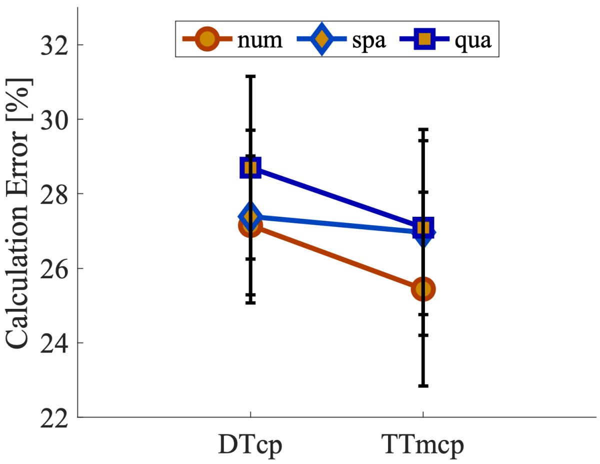

Figure 73: Mean calculation error in CLC for different conditions for the factors task load and probetype in experiment 2.

Analyzing the calculation error, the same task load conditions can be used, resulting in the same ANOVA. Normal distribution is given for all combinations of factor levels. No significant main effect for task load (*$F(1,19) = 0.255$, $p = 0.620$, η^2_p = 0.013*) or probetype (*$F(2,38) = 0.500$, $p = 0.611$, η^2_p* = 0.026) could be found. The interaction also reached no statistical significance (*$F(2,38) = 0.070$, $p = 0.932$,* η^2_p = 0.004).

Motor Time Lag

For the performance in the motor force tracking task for the differentiation of the different probetypes, the conditions DTmp and TT could be analyzed. Normal distribution was violated for three out of six combinations of factor levels, but Q-Q plots revealed that normal distribution of residuals was given. Therefore the 2 x 3 repeated-measures ANOVA for the factor task load (DTmp x TT) and probetype (num x spa x qua) is calculated and interpreted. No significant main effect was found for task load (*$F(1,20) = 4.350$, $p = 0.050$, $\eta^2_p = 0.179$*), for the probetype (*$F(2,40) = 0.069$, $p = 0.933$, $\eta^2_p = 0.003$*), or the interaction (*$F(1.450,28.996) = 0.431$, $p = 0.590$, $\eta^2_p = 0.021$*).

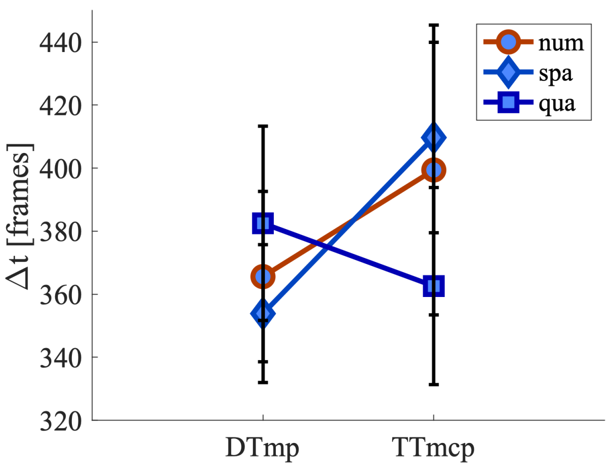

Figure 74: Mean Δt in FRC for different conditions for the factors task load and probetype in experiment 2.

Motor Error

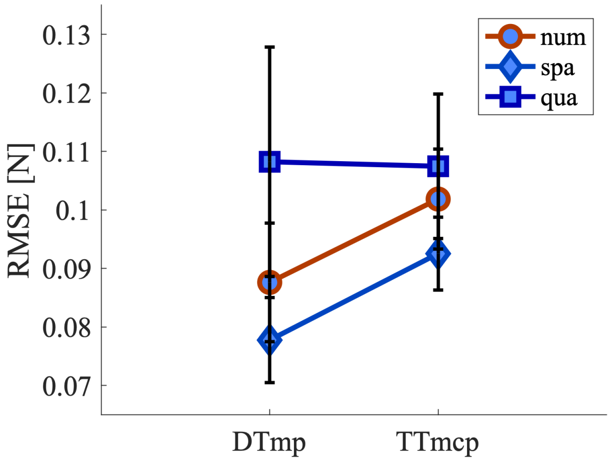

Figure 75: Mean normalized RMSE in FRC during different conditions for the factors task load and probetype in experiment 2.

For the motor error in the force tracking task, the same factors and factor levels as for the motor time lag can be analyzed in the repeated-measures ANOVA. All probetype combinations for the task load condition DTmp are not normally distributed. In the Q-Q plots it becomes clear, that normal distribution of the residuals here is only violated due to one outlier. This data lay below the accepted upper boundary of 0.5 N however and was therefore regarded a valid value. Thus, the ANOVA was interpreted. Only the factor task load reached statistical significance ($F(1,20) = 4.881$, $p = 0.039$, $\eta^2_p = 0.196$). The factor probetype ($F(1.369,27.388) = 2.520$, $p = 0.115$, $\eta^2_p = 0.112$) and the interaction ($F(1.182,23.636) = 0.162$, $p = 0.733$, $\eta^2_p = 0.008$) did not produce statistically significant effects.

8.4.4 Interim Conclusion for the Probetype Effect

For the reaction time of the PRT task, a probetype effect was found. Here, the quantitative probetype led to higher reaction times in all three task load combinations (DTmp, DTcp, TT). For all other dependent variables, no probetype effect could be found.

8.4.5 Event-related Effects

For the analysis of event related effects, all multiple task execution conditions can be compared with each other.

Reaction Time

Here, all conditions that held the probe reaction time task could be analyzed. For the DTmp this meant all ProbeForce events. In DTcp the reaction time could be differentiated between CalcProbe and CompProbe events and for TT CalcProbeForce and CompProbeForce events could be analyzed. Therefore, a repeated-measures analysis for the factor Event with 5 factor levels (ProbeForce x CalcProbe x CompProbe x CalcProbeForce x CompProbeForce) was conducted. Data of one participant had to be excluded from the analysis due to missing reaction times because of technical problems. There was a significant main effect for the factor event (*$F(2.104, 35.771) = 22.191$, $p < 0.001$, $\eta^2_p = 0.566$*). Post-hoc comparisons revealed numerous statistically significant differences, that are listed in Table 9.

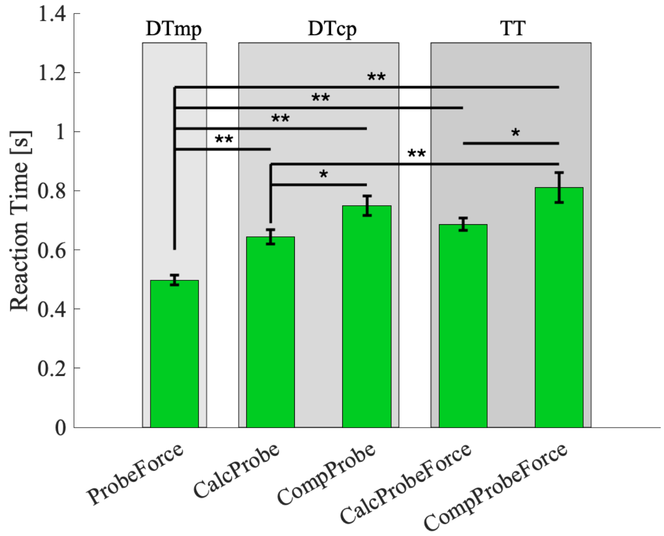

Figure 76: Mean reaction time in PRT for different events during different task loads in experiment 2.

Table 9: Statistical parameters for pairwise post-hoc comparisons of different events for the reaction time in PRT for experiment 2.

		t	p_{bonf}	Cohen's d
ProbeForce	CalcProbe	-4.179	<0.001	-1.103
	CompProbe	-7.080	<0.001	-1.868
	CalcProbeForce	-5.383	<0.001	-1.420
	CompProbeForce	-8.778	<0.001	-2.316
CalcProbe	CompProbe	-2.900	0.050	-0.765
	CompProbeForce	-4.598	<0.001	-1.213
CalcProbeForce	CompProbeForce	-3.395	0.012	-0.896

Reaction Error

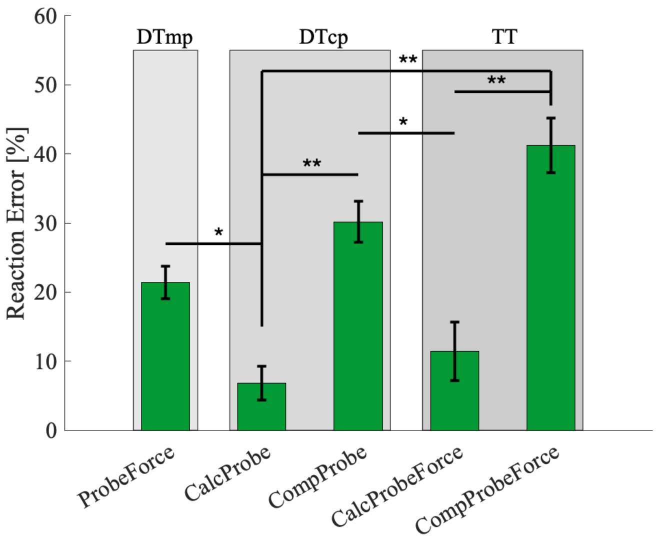

Figure 77: Mean reaction error of PRT for different events during different task loads in experiment 2.

For the ERA of the reaction error rates, the same task conditions and events can be analyzed. The analysis was conducted with the same 18 participants due to the missing data of one participant. Error rates were high for some participants because they did not respond to many probe stimuli. An exclusion of error rates higher than 50% would have led to a moderate data loss (data loss of participant's data per event: ProbeForce = 1; CalcProbe = 0; CompProbe = 1, CalcProbeForce = 6; CompProbeForce = 1), therefore all data was included in the analysis. The minimum of error rates were thus 0% and the maximum 82.759% (in CalcProbeForce). The S-W test showed a violation of normal distribution for all events. Q-Q plots also led to the assumption that residuals do not follow normal distribution. Therefore, the non-parametric equivalent, the Friedman test is used, to test for different reaction error rates in the events. A significant main effect for the factor event was found (*$\chi 2 = 54.470, p < 0.001, W = 0.757$*). The significant post-hoc comparisons found by the Conover test are listed in Table 10.

Table 10: Statistical parameters for post-hoc pairwise comparison using the Conover test of the reaction error in PRT for different events in experiment 2.

		t	p_{bonf}
ProbeForce	CalcProbe	4.105	0.001
CalcProbe	CompProbe	5.000	< 0.001
	CompProbeForce	6.473	< 0.001
CompProbe	CalcProbeForce	3.210	0.020
CalcProbeForce	CompProbeForce	4.684	<0.0011

Calculation Time

In the ERA for the calculation time, all Calc events could be analyzed and compared with each other. This was the CalcForce event from DTmc, the CalcOnly and CalcProbe events from DTcp and the CalcOnlyForce and CalcProbeForce events from TT execution. Data of 18 participants could be included in the repeated-measures analysis for the different events. Two outliers were detected that had consistently slow calculation times over all events (maximum for CalcProbeForce = 15.841 s). Values were however classified as valid and therefore left in the sample. The ANOVA showed a significant main effect for the factor event ($F(2.556, 43.455) = 22.970$, $p < 0.001$, $\eta^2_p = 0.575$). Significant pairwise post-hoc comparisons are listed in the following Table 11.

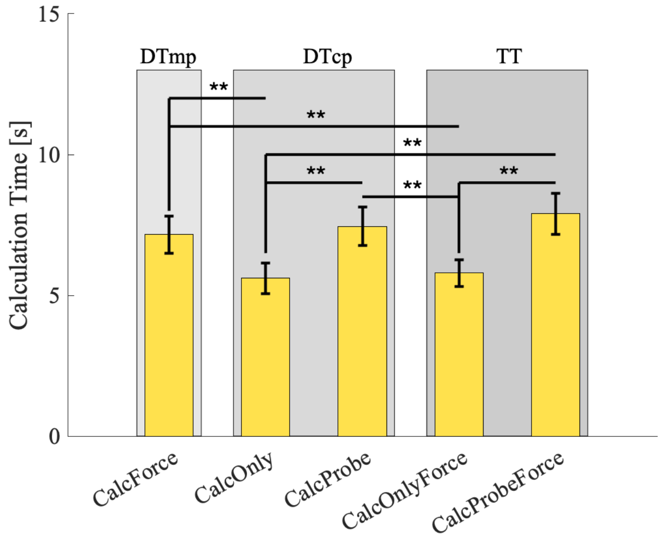

Figure 78: Mean calculation time during different events and task load conditions in experiment 2.

Table 11: Statistical parameters of the pairwise post-hoc comparisons for calculation times during different events in experiment 2.

		t	p_{bonf}	Cohen's d
CalcForce	CalcOnly	5.084	<0.001	0.553
	CalcOnlyForce	4.407	<0.001	0.479
CalcOnly	CalcProbe	-5.981	<0.001	-0.651
	CalcProbeForce	-7.722	<0.001	-0.840
CalcProbe	CalcOnlyForce	5.304	<0.001	0.577
CalcOnlyForce	CalcProbeForce	-7.046	<0.001	-0.766

Calculation Error

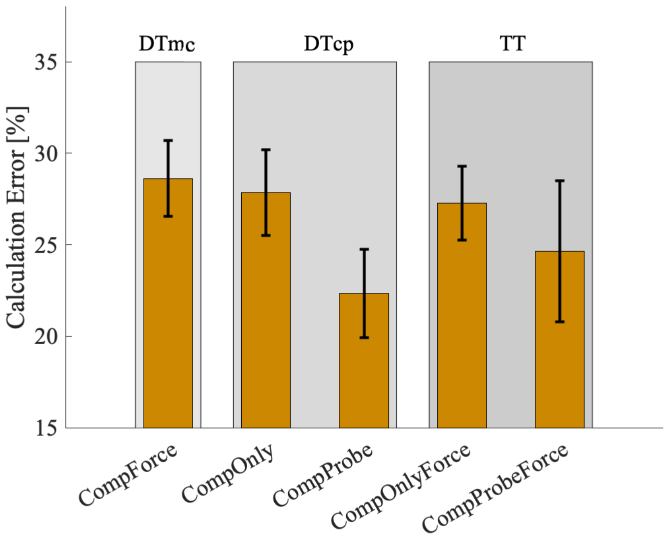

Figure 79: Mean calculation error for the comparison phase of CLC in different events under different task loads in experiment 2.

For the calculation error, all Comp events could be used in the ERA, resulting in the same event classifications as for the calculation time (with Comp instead of Calc). For the repeated-measures ANOVA, no significant effect for the factor event was found (F*(2.305, 41.483) = 1.131, p = 0.339*, η^2_p *= 0.059*).

Motor Error

For the ERA for the motor error (normalized RMSE subtracted by residuals after alignment), events from DTmc and TT execution were analyzed. Under the DTmc condition the CalcForce and CompForce events took place. For the TT condition CalcOnlyForce, CompOnlyForce, CalcProbeForce and CompProbeForce events could be differentiated. No significant main effect for the factor event was found in the repeated-measures ANOVA (*F(2.261, 38.438) = 2.118, p = 0.128*, η^2_p *= 0.111*).

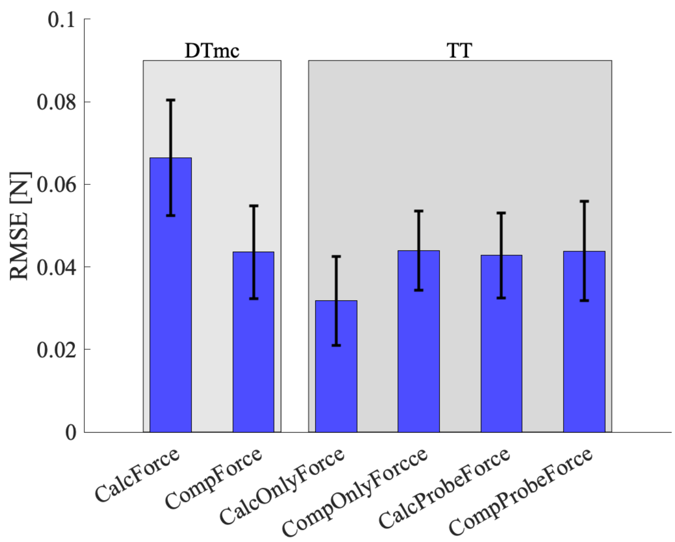

Figure 80: Mean normalized RMSE for FRC under different task loads and events in experiment 2.

Regions of Interest

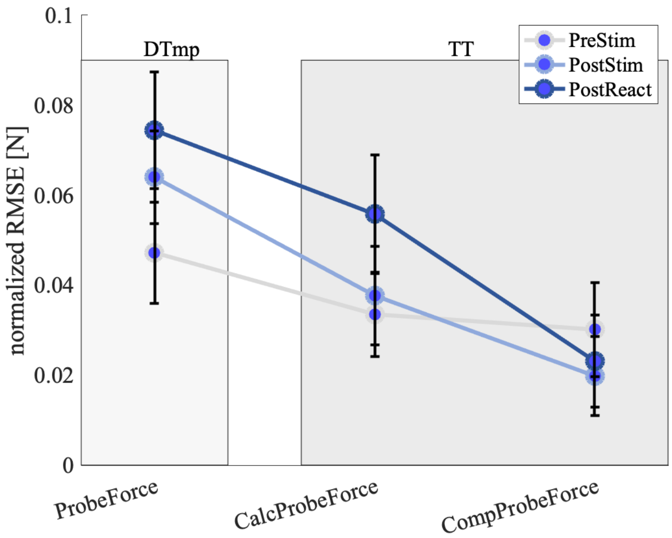

Figure 81: Mean normalized RMSE in FRC in different events and task loads for different ROI in experiment 2.

For the ROI in the ERA the motor error (RMSE) was again used as the dependent variable. Here, the ProbeForce event from DTmp execution was analyzed together with the CalcProbeForce and CompProbeForce events from TT execution. For each event the PreStim, PostStim and PostReact interval was determined. Thus a 2 x 3 repeated-measures ANOVA was calculated for the factor event and ROI. A significant main effect was found for the events ($F(2,32) = 3.356$, $p = 0.048$, $\eta^2_p = 0.173$) as well as for the ROI ($F(2,32) = 9.545$, $p < 0.001$, $\eta^2_p = 0.374$). The interaction between the two factors also showed statistical significance ($F(4,64) = 5.010$, $p = 0.001$, $\eta^2_p = 0.238$). Statistically significant differences in the post-hoc comparison for the both main effects are reported in Table 12.

Table 12: Statistical parameters for post-hoc pairwise comparisons of normalized RMSE in FRC for different events and regions of interest in experiment 2.

Events					Regions of Interest				
		Statistics					Statistics		
		t	p_{bonf}	d			t	p_{bonf}	d
ProbeForce	CompProbeForce	2.591	0.043	0.714	PreStim	PostReact	-4.311	<0.001	-0.323
					PostStim	PostReact	-2.773	0.028	0.028

8.4.6 Interim Conclusion for the Event-related Analysis

The ProbeForce event under DTmp execution has the shortest reaction times. For DTcp and TT similar result patterns arise: in each case, the reaction time is longer in the Comp calculation phase CompProbe event for DTcp and CompProbeForce event for TT) than in the Calc (CalcProbe for DTcp and CalcProbeForce for TT) calculation phase. For the reaction error, it can be seen even more clearly that in the Comp phase the stimuli often remain unresponded to, which is reflected in a significantly higher error rate compared to the Calc events.

For the calculation time it can be clearly observed that the presence of a probe stimulus leads to an increase for both DTcp and TT task load conditions. The CalcForce event for the DTmp condition has a calculation time that lies in between the different levels for the differentiation of whether a probe stimulus was present or not in the other two task load conditions. For the calculation error, no such observation could be made, because the events did not lead to a significant main effect.

The motor error did also not result in any significant differences for the different events under DTmc and TT execution. However, an interesting pattern of results was found for the differentiation of ROIs in the motor error. The event effect reached statistical significance revealing an increased RMSE in the ProbeForce (DTmp) event in comparison to the CompProbeForce (TT) event. Additionally, the PostReact ROI led to the highest RMSE in comparison to the PreStim and PostStim ROI.

8.5 Discussion Experiment 2

In the second experiment, the three tasks (PRT, CLC, FRC) were executed under single and triple task conditions. Additionally, all possible combinations of tasks were completed as dual task conditions. This allowed for an even more precise examination of task interference than in the first experiment. The biggest change here was that the curves in FRC became less predictable due to their inconsistency, and that CLC was now responded to verbally with a microphone to avoid a bimanual interference with PRT. Besides the examination of the classical performance decrement, the probetype effect and the ERA were also undertaken.

The following observations are made for multiple task execution in comparison to single task execution (performance decrement) in the time and PCU dimensions for the TR model in the six dependent variables (see also Table 13): For the combination of FRC and PRT (DTmp), there is no interference in the time dimension, but a large interference in the PCU dimension for PRT. There is no interference in the time or PCU dimension for FRC. Combining CLC and PRT (DTcp) resulted in an interference in both time and PCU dimensions for PRT, but no interference in both dimensions for CLC. In DTmc, where FRC and CLC are executed concurrently, there was an interference in the time, but not in the PCU dimension for CLC. For FRC the interference became visible in both dimensions. For the TT execution that combined all three tasks, there was an interference in the time, but not in the PCU dimension for PRT, no interference in CLC, and interference in both dimensions for FRC.

Table 13: Result pattern for experiment 2 comparing single task to multiple task execution for dependent variables resembling performance changes in Time and PCU dimensions. Red fields = Performance decrement under mutltiple task execution. Green fields = no performance decrement. Black fields = not measurable because single task was not included in multi-task condition.

		DTcp	DTmp	DTmc	TT
PRT	RT				
	RE				
CLC	CT				
	CE				
FRC	Δt				
	RMSE				

The PRT seems to be a good task to infer interference with one of the other tasks. For all task combinations, at least one of the dimensions for PRT (RT or RE) showed a performance decrement through task interference. While the combination of CLC and PRT did not affect performance in the cognitive task, the combination of CLC with FRC led to prolonged calculation times and under triple task execution also to higher calculation error rates. Performance in FRC is not affected by PRT, but by the combination with CLC in the motor time lag and the motor error. Strangely, this effect disappears for triple task execution.

This leads to the assumption, that the probe stimulus does not affect performance in FRC and CLC but that performance in PRT itself is affected under multiple task execution. This finding is again in line with Ells (1973) who found that reaction times to probe stimuli are prolonged while primary task performance remains untouched under dual task execution. The combination of CLC and FRC seems to lead to the most interference in the performance of both tasks. By combining all three tasks, the performance decrement (that was before mediated by CLC in DTmc) in both FRC dimensions vanishes. A very wild guess would be that this has to do with the probe stimuli somehow boosting motor performance in this case.

Regarding the probe stimuli, other than for the reaction time, where a quantitative probe stimulus led to the longest reaction times, no semantic probetype effect could be found for the cognitive or the motor task.This could either mean, that a breakthrough of the unattended semantics of the probetypes cannot be confirmed, either because it does not exist because participants did not process the semantic content, or it could also mean, it does exist, but the semantics of the words do not influence the performance of the tasks.

For the ERA, it is quite clear that for the time and PCU dimension of the PRT, the Comp phase of the calculation task leads to both higher reaction times and reaction error rates. Something similar can be seen for the calculation time (time dimension) for the CLC. Here, the performance is significantly decreased if a probe stimulus occurs in the event. However, there is no effect on the PCU dimension (calculation error). Also for the RMSE such an observation cannot be made. However, for the subdivision into different ROIs, the RMSE in the PostReact interval is higher for all three tested events. This could be in accordance with an effect monitoring hypothesis, which would state that by monitoring the manual response to the probe stimulus, cognitive resources are tied up for a certain time, which are then no longer available for the FRC task and thus briefly cause a drop in performance here.

To investigate whether the probe stimuli were actually semantically processed, the next experiment is conducted in a Go-NoGo design to force participants to process the semantics of the probe. If this leads to the finding of a probe effect, the semantic processing must have not taken place in the studies before. Also this experiment has revealed intereseting effects on performance at specific events. Therefore, the ERA should be kept up in the next experiment.

9 Experiment III - Probe Reaction Time Task as a Go-NoGo Design

In this experiment, a final effort was undertaken to demonstrate a semantic probe-type effect. The objective was to elicit higher-level semantic processing by employing the PRT within a Go-NoGo design framework. Here, only one word of a word pair required a manual reaction while the response to the other word had to be withheld. This way the involvement of higher processing levels becomes more likely, possibly leading to compatibility effects (Verghese et al., 2018).

This approach furhter allows for the investigation of how the processing demands of executed responses to Go stimuli differ from the inhibited responses to NoGo stimuli, particularly in terms of their interference with other tasks.

9.1 Methods Experiment 3

Given that experiment 2 clearly demonstrated that the DTmc combination results in decreased performance regarding calculation time, calculation error, and RMSE, this condition was not re-evaluated, nor was the TT condition. To further address research line R2, the DTmp and DTcp combinations, augmented with the aforementioned Go-NoGo component, were included in the PRT alongside the ST conditions.

9.1.1 Participants

In the experiment, 30 healthy undergraduate students participated. Due to technical problems during data acquisition, data of two participants had to be excluded from further analysis, leading to a sample size of 28 (19 female, 9 male). All participants can be classified as right-handed using the Edinburgh Handedness Inventory with a median score of 90 and an interquartile range of 20. The exclusion criteria were acute injuries of the lower exrtremities or neurological problems. Participants were required to understand German on a native speaker level, which was a necessity to ensure that the semantiv conten of the probe stimuli was fully understood. The local ethics committee had approved the experiment and all participants gave their written informed consent in accordance with the Declaration of Helsinki before the beginning of data acquisition.

9.1.2 Motor Task

The curve in the motor force tracking task was varied between participants, but stayed the same over all trials within each participant. The procedure of designing the curves around 10% of each paticipants maximum force with a maximum amplitude of ±5% remained the same. The templates were created by different Fourier sums of sine and cosine functions.

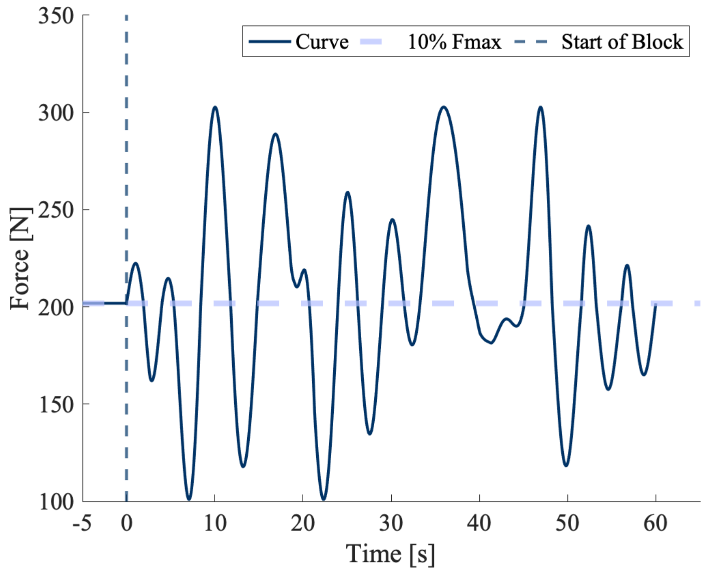

Figure 82: Exemplary generated force curve for participant 1 as used in the third experiment. Mean value of the curve was at 10% of the participants individual F_{max}. Curve characteristics started with the beginning of the task block. Five seconds before the start of the block, an adjustment phase was executed to reach the desired baseline.

At the beginning of each task block, a habituation phase of 5 seconds was executed, during which a countdown was presented on the experimental projection area and the bar for the force tracking task indicated the start value for the following 60 seconds block.

9.1.3 Cognitive Task

For the mental subtraction, the familiar task from the previous experiments was used. As in experiment 2, the result comparison phase always lasted 2.5 seconds regardless of whether participants indicated a registered error by giving the “pa” answer into the microphone, or not.

9.1.4 Probe Reaction Time Task

In the probe reaction time task three different semantic stimulus pairs were used to test for the probetype effect. The numeric (num) word pair had the same spoken numbers “four” and “six” with a stimulus duration of 552 ms for each word, as the previous two experiments and was expected to interfere with the cognitive task. The spatial (spa) word pair held the stimuli “hoeher” (532 ms, engl.: “higher”) and tiefer (524 ms, engl.: “lower”) and the quantitative (qua) word pair held the stimuli “staerker” (781 ms, “stronger”) and “schwaecher” (725 ms; “weaker”). Both spatial and quantitative semantics were expected to interfere with the motor task. The spatial stimuli were selected because they target the spatial requirements of the motor task. Since the visual display of the bar has the target value in the middle of the screen, the task is to use muscle strength to move the bar either “higher” or “lower”. This is realized by a “stronger” or “weaker” force appliance to the leg press.

Another new aspect to this task was its Go NoGo character. Each word pair had a defined Go stimulus that participants had to react to as quickly as possible and a NoGo stimulus, where the manual response had to be inhibited. The assignment of which stimulus within a word pair was the Go and which the NoGo stimulus was counterbalanced between participants.

9.2 Procedure Experiment 3

The experiment was executed on only one day. Participants came to the lab and signed the written informed consent form. Afterwards, the measurement of the maximum force was executed the same way as in the first two experiments. The actual experiment began by executing one block of the motor task as single-task condition and then the calculation task. After that, three blocks of the probe reaction time task were completed; each block with a different probetype. Then again, the motor and then the cognitive task were executed as single task conditions, one block each. When this habituation phase was finished, the dual tasking conditions began. A total of 12 blocks had to be executed here. Six blocks were the combination of the motor task with the probe reaction time task (DTmp), and six blocks held the combined cognitive and probe reaction time task (DTcp). Since there were three different probetypes (num, spa, qua), each dual tasking condition was carried out twice with each probetype. DTmp and DTcp were always alternating, while the probetype stayed the same for two task blocks in a row. Probe stimulus order for the blocks was pseudo-randomized over all subjects. Participants with an odd subject number began with the DTmp condition and participants with odd subject numbers began with DTcp. An exemplary protocol is shown in Table 14.

Table 14: Exemplary protocol for the procedure of experiment 3 combining single and dual task conditions.

<table>
<tr><th>1. Single Task Conditions</th><th colspan="6">2. Dual Task Conditions – First Passage</th></tr>
<tr><td rowspan="4">STm
STc
STp num, spa, qua (3 blocks)
STm
STc</td><td>DTmp num</td><td>DTcp num</td><td>DTmp spa</td><td>DTcp spa</td><td>DTmp qua</td><td>DTcp qua</td></tr>
<tr><th colspan="6">3. Dual Task Conditions – Second Passage</th></tr>
<tr><td>DTmp num</td><td>DTcp num</td><td>DTmp spa</td><td>DTcp spa</td><td>DTmp qua</td><td>DTcp qua</td></tr>
</table>

Between each block, a minimum break of 60 seconds was mandatory. If participants felt the need to pause a bit longer, they could do so any time and would indicate the experimentator when they felt ready to commence.

9.3 Data Pre-processing Experiment 3

The pre-processing of collected data from experiment 3 did not differ from experiment 2. However, it should be noted that the FSR malfunctioned intermittently and therefore reaction times and reaction errors are not available for all test subjects throughout all conditions. The amount of missing data for respective analyses is stated directly with corresponding dependent variables in the next subchapters of chapter 9.4.

9.4 Results Experiment 3

As in the two experiments before, first, the blockwise performance decrement is analyzed. For this, single task execution of each task is compared with the according dual task condition. Then dual task execution conditions are tested for the probetype effect. Last, the ERA is accomplished to investigate task interference on an ever smaller scale.

9.4.1 Blockwise Performance Decrement

To test for the blockwise performance decrement, single- and dual-task conditions are compared using t-test for paired samples and a rmANOVA for the Reaction Time.

Reaction Time

Data from five participants had to be excluded because it was not complete due to technical problems. For the reaction time a rmANOVA for the factor task load with the factor levels STp, DTmp and DTcp was conducted. The SW test showed normal distribution for DTmp ($p = 0.521$) and DTcp ($p = 0.053$), but not for STp ($p = 0.038$). The Q-Q plot for STp however shows that normal distribution for the residuals can be accepted . Sphericity was given with $p = 0.547$. The rmANOVA revealed a significant main effect for task load ($F(2,44) = 90.541$, $p < 0.001$, $\eta^2_p = 0.805$). A post-hoc comparison revealed a statistically significant difference between ST and DTcp ($t = -10.971$, $p_{bonf} < 0.001$, $d = -2.020$) and DTmp and DTcp ($t = -12.234$, $p_{bonf} < 0.001$, $d = -2.252$) but not between ST and DTmp.

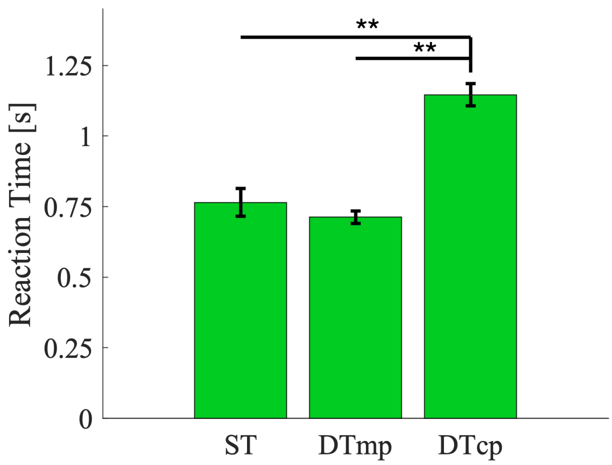

Figure 83: Mean reaction times of PRT under different task loads in experiment 3.

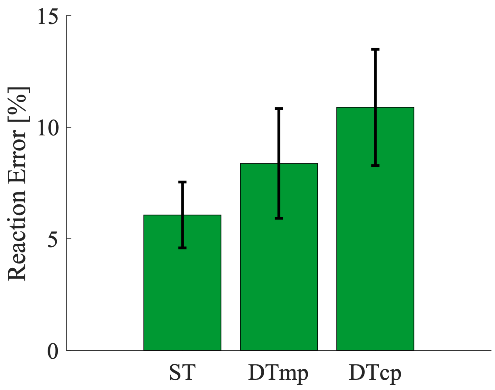

Figure 84: Mean reaction error of PRT for different task loads in experiment 3.

Reaction Error

The same analysis was conducted for the reaction error with the same sample size. Here, normal distribution was violated in the S-W test, as well as in the Q-Q plots. Therefore the non-parametric Friedman test is calculated. Here, no task load effect was found ($\chi^2 = 1.811, p = 0.404, W = 0.045$).

Calculation Time

For the calculation time a paired samples t-Test was calculated to test for differences between STc and DTcp. Bivariate normal distribution was given (p = 0.619). For the one-sided hypothesis, that calculation time would be smaller in ST compared to DTcp, the test showed no statistical significance ($t(27) = 1.920, p = 0.967$).

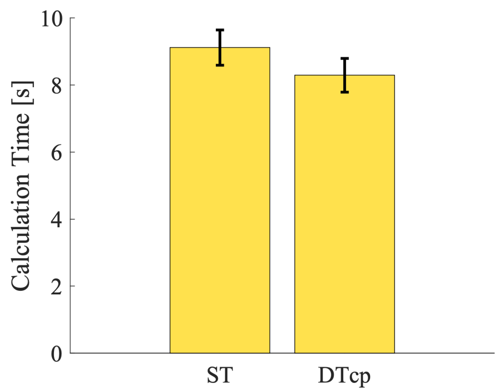

Figure 85: Mean calculation time in CLC for different task loads in experiment 3.

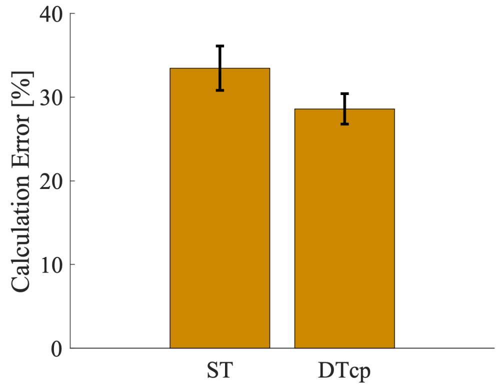

Figure 86: Mean calculation error of CLC during different task loads in experiment 3.

Calculation Error

The same t-Test was calculated for the error rate of CLC. A statistically significantly higher error rate under DTcp execution in comparison to ST execution, could also not be found for this dependent variable ($t(27) = 1.399, p = 0.912$).

Motor Time Lag

To test for performance decrement in the dependent variable Δt (corrected by residuals after alignment), ST and DTmp execution was tested against each other in a pairwise t-Test for the hypothesis, that time lag was slower under ST. However, this was not confirmed by the statistics: $t(27) = 3.554, p = 0.999, d = 0.672$.

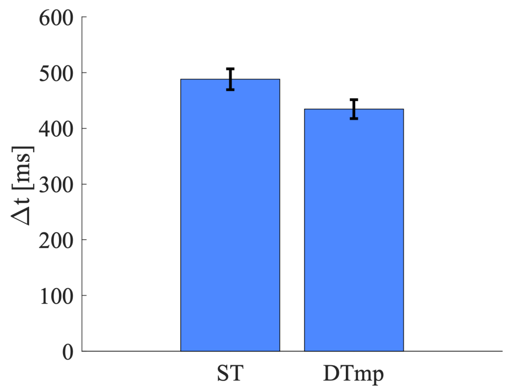

Figure 87: Mean Δt for FRC under different task loads in experiment 3.

Motor Error

For the performance decrement for the motor error, the same pairwise t-test was conducted for the normalized RMSE (corrected by residuals after alignment) in ST and DTmp. The one-sided hypothesis, that RMSE would be smaller under ST in comparison to DTmp execution could not be validated ($t(27) = 3.637, p = 0.999, d = 0.687$).

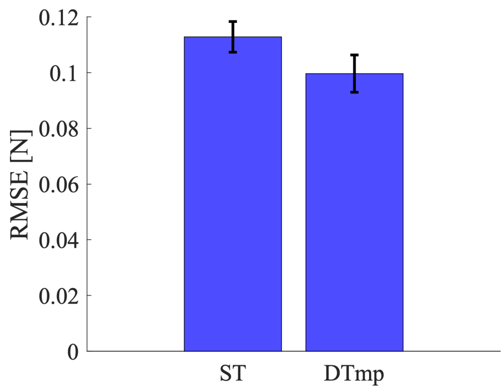

Figure 88: Mean RMSE for FRC under different task loads in experiment 3.

9.4.2 Interim Conclusion for the Performance Decrement

For the anaylysis of the blockwise performance decrement, the hypothesized effect was only found for the dependent variable reaction time in the DTcp condition, where reaction times were prolonged in comparison to ST and DTmp. No further performance decrement effect could be shown for any of the other dependent variables. However, all pairwise comparisons would have shown a statistically significant difference if tested one-sided for the hypothesis, that performance under ST execution would be worse, because performance in ST was decreased for calculation time, calculation error, Δt and RMSE in comparison to their belonging DT conditions. An explanation for this unexpected finding will be given in the Discussion (chapter 9.5).

9.4.3 Semantic Probetype Effect

To test for the semantic probetype effect in the third experiment, the DTmp condition can be used for the dependent variables of FRC and PRT and the DTcp condition for CLC and PRT. For the three different probetypes (num, spa, qua), rmANOVAs are executed to test whether performance in one of the tasks is specifically influenced by the semantic content of the stimuli.

Reaction Time

Reaction time data differentiated by probetypes was normally distributed (exception for DTmp num: $p = 0.033$ and DTcp spa: $p < 0.001$) and Q-Q plots also showed normal distribution of residuals. Data of six participants was missing. Therefore, the rmANOVA with the factor task load (DTmp x DTcp) and probetype (num x spa x qua) was only calculated with $n = 22$. Only the factor task load showed statistical significance ($F(1,19) = 114.598$, $p < 0.001$, $\eta^2_p = 0.858$). The factor probetype ($F(2,38) = 2.608$, $p = 0.087$, $\eta^2_p = 0.121$) and the interaction ($F(2,38) = 0.425$, $p = 0.657$, $\eta^2_p = 0.022$) were not statistically significant. Post-hoc comparisons for the task load are not reported here, because the results were already described under the performance decrement (see Chapter 9.4.1).

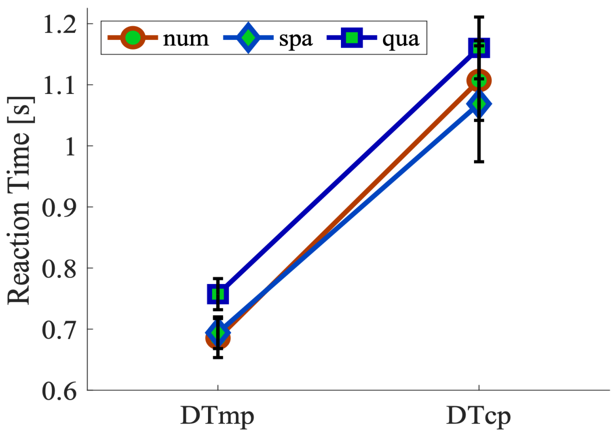

Figure 89: Mean reaction time in PRT of dual-task conditions with num, spa, or qua probe stimuli in experiment 3.

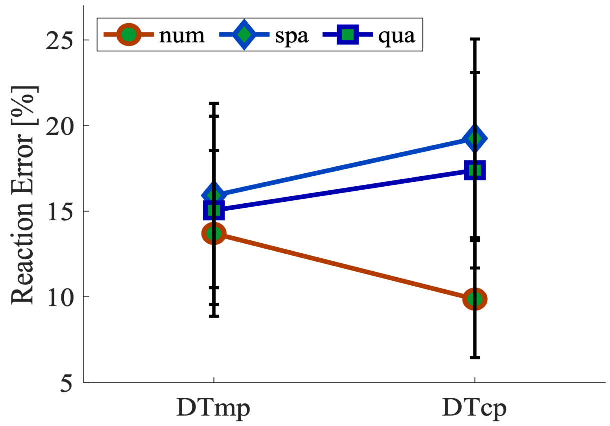

Figure 90: Mean reaction error in PRT of dual-task conditions DTmp and DTcp with num, spa, or qua probe stimuli in experiment 3.

Reaction Error

For the analysis of semantic probetype effects for the reaction error, the same data is used. Because normal distribution of the reaction error and normal distribution of the residuals are violated, the Friedman test is calculated. No effect for the probetype was found ($\chi^2 = 0.830$, $p = 0.659$), as well as for task load, or the interaction.

Calculation Time

To test for the probetype effect in the calculation time, data from the DTcp condition was used and differentiated for the three probetypes. Data of 27 participants could be used and normal distribution was given only for DTcp spa ($p = 0.173$) but normal distribution of residuals could be accepted. The rmANOVA for the factor probetype (num x spa x qua) did not lead to the finding of a significant main effect ($F(1.297, 33.731) = 1.220$, $p = 0.291$, $\eta^2_p = 0.045$).

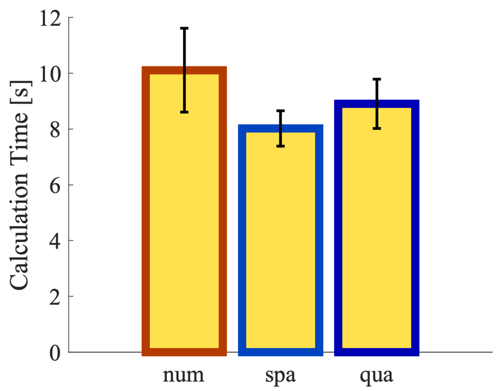

Figure 91: Mean calculation time in CLC for DTcp with probe stimuli num, spa, or qua in experiment 3.

Calculation Error

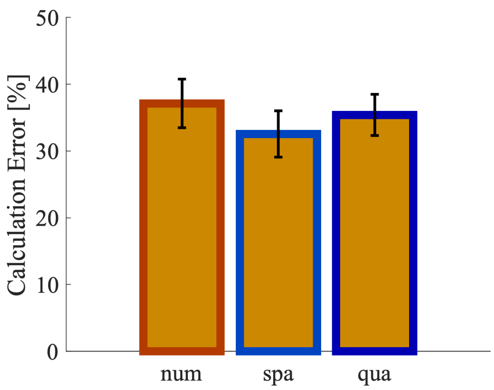

Figure 92: Mean calculation error in CLC for DTcp with probe stimuli num, spa, or qua in experiment 3.

For the probetype effect in the calculation error, the same data can be used. Data showed normal distribution and therefore the rmANOVA for the factor probetype could be executed. No probetype effect was found ($F(2,52) = 0.683$, $p = 0.510$, $\eta^2_p = 0.026$).

Motor Time Lag

To analyze the motor time lag (subtracted by aligned data residuals) for the probetype effect, the DTmp condition data was used. The rmANOVA for the factor probetype showed no statistically significant effect ($F(2,54) = 0.121$, $p = 0.886$, $\eta^2_p = 0.004$).

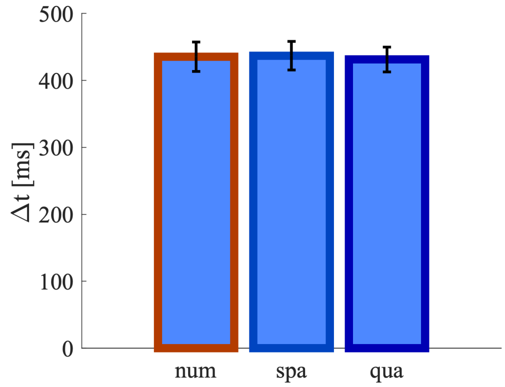

Figure 93: Mean Δt in FRC under DTmp for num, spa, or qua probe stimuli in experiment 3.

Motor Error

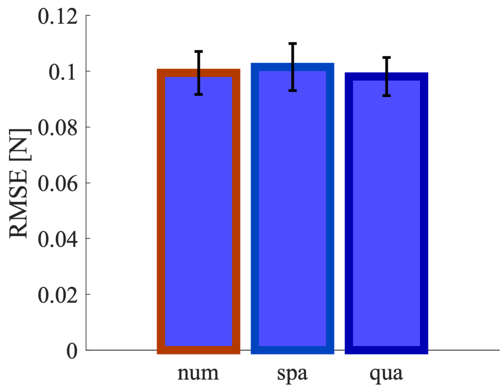

Figure 94: Mean normalized RMSE in FRC for DTmp with probe stimuli num, spa, or qua in experiment 3.

For the motor error (as normalized RMSE, subtracted by aligned data residuals), the same data was analyzed. Normal distribution of the data in the SW test was violated, but normal distribution of residuals in the Q-Q plots looked given. The rmANOVA was calculated for the factor probetype, but no statistically significant effect was revealed ($F(1.446, 39.033) = 1.150$, $p = 0.312$, $\eta^2_p = 0.041$).

9.4.4 Event-related Effects

For the ERA in the third experiment, DTmp and DTmc conditions were used.

ReactionTime

For the ERA of the reaction time, only Go probe stimuli could be analyzed. Therefore, the possible events during DTmp and DTcp task execution were GoProbeForce (DTmp) and CalcGoProbe and CompGoProbe (DTcp). Due to the existing technical problems and the decimation because only Go probe stimuli could be used, there was only few existing data. Additionally, CompGoProbe events were missing for some participants. Therfore, data of only eight participants could be analyzed. In the rmANOVA

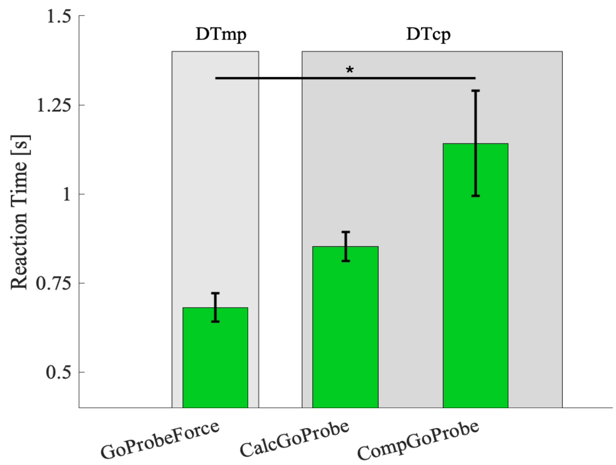

Figure 95: Mean RT in PRT for DTmp and DTcp conditions and three different events in experiment 3.

a significant effect for the factor probetype was found ($F(1.074, 7.521) = 7.831, p = 0.023, \eta^2_p = 0.528$). The pairwise post-hoc comparison revealed a statistically significant difference between the GoProbeForce and the CompGoProbe event ($t = -3.915$, $p_{bonf} = 0.005$, $d = -1.786$).

Reaction Error

The same conditions and events could be used for the ERA of the reaction errors. Additionally, the NoGo events (NoGoProbeForce, CalcNoGoProbe, CompNoGoProbe) could be used, too. Thus, more data was available here. In some cases, data for the CompGoProbe events were available for the reaction error that had been missing before because subjects did not react to these probe stimuli, resulting in no measurable reaction time, but a quantifiable reaction error. Thus, data of 27 participants could be used for this analysis.

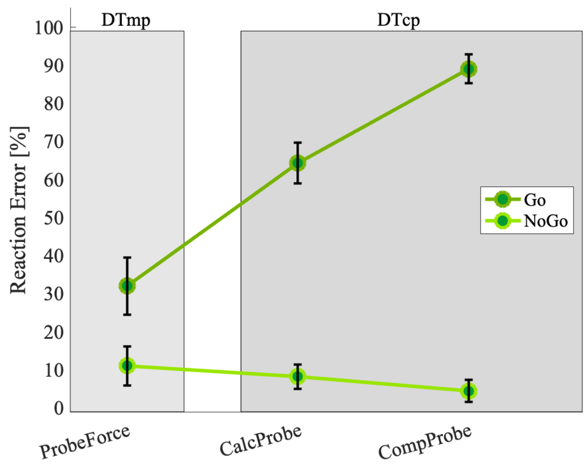

Figure 96: Mean reaction error in PRT for the conditions DTmp and DTcp as three different events in experiment 3.

Errors in the Go events meant, that there were no responses to the stimuli as intended. An error in the NoGo event meant that a response was executed, even though it should not have been. A 2 x 3 rmANOVA could be calculated for the factor event (ProbeForce x CalcProbe x CompProbe) and the factor probe response (Go x NoGo). A statistically significant main effect was found for the factor event ($F(1.312, 30.187) = 20.243, p < 0.001, \eta^2_p = 0.468$) and for the factor probe response ($F(1,23) = 136.860, p < 0.001, \eta^2_p = 0.856$), and also for the interaction ($F(1.858,42.725) = 56.662, p < 0.001, \eta^2_p = 0.711$). The pairwise post-hoc comparison revealed significant differences between all three events (see Table 15).

Table 15: Statistical parameters for post-hoc pairwise comparison for reaction errors in PRT under different events in experiment 3.

		t-value	p_{bonf}	Cohen's d
ProbeForce	CalcProbe	-3.782	0.001	-0.621
	CompProbe	-6.322	<0.001	-1.038
CalcProbe	CompProbe	-2.540	0.044	-0.417

Calculation Time

For the ERA of the calculation time, three events from the DTcp condition could be used: CalcOnly, CalcGoProbe and CalcNoGoProbe. Data of 27 participants could be included in the analysis. Normal distribution for the calculation time was only given for CalcNoGoProbe data ($p = 0.225$), but normal distribution of residuals in the Q-Q plots looked acceptable. A rmANOVA was calculated for the factor event, where a statistically significant effect was found ($F(1.534, 39.885) = 18.074, p < 0.001, \eta^2_p = 0.410$). The post-hoc comparison showed significant differences between CalcOnly and CalcGoProbe ($t = -5.564, p_{bonf} < 0.001, d = -1.024$) and CalcOnly and CalcNoGoProbe ($t = -4.755, p_{bonf} < 0.001, d = -0.875$); but not for the comparison of CalcGoProbe and CalcNoGoProbe ($t = 0.808, p_{bonf} = 1.000, d = 0.149$).

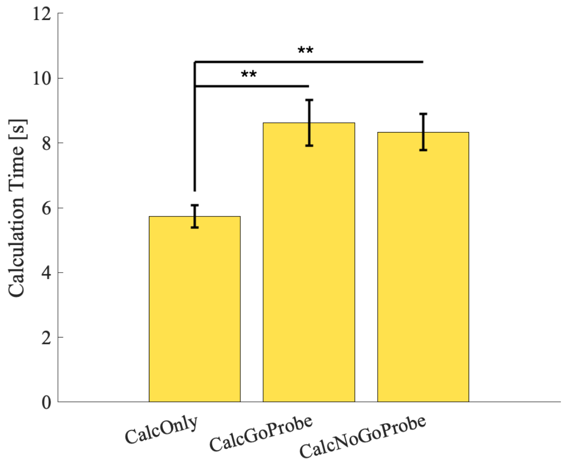

Figure 97: Mean calculation time in CLC for event-related-analysis under DTcp for different events in experiment 3.

Calculation Error

To analyze the calculation error in the ERA, the same events as for the calculation time could be used, but for the comparison phase of the calculation task (Comp); resulting in the CompOnly, CompGoProbe and CompNoGoProbe events. Data of 26 participants could be included in the analysis. Because normal distribution was only given for the CompOnly event (p = 0.339), and normal distribution of the residuals looked questionable, the Friedman test for the factor event was calculated. No statistically significant effect was found ($\chi^2 = 4.796, p = 0.091, W = 0.096$).

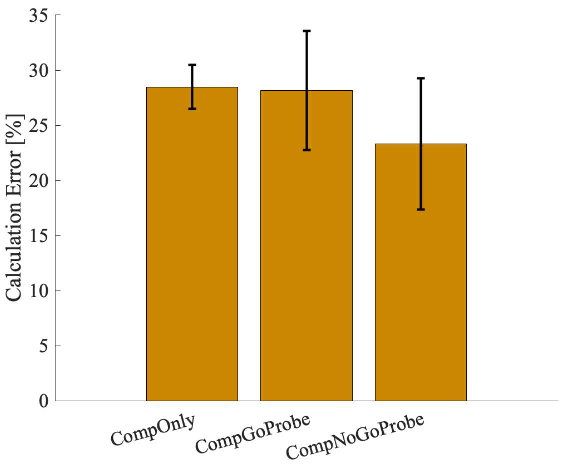

Figure 98: Mean calculation error in CLC for event-related analysis over different events during DTcp for experiment 3.

Motor Error and Regions of Interest

Since the only dual task combination was DTmp, the ERA for the dependent variables of the motor task is only appropriate in relation to the probe stimuli. Thereby the RMSE is directly calculated for the different ROIs for the ProbeForce event for either the Go stimulus (GoPreStim, GoPostStim, GoPostReact), or the NoGo stimulus (NoGoPreStim, NoGoPostStim). For the NoGo stimulus, no PostReact is calculated because there ought to be no reaction to the NoGo stimulus. Therefore, the factor ROI cannot be used together with a Go/NoGo factor for a rmANOVA, because there would be a missing NoGoPostReact condition. Thus, the rmANOVA is calculated for the factor event with the five given combinations (GoPreStim, GoPostStim, GoPostReact, NoGoPreStim, NoGoPostStim). Data of 23 participants were included in the analysis. Normal distribution was violated, but Q-Q plots revealed normal distribution of residuals with the exception of some outliers, that were nevertheless valid values. A significant main effect was revealed for the factor event (*$F(1.394, 30.664) = 5.323$, $p = 0.019$, $\eta^2_p = 0.195$*). The statistically significant pairwise post-hoc comparisons are listed in Table 16 below:

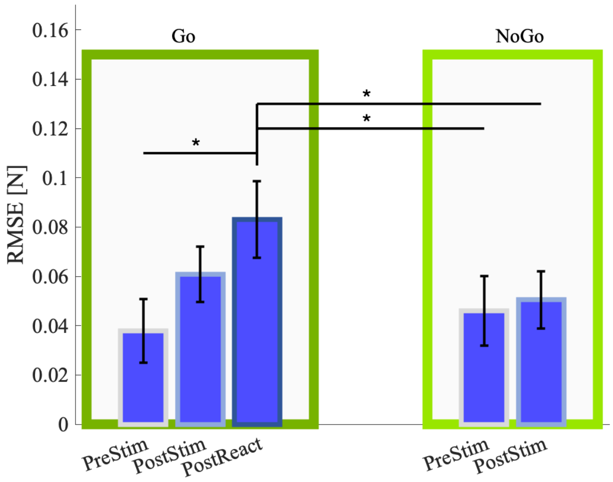

Figure 99: Mean normalized RMSE for FRC in DTmp for different events under Go and NoGo probe stimuli in PRT in experiment 3.

Table 16: Statistically significant differences in experiment 3 for RMSE under DTmc for events and their ROI's: PreStim, PostStim and PostReact for Go stimuli such as PreStim and PostStim for NoGo probe stimuli.

		t-value	**p_{bonf}**	**Cohen's d**
GoPostReact	GoPreStim	3.890	0.002	0.900
	NoGoPreStim	3.648	0.004	0.844
	NoGoPostStim	3.581	0.006	0.829

9.4.5 Interim Conclusion for the Event-related Analysis

For the reaction time, it was already known from the performance decrement analysis that there would be a difference between DTmp and DTcp. Now, interestingly, it becomes clear that the performance only deteriorates during the Comp and not during the Calc phase. This trend is also clearly visible for the reaction error. The already known higher reaction error rate under DTcp differs again within DTcp between the Calc and Comp condition and is clearly higher under Comp. Furthermore, it can be clearly seen that the error rate only increases under the Go stimuli. In this case, an error results if the required manual response is not executed on these stimuli. Thus, Go stimuli are missed resulting in higher error rates, whereas the NoGo stimuli do not result in higher error rates in the form of an incorrect response.

For the calculation time, it was observed that within the DTcp condition, the time required to solve a task increases whenever a probe stimulus occurs, in comparison to trials without presented stimuli (CalcOnly). It is irrelevant whether this stimulus has to be reacted to (Go) or not (NoGo). For the calculation error, no such influence of the probe stimulus is found.

The ERA for the motor error was executed as the analysis for the ROI under the DTmp condition. Here, the PostReact ROI for Go probes differed statistcally from the PreStim ROIs for both Go and NoGo stimuli and from the PostStim ROI of the NoGo stimuli. Thus, PreStim and PostStim seem to not differ from another, suggesting that the perception and response selection for the probe stimulus do not lead to a discrete performance decrement for the motor task. But a decrement can be observed for the PostReact ROI of the Go stimuli pointing towards a declined performance in the motor task, after the manual response to the stimulus from the PRT task was executed.

9.5 Discussion Experiment 3

For the third experiment, only dual task conditions that combined the PRT with either FRC, or CLC were executed to further investigate the influence of the semantic probe stimuli on the motor and the cognitive task (R2). To also reinforce the semantic processing of the probe stimuli, the PRT task was executed as a Go-NoGo task, where participants had to either manually react to the probe stimulus (Go), or withhold their reaction (NoGo).

The classical performance decrement (R1) could be shown in this experiment only in the reaction time for the comparison of the single task condition with DTcp. The performance in the dependent variables of the calculation and the motor task under single task execution was this time below the performance in the dual task conditions. However, this can be explained by the fact that the single task conditions for the motor and cognitive task were only conducted with two blocks each at the beginning of the experiment to make the functioning for the tasks clear to the subjects. Therefore, the performance in the single task blocks contains only these familiarization trials, which were not averaged over a large number of trial blocks, resulting in the poorer performance compared to the dual tasking conditions run later. This is not clearly visible in the reaction time task, since it can be assumed that no major habituation effects occur here, since the task has a rather simple character. An attempt was made to extract this habituation trend from the data with the help of a regression. However, the consideration of the sequential effects only contributed to a variance elucidation for the dependent variables of the motor task, so that for all other variables it was dispensed with. At this point, however, it should be mentioned that this does not detract from the quality of the experiment, since the performance decrement effect was already clear from the first two experiments. The focus of this study was on the combination of the semantic probe stimuli with the motor and the cognitive task.

This hypothesized probetype effect could however not be found in the data for any of the dependent variables because no specific performance changes could be observed for the combination of the reaction time task with numeric stimuli on the cognitive task and with the reaction time task with spatial or quantitative stimuli on the motor force tracking task. Rüschemeyer et al. (2007) and Pulvermüller et al. (2005) have shown that processing motor verbs describing a specific motor performance require resources in M1 and PMC in a somatotopically-organized manner. Verbs describing leg actions would thus specifically activate leg areas. This could mean, that the probe stimuli that are hypothesized to lead to a performance decrement in the motor task are still not appropriate because they lack a somatotopical link to the actual task.

The event related analysis was suitable to uncover that specifically Go probe stimuli during DTcp execution led to a discrete performance decrement in the comparison phase of the calculation task for the reaction time and error. The performance in the calculation time suffered specifically under probe presentation, independent from the neccessity of the sitmulus requiring a motor response or not. The most interesting finding occurred in the DTmp task combination, where it could be shown that performance in the RMSE did specifically decrease in the interval after a motor response was carried out to a Go probe stimulus. This can be seen as evidence for an effect monitoring of the manual response in the reaction time task that claims cognitive resources that create a bottleneck for the motor task execution for a very short period of time.

10 Summary of the Results from Experiment 1-3

Three experiments were conducted, each integrating a PRT, a CLC, and an FRC task, with variations in their execution across dual-task and triple-task conditions. In certain instances, the patterns of results varied significantly across the different experiments within the research interests (R1-3). However, these differences are quite understandable due to the slight variations in the paradigm, which led to different events. The absolute measured values for the respective dependent variables from the three tasks of the experiments should not be directly compared with each other. Nevertheless, generalizing statements can be made based on the comparison of the inferential statistical results.

Beyond examining the performance decrement (R1), the experiments also aimed to explore the extent to which the semantics of probe stimuli cause specific interference with either the motor or cognitive tasks (R2). Given the critical importance of the temporal organization of tasks and their subcomponents in self-organized multitasking, an event-related analysis was developed to further evaluate the previously proposed time-regime model (R3). Therefore, different combinations of PRT, CLC and FRC were used as dual and triple-task loads. The used task combinations in the three experiments can be derived from Table 17. Distinctive characteristics here were that the CLC task was answered manually in experiment 1 and verbally in experiments 2 and 3. In experiment 3, the PRT was carried out as a Go-NoGo paradigm.

Table 17: Conditions over the three experiments combining the probe reaction time task (PRT, p) the cognitive calculation task (CLC, c) and the motor force tracking task (FRC, m) in different task load conditions.

Experiment 1	STp	STc	STm	DTcp	DTmp	-	TT
Experiment 2	STp	STc	STm	-	-	DTmc	TT
Experiment 3	STp	STc	STm	DTcp	DTmp	-	-

To test task interference in the three experiments under the assumption of the time-regime model, a performance parameter in at least one of the dimensions of the time-regime model was required for each task. For PRT, these were response time and response error rate. For CLC these were calculation time for the time dimension and calculation error rate for the PCU dimension. In addition to the RMSE in the PCU dimension, Δt was also introduced as a performance parameter in the time dimension for FRC. It could be shown that Δt is a suitable performance measure for continuous tracking tasks, that is sensitive to multitasking conditions by resulting in a performance decrement. It also shows a different result pattern than the RMSE and thus forms a useful addition (as shown in experiment 1).

10.1 Performance Decrement (R1)

In the first experiment, a notable performance decrement was evident, as performance parameters in time and PCU dimension were statistically significantly superior under single-task conditions compared to triple-task conditions. The only exception was the response error in PRT, where no statistically significant difference was observed between the two task demands. Additionally, no data were obtainable for the calculation error in the CLC in this experiment.

Experiment 2 similarly demonstrated a substantial decline in performance, with outcomes in single-task conditions surpassing those in DT and TT conditions across nearly all parameters. The only exception was the calculation time in CLC, where no difference was measured.

In experiment 3, statistically different reaction times in PRT were evident only in the comparison where (ST=DTmp) < DTcp. All other comparisons failed to show a performance decrement under heightened task loads. However, since only DTmp and DTcp conditions were examined, it is possible that the performance decrement was minimal, potentially being averaged out in the block-by-block analysis based on averaged performance metrics. This aspect will be further explored in chapter 11.

10.2 Semantic Probetype Effect (R2)

In experiment 1, the proposed effects of semantic probe stimuli were observed for several dependent variables. Regarding reaction times in PRT, the hierarchy was beep < num < (int = dir = word). For the calculation time in the CLC task, it was (int = dir) < beep, with num and word not displaying statistically significant differences from the other probe types. Notably, it was found that for Δt, the sequence was (num = int = dir = word) < beep, indicating that the beep stimulus caused a greater temporal delay in the FRC task compared to the semantic stimuli. However, the reverse was noted for the RMSE, where the beep resulted in smaller deviations than the semantic stimuli.

In experiment 2, only one probetype effect could be shown for the reaction time in PRT, in which (num=spa)<qua. All other dependent variables of the tasks in the time and PCU dimension showed no specific statistically significant effects for the different semantics of the probe stimuli. Further, no probetype effect could be found in experiment 3.

10.3 Event-related Analysis (R3)

This type of analysis yielded the most insightful findings concerning the temporal interdependence of the tasks and their associated performance decrement.

In experiment 1, it was demonstrated that for the calculation time under TT condition, CalcOnlyForce events had lower durations than the CalcProbeForce events. Thus, the computing time per task for CalcOnlyForce decreased to the level observed in single-task (ST) conditions when no probe stimulus required a response during the TT events, even though a stimulus onset could have been expected.

In experiment 2, it was demonstrated for reaction time in the PRT task, that ProbeForce<(CalcProbe=CalcProbeForce)<(CompProbe=CompProbeForce). This indicates that response time is particularly prolonged due to the additional processing in the CLC task, necessitating a distinct differentiation between the calculation phase and the response comparison phase. The latter contributes to a further extension of calculation time. A similar pattern was observed for response errors in PRT, with more errors occurring during the comparison phase as opposed to the calculation phase. Regarding calculation time in CLC, it was also noted that calculation time is specifically extended during events containing a probe stimulus. For the region of interest (ROI) concerning the RMSE, it was shown that RMSE increases in the PostReact ROI compared to the PreStim and PostStim periods.

Experiment 3 also utilized event-related analysis to identify specific interferences between tasks in DTmp and DTcp conditions. It was found that the reaction error in PRT for NoGo stimuli was smaller than for Go stimuli, indicating that inhibiting a response to the probe stimulus leads to less interference than executing the response. For calculation time in CLC it was shown that CalcOnly<(CalcGoProbe=CalcNoGoProbe), implying that the mere presence of a probe stimulus increases calculation time, regardless of whether a response is required. In terms of the root mean square error (RMSE) for the region of interest (ROI), similar to findings in experiment 2, the PostReact RMSE was higher than during PreStim or PostStim periods, irrespective of whether these periods involved a Go or NoGo stimulus.

11 Discussion

The three experiments conducted for this work were used to replicate performance decrement (R1), to test whether semantic probe stimuli can specifically influence task performance (R2), and if an event-related approach of analyzing task performance (R3) leads to more specific insights into the allocation of processing resources.

For many of the dependent variables during experiments 1 and 2, the classic performance decrement (R1) can be replicated, but surprisingly, for some variables, this approach does not yield a significant performance decrement in multi-task compared to single-task execution. In experiment 3, no performance decrement was observed for any dependent variable. However, the event-related analysis revealed significant differences between events within the multi-tasking conditions. For some events, performance was comparable to single-task execution, while for others, task interferences became visible. This suggests that these differences are highly nuanced and that any resulting interference can be effectively compensated for, possibly through the execution of discrete tasks over an extended period of time. This finding can be interpreted as further evidence of flexible capacity distribution, underscoring the importance of conducting event-related analyses. This clearly indicates that the block-by-block method of assessing performance is not suitable for representing the flexibility with which the allocation of processing resources actually takes place. Evidence could be gathered through the ERA that even during a triple-task block, performance in individual tasks can temporarily reach the level of single task execution at discrete times, provided that the other tasks do not demand high processing capacity at those moments.

The hypothesized probetype effect (R2) could not be found. Only in experiment 1 did the beep stimulus differ in part from the other semantic stimuli in its influence on the dependent variables, resulting in higher calculation times and motor time lag but lower motor error in FRC. It seemed plausible that the semantic num probe stimuli exerted a greater influence on arithmetics, particularly in the second and third experiments, as the response to the arithmetic task involved verbalization with the sound "pa" rather than a motor action like a click. The absence of interference from number words in arithmetic tasks may be attributed to the processing of these words in a different brain region (related to speech) than the digits on the screen (related to arithmetic), in accordance with Campbell (2009) and Cipolotti and Butterworth (1995), who proposed the multi-route model of numerical processing. However, specific interference could have occurred due to the verbal nature of the response. Conversely, it is possible that the "pa" sound, lacking semantic content, did not sufficiently provoke interference as it may not adequately stimulate the speech centers. In future experiments, it may be advisable to require a fully verbal response by naming the result of CLC to reassess whether a number word as a probe stimulus could introduce interference at this stage.
The line of research that yielded the greatest insights was the event-related analysis (R3). Here, very fine-grained task interference was detected at discrete points in time during multi-task execution, suggesting a

dynamic allocation of processing resources during all three experiments. To categorize the ERA results from the three experiments within the time regimes model, it is essential to account for interference at the task level during the individual processing phases of the PRT, CLC, and FRC tasks. This involves drawing appropriate conclusions regarding the utilization of resources in both the temporal and processing capacity utilization dimensions.

The underlying premise is that each of the three tasks (with the CLC further subdivided into Calc and Comp phases and FRC into its processing stages) imposes specific demands on the processing capacity, which can be represented by the resulting area of time and processing capacity utilization for the task (see Figure 100).

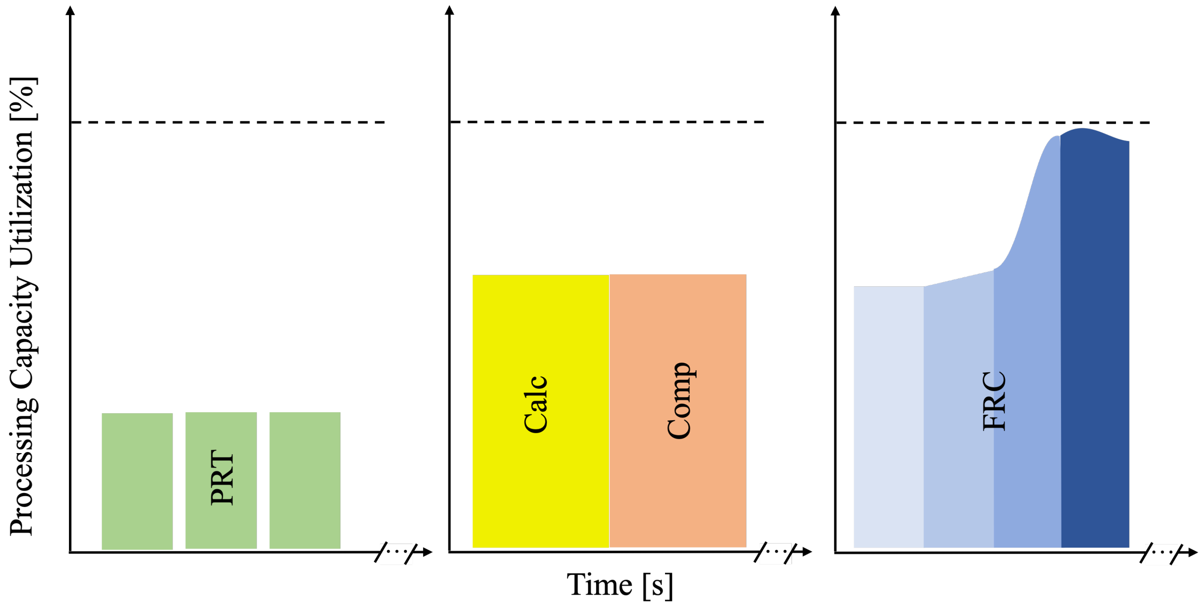

Figure 100: Basic Assumption of required PC for the three tasks PRT, CLC (further divided into Calc and Comp Phase) and FRC (divided into perception, response selection, response execution and effect monitoring phase) in the three experiments.

The experiments investigated the extent to which task interference occurs during multitasking due to the varying demands and time regimes of the tasks. This will now be examined in more detail within the context of the Time Regime (TR) model.

11.1 Probe Reaction Time Task in the Time Regime Model

For the reaction time in the PRT task, it could be shown in experiment 2 that ProbeForce < (CalcProbe = CalcProbeForce) < (CompProbe = CompProbeForce) and in experiment 3 (GoProbeForce = CalcGoProbe) < CompGoProbe. This implies that processing the Comp phase of the CLC results in the longest reaction times in PRT. Since probe tasks can be used to assess capacity utilization in the primary tasks (Ells, 1969; McLeod, 1978), it can be inferred that the Comp phase demands the highest capacity utilization. Consequently, this corresponds to the largest area required in the time regime model.

The result pattern for the reaction error of the PRT task in experiment 2 ((CalcProbe = CalcProbeForce) < (ProbeForce = CompProbe = CompProbeForce)) and experiment 3 (ProbeForce < CalcProbe < CompProbe) showed a slightly different pattern of results. In experiment 2, the FRC involved a complex curve where the beginning and end varied for each task block, while only the middle section was consistent within a subject. In contrast, experiment 3 featured a similarly complex curve that remained identical for each subject across all task blocks. This difference in curve complexity could account for the varying patterns of results between the two experiments. In experiment 2, the Calc phase of the CLC exhibited a (partially statistically significant) lower reaction error compared to events requiring concurrent force tracking. In experiment 3, the reaction error in PRT was lowest for the expected result pattern in FRC and highest for the Comp phase in CLC, with the Calc phase of CLC falling between these two. Therefore, within the TR model, it can be inferred that the more complex curve in experiment 2 resulted in a higher capacity requirement (Time * PCU) for the FRC compared to the Calc phase. Conversely, in experiment 3, with the simpler curve, the requirement was lower than that for the Calc phase of CLC.

Interestingly, the PRT reveals two key insights: the Comp phase of the CLC generally appears to demand a larger area in the TR model compared to the Calc phase of the CLC, leading to higher reaction times. However, the results of reaction time and calculation time must be interpreted in context, as the Calc phase of the CLC, from which calculation time is derived, has a FlexTR, potentially extending calculation time. Examining the reaction error in PRT provides further clarity, as it reflects missed reactions or incorrect responses (such as in a NoGo probe in experiment 3) to the probe stimulus. This highlights when a response to a probe stimulus was not measured due to excessive duration (longer than 2 seconds) or incorrect execution (only in experiment 3). A comparison between experiments 2 and 3 clearly demonstrates that the FRC can also impose similarly high demands in the TR model, depending on the complexity of the curve to be tracked. This underscores that even minor adjustments to a task can significantly impact the level of task interference when evaluating performance in individual events.

This is consistent with the findings of a study by Laessoe et al. (2008), who observed increased gait variability in older adults walking in a figure-of-eight pattern while performing simultaneous a motor and a cognitive task as dual-tasks or triple-tasks. Although walking is typically a highly automated activity, following a prescribed pattern can lead to gait disturbances (Shkuratova et al., 2004). This underscores the importane of carefully selecting task specificities and characteristics for both motor (Beauchet et al., 2005) and cognitive tasks (Laessoe et al., 2008), as they affect capacity demands. Applying this to the time regimes model suggests, that the dimensions of task areas vary based on time and PCU requirements. More specifically, this implies that even tasks with minor differences in characteristics can have significantly varying capacity demands.

11.2 Calculation Task in the Time Regime Model

Calculation time and calculation error are the deviated dependent variables for the CLC task. Calculation time can be derived from the Calc phase, while calculation error can be assessed from the Comp phase of the calculation task. Experiment 1 demonstrates that CalcOnlyForce < CalcProbeForce, indicating that during triple-task execution, the calculation time increases if a probe stimulus must be responded to during the Calc phase. This finding is corroborated in experiment 2 ((CalcOnly = CalcOnlyForce) < (CalcForce = CalcProbe = CalcProbeForce)) and experiment 3 (CalcOnly < (CalcGoProbe = CalcNoGoProbe)).

This pattern of results seems feasible, as the PRT acts as a discrete task that temporarily diverts capacity, which must then be allocated away from the other tasks. Given that the Calc phase of the CLC is a task with FlexTR, it is reasonable that the time required to complete the task extends to accommodate the increased demands on the processing capacity. This finding aligns with the observation that, for reaction time and error in the PRT, performance in the CalcPhase events does decline compared to the single task, but the decrement is not as pronounced as, for example, in the Comp phase of the CLC. This suggests that while a probe stimulus can be used to assess capacity utilization in a primary task, the primary task itself may also exhibit a performance decrement if it has a FlexTR.

When considering the pattern of results for calculation error from experiments 2 and 3 for the CLC, the importance of the flexibility of the time regimes of the tasks becomes even more apparent. Neither experiment showed statistically significant differences in calculation error in the ERA. This can be attributed to the fact that the Comp phase is a task with a fixed TR. This is due to the limited duration of the stimulus, specifically the display of a new digit that must be verified as correct or incorrect. Although the FRC also has a fixed TR where errors can occur, these errors may not be as explicitly recognizable for the subjects. Therefore, it can be assumed that subjects are most likely to anticipate errors in the Comp phase as a negative consequence and thus strive to avoid them at all costs. This means that performance remains stable across all events (the area in the TR model does not change), and the other tasks must be adjusted to meet the processing capacity (PC) requirements.

11.3 Force Tracking Task in the Time Regime Model

For the ERA of the FRC, the motor error as normalized RMSE is used as the dependent variable. In experiment 1, the results show that (CalcProbeForce = CompProbeForce) < (CalcOnlyForce = CompOnlyForce). Experiment 2 exhibits a similar pattern: (CompProbeForce = CalcProbeForce) < ProbeForce. Surprisingly, the motor error appears to decrease in the ERA as the number of additional tasks increases, contrary to the expected increase in motor error. The reason for this finding is not clear. However, the anticipation of negative consequences from errors (i.e., an increase in RMSE) could lead to the prioritization of the FRC task. This would further support the notion that tasks with a fixed TR are allocated the greatest processing capacity in terms of error avoidance, as seen in the Comp phase of the CLC.

Additionally, the ERA for the motor error allows for the examination of differences in the PreStim, PostStim, and PostReact intervals of the probe stimuli from PRT. In experiment 2, this analysis yielded the result (PreStim = PostStim) < PostReact. In other words, the RMSE increases in a time-discrete manner, particularly after responding to a probe stimulus. This indicates that effect monitoring is occurring, which appears to require significantly more capacity than the other phases of task execution of FRC, such as perception, response selection, and response execution. This finding is also supported by the results from experiment 3: (GoPreStim = NoGoPreStim = NoGoPostStim) < (GoPostStim = GoPostReact). Notably, during the PostStim phase, which encompasses response selection until response execution, the RMSE places similarly high demands as the PostReact region during effect monitoring.

11.4 Tasks under the Time Regime Model

Upon reviewing the results, the three tasks can now be integrated under dual and triple task conditions and roughly classified within the time regimes model to visualize the findings succinctly. It is crucial to reiterate that the fundamental assumption is that performance generally deteriorates in the transition from single-task to dual-task and triple-task conditions across all analyzed variables and both measured dimensions. However, it is essential to focus specifically on the comparisons within the different events (see Figure 101). Therefore, DTmc is not included here either, as although the ST to DT comparison showed a performance decrement in both dependent variables in both dimensions, the event-related analysis within the multi-tasking conditions only revealed that the extended calculation time differs significantly from the other events. This aligns with the assumption that an extension of calculation time can be tolerated, as it is of a FlexTR, and both the calculation errors in the CompPhase and the RMSE in the tracking task could be minimized as much as possible.

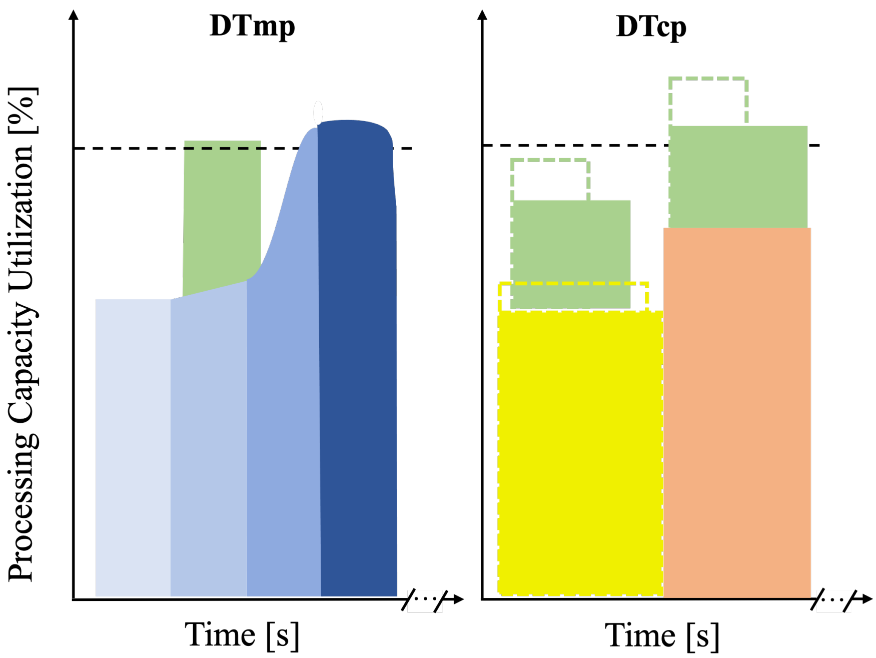

Figure 101: Time Regime Model for the combination of FRC with PRT (DTmp) and CLC with PRT (DTcp) as a suggested organizational structure for processing capacity allocation during multiple-task execution as indicated by experiments 1-3.

Interestingly, the combination of FRC and PRT does not result in an increase in reaction time in the results pattern of the multitasking comparisons, but rather in an increase in reaction error and RMSE. The latter specifically caused by a decreased tracking performance during the post stimulus phase of the probe stimulus indicating that effect monitoring of the probe response might be responsible for this effect. One possible explanation for the indifferent probe reaction times is that the FRC already consumes a significant amount of processing capacity, leading to delayed responses to the stimulus slower than two seconds, which are then registered as errors. Assuming that tasks with a FlexTR are always extended in the time dimension first, this would provide a logical explanation.

This aligns with the findings of a study by Langhanns & Müller (2020), which examined the integration of a walking task with an n-back task. In their study, the stimuli for the n-back task were presented either at consistent, fixed events during the walking task or at random intervals. The results indicated that there was no task interference when stimuli appeared at predictable intervals during walking. It is suggested that when stimulus timing is known in advance, processing resources can be allocated effectively, allowing the walking task to operate significantly below the PC_{100} threshold and thus preserve sufficient capacity for the anticipated n-back task. However, if stimuli are presented irregularly and unpredictably, this organized allocation is not feasible, potentially leading to inadequate processing capacity. The same phenomenon is anticipated in the current work with the combination of FRC and PRT tasks. Due to the unpredictable timing of stimuli, prescheduled resource allocation (a term coined by Langhanns & Müller, 2020) is unattainable, quickly depleting the PC100 and potentially causing errors in both tasks.

The analysis of the DTcp result pattern also adheres to the suggested strategy where tasks with a Flex TR are extended to prevent execution errors. It is evident that both tasks or phases of tasks with a FlexTR are extended to avoid or minimize errors as much as possible. In the CalcPhase, this strategy is successful, as no increase in reaction error is observed. However, it is important to note that no errors are measured in the CalcPhase for the CLC, but only in the CompPhase, which corresponds to the calculation error. Nevertheless, performance can be kept below the processing capacity limit (PC_{100}) compared to other task combinations. Conversely, the processing of the PRT task cannot be extended sufficiently to avoid errors, resulting in an increase in reaction error.

For the combination of all three tasks into a triple task condition, the figure below (see Figure 102) again differentiates between the TT condition during the CalcPhase of the CLC (left) and the TT condition during the CompPhase of the CLC (right).

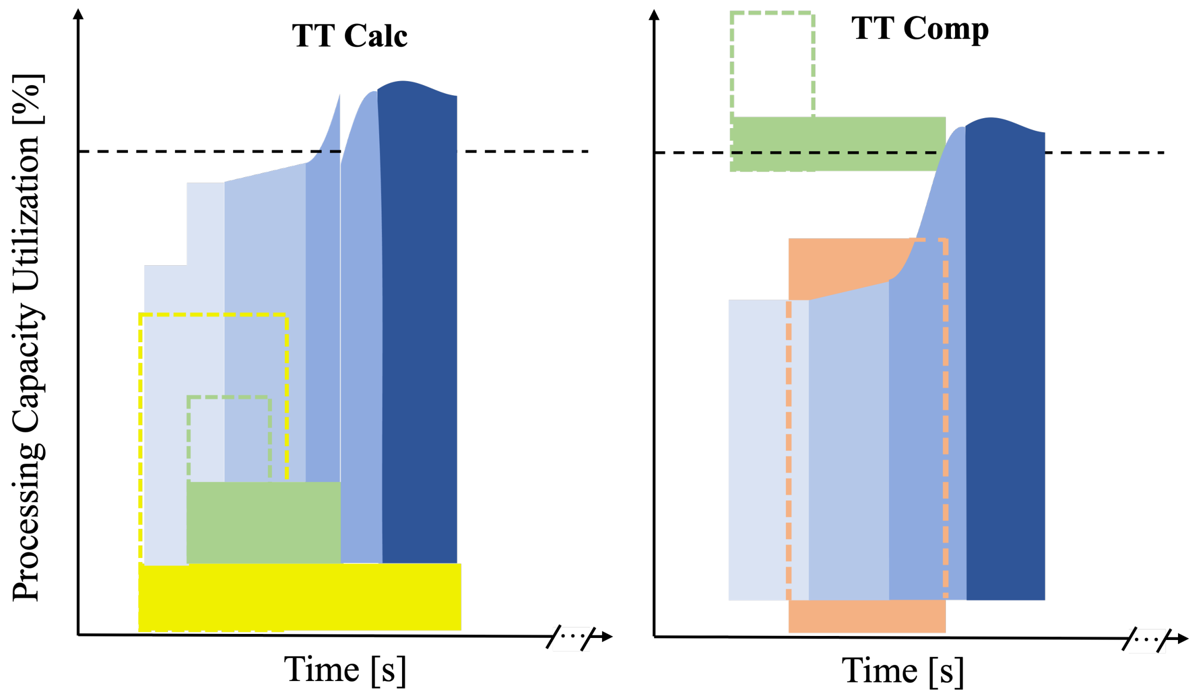

Figure 102: Time Regime Model of triple-task execution for the combination of the calculation phase of CLC with FRC and PRT (TT Calc) and the comparison phase of CLC with FRC and PRT (TT Comp) by result pattern yielded from experiments 1-3. Note that the model cannot sufficiently visually account for the result pattern found for TT Comp.

Here, it is also evident that the evidence-based assumption that tasks with a FlexTR are always extended first to avoid PC_{100} overruns as much as possible holds true. Consequently, the CalcPhase of the CLC and the response to the PRT are significantly extended. However, the timely extensions do not suffice completely to prevent an increase in the motor error for FRC.

Such an extension is not feasible for the CompPhase of CLC. This results in a problem in the representation in the image on the right and thus also a problem for the model. The fixed Time Regime (FixTR) CompPhase and the FRC cannot be combined without the individual tasks overlapping. This issue may arise because the illustration relies on theoretically assumed areas rather than the actual processing capacities required. However, this limitation suggests that PCU_{100} may have been presented in a too one-dimensional manner up to this point. A new proposal for expanding the model would therefore be to also integrate the negative range of the y-axis as shown in Figure 103.

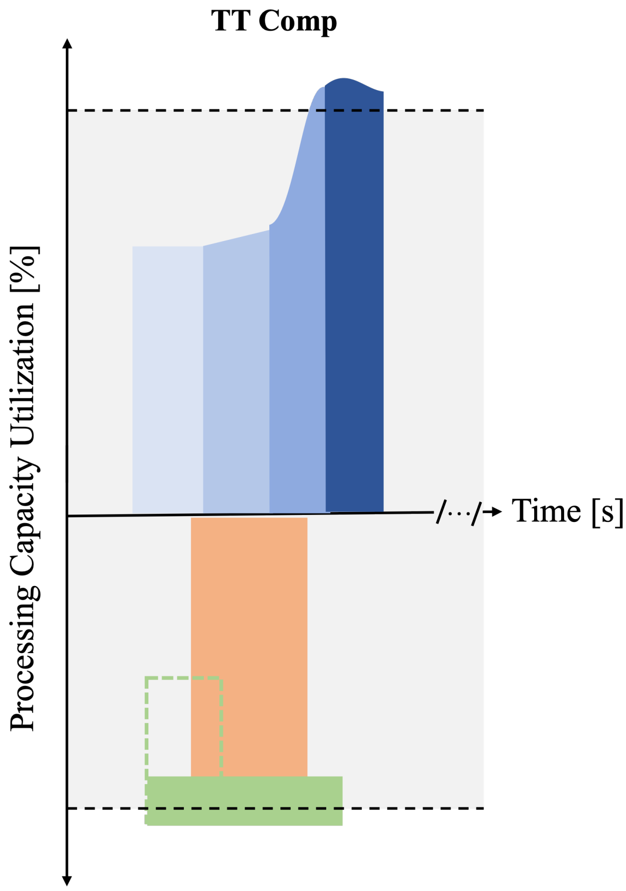

Figure 103: Time Regime Model for the combination of the calculation phase of CLC with FRC and PRT (TT Comp) with a new proposed addition to the model, where the x-axis is used to divide between tasks. The PC_{100} lies between both dashed lines as indicated by the grey area.

This new visualization results in a PC_{100} as depicted by the gray area. As illustrated in the figure, the x-axis could merely separate tasks visually, but this would also imply that errors (PC_{100} depletion) could occur in two tasks simultaneously. This could be interpreted as evidence that there must be several different resources (as proposed by Wicken's (2002) Multiple Resource Model) from which capacity is drawn, and that different types of errors can occur when both resources are exhausted.

Previously, the model featured only two axes. Therefore, it is not yet possible to fully capture what happens when errors occur in two tasks being processed simultaneously. The representation for DT was adequate for mapping the performance of the three tasks used here. However, this representation was not quite sufficient to display the results for the TT condition with the Comp phase of the CLC. With the newly selected representation, it is now possible to investigate whether the limits result from the same (as proposed by Kahneman's (1973) Capacity Theory) or different resources (Multiple Resource Model). However, this question was initially neglected in the experiments conducted so far and therefore not visualized accordingly.

Terms and concepts such as task prioritization, compatibility or congruency, and automation are also not yet incorporated into the model. However, the experiments based on the TR model were initially designed to break down the more precise allocation of resources over time. Therefore, these fundamental concepts were initially excluded. Nevertheless, it is conceivable that the model can be applied to other issues with different paradigms, as it is also based on capacity and multiple resource models and would therefore predict similar results. The gain in knowledge to this point is derived solely from a better temporal resolution. Therefore, further experiments could now extend this approach to other paradigms, taking into account the TR model for the tasks to be performed.

12 Conclusion

Earlier studies in cognitive psychology have employed various paradigms to demonstrate significant performance declines in multitasking due to task interference. Numerous models have been utilized repeatedly to explain the mechanisms behind this interference. Nevertheless, none of these models offers a universally applicable and transferable explanation for how processing capacity is allocated. Additionally, given that many studies focus predominantly on discrete tasks or that continuous task performance metrics are typically aggregated at the level of trials or blocks, it has been challenging to precisely describe the temporal dynamics of resource allocation. To address this gap, the time regimes model was developed in this research. This model enables a two-dimensional description of task performance across different processing stages, primarily achieving high temporal resolution by segmenting tasks into distinct events during concurrent processing.

The experiments aimed to gain a clearer understanding of how motor-cognitive multitasking leads to interference within a self-organized framework. Therefore, three experiments combined a probe reaction time, calculation, and force tracking task in different single, dual, and triple-task conditions. Auditory stimuli with specific semantics were used in the probe task to test for task interference at the linguistic level. However, for the latter, no evidence could be found.

The three conducted experiments revealed that assumptions made under the time regimes model allow for an effective representation of the flexible and dynamic allocation of processing capacity. The distinction between the temporal and processing capacity utilization dimensions is crucial, as both performance metrics, despite their interdependence, exhibit divergent behaviors across various contexts and task characteristics. Moreover, tasks can be categorized into fixed and flexible time regimes. In the flexible time regime, processing duration is preferentially extended to prevent errors, notably as exceeding the PC_{100}, emphasizing error avoidance as the highest priority. This implies an additional layer of prioritization, where tasks with clearly visible error consequences are executed without limitation. Consequently, other tasks must be allocated within the constraints of available time and PCU_{100}. Tasks may also require varying levels of processing capacity at different processing stages, underscoring the necessity for continuous monitoring and analysis of performance changes in both dimensions, similar to the approach adopted in the event-related analysis.

It is essential to first establish foundational assumptions regarding the processing phases and their utilization concerning time and processing capacity for each task type. This approach facilitates the identification of specific events and the examination of task interferences at these distinct time points. Additionally, consideration should be given to whether the task involves a fixed or flexible time regime. If possible, logical considerations should also be made in advance regarding which of the four processing phases represents a particularly high or low processing requirements for the task (and also in connection with other tasks).

Beurskens and Bock (2013), for example, also demonstrated that task interference arising from the combination of different walking conditions with various cognitive tasks strongly depends on the specific characteristics of the tasks involved. When calculating dual-task costs (McDowd, 1986), it became evident that, during walking, step duration - representing a performance measure in the temporal domain - exhibits particularly high dual-task costs, which in turn affects other gait parameters (Winter et al., 1990) within the PCU dimension but only for specific task combinations. These findings further emphasize the necessity of accurately identifying the specific demands of a task across different processing stages and with regard to its time regime in advance. Moreover, this suggests that future research might benefit from calculating dual-task costs not only in aggregated blocks but also for discrete events, thereby enabling a more detailed analysis of how cognitive tasks modulate gait in the temporal domain and in parameters informative of the PCU dimension.

The approach to creating a very precise requirements profile for a task in terms of processing capacity utilization (PCU) and time dimensions, such as requirements throughout processing stages and temporal flexibility (e.g., fixed or flexible time regime), is based on the idea of creating a requirements profile for a specific action using a coordination requirements controller according to Neumaier (2016). This allows, through a precise analysis of information requirements and pressure conditions, the identification of areas that place high demands. These areas can then be specifically trained, thereby improving performance in a probationary situation. Transferable to multi-tasking research, this also means that, through the precise classification of a task within the model, high PC utilization and task interferences can be better predicted or even explicitly induced.

The results of the three experiments presented in this work, considering the time regimes, suggest that tasks with a flexible time regime are always extended to avoid errors, while tasks with a fixed time regime are significantly affected by even minor adjustments to task complexity.

Further investigations employing event-related analysis are necessary to determine if this analytical approach applies to other task types and is capable of addressing additional research questions. One question that remains unanswered and was neglected in this work is whether the model can actually represent the different resources of various modalities. Given that the analysis has only been utilized in three experiments thus far, the current findings have limited applicability and generalization. Consequently, this represents a potential direction for future research endeavors.

13 References

Arbib, M. A. (2003). The Evolving Mirror System: A Neural Basis for Language Readiness. In M. H. Christiansen & S. Kirby (Eds.), *Language Evolution* (pp. 182–200). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199244843.003.0010

Attneave, F. (1959). *Applications of Information Theory to Psychology*. Henry Holt and Company.

Beauchet, O., Dubost, V., Aminian, K., Gonthier, R., & Kressig, R. W. (2005). Dual-task-related gait changes in the elderly: Does the type of cognitive task matter? *Journal of Motor Behavior*, *37*(4), 259–264.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity.* McGraw-Hill Book Company. https://doi.org/10.1037/11164-000

Bernstein, N. A. (1967). *The Coordination and Regulation of Movements.* Pergamon Press.

Beurskens, R., & Bock, O. (2013). Does the walking task matter? Influence of different walking conditions on dual-task performances in young and older persons. *Human Movement Science*, *32*(6), 1456–1466. https://doi.org/10.1016/j.humov.2013.07.013

Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler: Mit 163 Tabellen* (7., vollst. überarb. und erw. Aufl). Springer.

Brand, T. K., Maurer, L. K., Müller, H., Döhring, F. R., & Joch, M. (2022). Predictability shapes movement kinematics and grip force regulation in human object handovers. *Human Movement Science*, *85*, 102976. https://doi.org/10.1016/j.humov.2022.102976

Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, *47*(3), 191–196. https://doi.org/10.1037/h0054182

Broadbent, D. E. (1958). *Perception and Communication: Applied Psychology Research Unit of the Medical Research Council*. Pergamon Press.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., & Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study: Cortical activation during action observation. *European Journal of Neuroscience*, *13*(2), 400–404. https://doi.org/10.1111/j.1460-9568.2001.01385.x

Campbell, J. I. D. (2009). Selective effects of number processing context on number naming. *European Journal of Cognitive Psychology*, *21*(5), 686–702. https://doi.org/10.1080/09541440802214846

Campbell, J. I. D., & Metcalfe, A. W. S. (2008). Arabic digit naming speed: Task context and redundancy gain. *Cognition*, *107*(1), 218–237. https://doi.org/10.1016/j.cognition.2007.10.001

Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, *25*(5), 975–979. https://doi.org/10.1121/1.1907229

Cipolotti, L., & Butterworth, B. (1995). Toward a multiroute model of number processing: Impaired number transcoding with preserved calculation skills. *Journal of Experimental Psychology: General*, *124*(4), 375–390. https://doi.org/10.1037/0096-3445.124.4.375

Cisek, P., & Kalaska, J. F. (2010). Neural Mechanisms for Interacting with a World Full of Action Choices. *Annual Review of Neuroscience*, *33*(1), 269–298. https://doi.org/10.1146/annurev.neuro.051508.135409

Damian, M. F. (2004). Asymmetries in the processing of Arabic digits and number words. *Memory & Cognition*, *32*(1), 164–171. https://doi.org/10.3758/BF03195829

De Marco, D., De Stefani, E., Bernini, D., & Gentilucci, M. (2018). The effect of motor context on semantic processing: A TMS study. *Neuropsychologia*, *114*, 243–250. https://doi.org/10.1016/j.neuropsychologia.2018.05.003

Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Rev. and updated ed). Oxford University Press.

Dehaene, S., & Cohen, L. (1997). Cerebral Pathways for Calculation: Double Dissociation between Rote Verbal and Quantitative Knowledge of Arithmetic. *Cortex*, *33*(2), 219–250. https://doi.org/10.1016/S0010-9452(08)70002-9

Dehaene, S., Molko, N., Cohen, L., & Wilson, A. J. (2004). Arithmetic and the brain. *Current Opinion in Neurobiology*, *14*(2), 218–224. https://doi.org/10.1016/j.conb.2004.03.008

Desantis, A., Roussel, C., & Waszak, F. (2014). The temporal dynamics of the perceptual consequences of action-effect prediction. *Cognition*, *132*(3), 243–250.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*(1), 80–90. https://doi.org/10.1037/h0039515

Ells, J. G. (1969). *Attention requirements of movement control*. University of Oregon (unpublished).

Ells, J. G. (1973). Analysis of temporal and attentional aspects of movement control. *Journal of Experimental Psychology*, *99*(1), 10–21. https://doi.org/10.1037/h0034740

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (n.d.). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381–391. https://doi.org/10.1037/h0055392

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593–609. https://doi.org/10.1093/brain/119.2.593

Gold, B., & Radar, C. M. (1969). *Digital Proessing of Signals*. McGraw-Hill.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25. https://doi.org/10.1016/0166-2236(92)90344-8

Grafton, ScottT., Arbib, MichaelA., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography: 2. Observation compared with imagination. *Experimental Brain Research*, *112*(1). https://doi.org/10.1007/BF00227183

Interlink Electronics. (2023). *FSR Force Sensing Resistor Integration Guide and Evaluation Parts Catalog*. https://cdn-learn.adafruit.com/assets/assets/000/010/126/original/fsrguide.pdf

James, W. (1950). *The principles of psychology: In two volumes. Vol. 1.*(Facsim. of ed. New York, Henry Holt, 1890). Dover.

Jentzsch, I., Leuthold, H., & Ulrich, R. (2007). Decomposing sources of response slowing in the PRP paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(3), 610–626. https://doi.org/10.1037/0096-1523.33.3.610

Jiang, T., Zhang, W., Wen, W., Zhu, H., Du, H., Zhu, X., Gao, X., Zhang, H., Dong, Q., & Chen, C. (2016). Reevaluating the two-representation model of numerical magnitude processing. *Memory & Cognition*, *44*(1), 162–170. https://doi.org/10.3758/s13421-015-0542-2

Kaas, J. H. (1997). Topographic Maps are Fundamental to Sensory Processing. *Brain Research Bulletin*, *44*(2), 107–112. https://doi.org/10.1016/S0361-9230(97)00094-4

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kaye, D. B., deWinstanley, P., Chen, Q., & Bonnefil, V. (1989). Development of Efficient Arithmetic Computation. *Journal of Educational Psychology*, *81*(4), 467–480.

Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking—An integrative review of dual-task and task-switching research. *Psychological Bulletin*, *144*(6), 557–583. https://doi.org/10.1037/bul0000144

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons. *Science*, *297*(5582), 846–848. https://doi.org/10.1126/science.1070311

Künstler, E. C. S., Finke, K., Günther, A., Klingner, C., Witte, O., & Bublak, P. (2018). Motor-cognitive dual-task performance: Effects of a concurrent motor task on distinct components of visual processing capacity. *Psychological Research*, *82*(1), 177–185. https://doi.org/10.1007/s00426-017-0951-x

Laessoe, U., Hoeck, H. C., Simonsen, O., & Voigt, M. (2008). Residual attentional capacity amongst young and elderly during dual and triple task walking. *Human Movement Science*, *27*(3), 496–512. https://doi.org/10.1016/j.humov.2007.12.001

Langhanns, C., Ewolds, H., Künzell, S., & Müller, H. (2022). Contributions of open-loop and closed-loop control in a continuous tracking task differ depending on attentional demands during practice. *Human Movement Science*, *85*, 103001. https://doi.org/10.1016/j.humov.2022.103001

Langhanns, C., & Müller, H. (2020). Prescheduled Interleaving of Processing Reduces Interference in Motor-Cognitive Dual Tasks. *Journal of Cognition*, *3*(1), 33. https://doi.org/10.5334/joc.122

Laurienti, PaulJ., Kraft, RobertA., Maldjian, JosephA., Burdette, JonathanH., & Wallace, MarkT. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research*, *158*(4). https://doi.org/10.1007/s00221-004-1913-2

Lund, N. (2001). *Attention and pattern recognition*. Routledge ; Taylor & Francis.

Mattes, S. (2001). Response force in probe reaction time tasks: Reduction in capacity, response inhibition or bottle neck? *Zeitschrift Fur Experimentelle Psychologie: Organ Der Deutschen Gesellschaft Fur Psychologie*, *48*(3), 201–206.

McDowd, J. M. (1986). The Effects of Age and Extended Practice on Divided Attention Performance. *Journal of Gerontology*, *41*(6), 764–769. https://doi.org/10.1093/geronj/41.6.764

McLeod, P. (1978). Does Probe RT Measure Central Processing Demand? *Quarterly Journal of Experimental Psychology*, *30*(1), 83–89. https://doi.org/10.1080/14640747808400656

Miller, E. K., & Cohen, J. D. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience*, *24*(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, *6*, 414–417. https://doi.org/10.1016/0166-2236(83)90190-X

Müsseler, J., & Rieger, M. (Eds.). (2017). *Allgemeine Psychologie* (3. Auflage). Springer.

Neely, G., Sörman, D. E., & Ljungberg, J. K. (2018). The impact of spoken action words on performance in a cross-modal oddball task. *PLOS ONE*, *13*(11), e0207852. https://doi.org/10.1371/journal.pone.0207852

Neumaier, A. (2016). *Koordinatives Anforderungsprofil und Koordinationstraining: Grundlagen, Analyse, Methodik* (5., korr. Aufl). Sportverlag Strauß.

Newell, A., & Rosenbloom, P. (1981). Mechanisms of Skill Acquisition and The Law of Practice. In J. R. Anderson (Ed.), *Cognitive Skills and Their Acquisition* (pp. 1–55). Erlbaum.

Noël, M.-P., Fias, W., & Brysbaert, M. (1997). About the influence of the presentation format on arithmetical-fact retrieval processes. *Cognition*, *63*(3), 335–374. https://doi.org/10.1016/S0010-0277(97)00009-7

Nolting, W. (2018). *Grundkurs Theoretische Physik 1: Klassische Mechanik und mathematische Vorbereitungen* (11. Aufl. 2018). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-57584-0

Norman, D. A. (1969). *Memory and attention: An introduction to human information processing*. Wiley.

Pashler, H. (Ed.). (2004). *Attention* (repr). Psychology Press.

Passingham, R. E., & Toni, I. (2001). Contrasting the Dorsal and Ventral Visual Systems: Guidance of Movement versus Decision Making. *NeuroImage*, *14*(1), S125–S131. https://doi.org/10.1006/nimg.2001.0836

Penfield, W., & Boldrey, E. (1937). Somatic Motor and Sensory Representation in the Cerebral Cortex of Man as Studied by Electrical Stimulation. *Brain*, *60*(4), 389–443. https://doi.org/10.1093/brain/60.4.389

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review*, *78*(5), 391–408. https://doi.org/10.1037/h0031333

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems: Functional links between motor and language systems. *European Journal of Neuroscience*, *21*(3), 793–797. https://doi.org/10.1111/j.1460-9568.2005.03900.x

Rasch, B., Friese, M., Hofmann, W., & Naumann, E. (2008). *Quantitative Methoden. Bd. 2: Mit 29 Abbildungen und 61 Tabellen* (2., erw. Aufl., korr. Nachdr). Springer Medizin Verl.

Reynvoet, B., Brysbaert, M., & Fias, W. (2002). Semantic priming in number naming. *The Quarterly Journal of Experimental Psychology Section A*, *55*(4), 1127–1139. https://doi.org/10.1080/02724980244000116

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, *21*(5), 188–194. https://doi.org/10.1016/S0166-2236(98)01260-0

Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, *27*(1), 169–192. https://doi.org/10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*(2), 131–141. https://doi.org/10.1016/0926-6410(95)00038-0

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., & Fazio, F. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental Brain Research*, *111*(2). https://doi.org/10.1007/BF00227301

Rüschemeyer, S.-A., Brass, M., & Friederici, A. D. (2007). Comprehending Prehending: Neural Correlates of Processing Verbs with Motor Stems. *Journal of Cognitive Neuroscience*, *19*(5), 855–865. https://doi.org/10.1162/jocn.2007.19.5.855

Schieber, M. H. (2001). Constraints on Somatotopic Organization in the Primary Motor Cortex. *Journal of Neurophysiology*, *86*(5), 2125–2143. https://doi.org/10.1152/jn.2001.86.5.2125

Schliephake, A., Bahnmueller, J., Willmes, K., Koch, I., & Moeller, K. (2022). Cognitive control in number processing: New evidence from number compatibility effects in task-switching. *Cognitive Processing*, *23*(2), 191–202. https://doi.org/10.1007/s10339-022-01074-5

Schott, N., & Klotzbier, T. J. (2018). Profiles of Cognitive-Motor Interference During Walking in Children: Does the Motor or the Cognitive Task Matter? *Frontiers in Psychology*, *9*, 947. https://doi.org/10.3389/fpsyg.2018.00947

Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. *Experimental Brain Research*, *185*(3), 359–381. https://doi.org/10.1007/s00221-008-1280-5

Shkuratova, N., Morris, M. E., & Huxham, F. (2004). Effects of age on balance control during walking. *Archives of Physical Medicine and Rehabilitation*, *85*(4), 582–588. https://doi.org/10.1016/j.apmr.2003.06.021

Simon, J. R., Rudell, A. P., Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*(3), 300–304. Business Source Premier.

Stanciu, I., Biehl, S. C., & Hesse, C. (2017). Increased cognitive demands boost the spatial interference effect in bimanual pointing. *Psychological Research*, *81*(3), 582–595. https://doi.org/10.1007/s00426-016-0762-5

Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology*, *14*(1), 1–36. https://doi.org/10.1037/h0073262

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. F., & Perani, D. (2005). Listening to Action-related Sentences Activates Fronto-parietal

Motor Circuits. *Journal of Cognitive Neuroscience*, *17*(2), 273–281. https://doi.org/10.1162/0898929053124965

Treisman, A. M. (1964). Verbal Cues, Language, and Meaning in Selective Attention. *The American Journal of Psychology*, *77*(2), 206. https://doi.org/10.2307/1420127

Ulrich, R., Fernández, S. R., Jentzsch, I., Rolke, B., Schröter, H., & Leuthold, H. (2006). Motor Limitation in Dual-Task Processing Under Ballistic Movement Conditions. *Psychological Science*, *17*(9), 788–793. https://doi.org/10.1111/j.1467-9280.2006.01783.x

Verghese, A., Mattingley, J. B., Palmer, P. E., & Dux, P. E. (2018). From eyes to hands: Transfer of learning in the Simon task across motor effectors. *Attention, Perception, & Psychophysics*, *80*(1), 193–210. https://doi.org/10.3758/s13414-017-1427-1

Welch, J. C. (1898). On the Measurement of Mental Activity through Muscular Activity and the Determination of a Constant of Attention. *American Journal of Physiology*, *3*(1), 24.

Welford, A. T. (1952). The "psychological refractory period" and the timing of high speed performance—A review and a theory. *British Journal of Psychology*, *43*, 2–19.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

Wickens, C. D., & Liu, Y. (1988). Codes and Modalities in Multiple Resources: A Success and a Qualification. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *30*(5), 599–616. https://doi.org/10.1177/001872088803000505

Winter, D. A., Patla, A. E., Frank, J. S., & Walt, S. E. (1990). Biomechanical Walking Pattern Changes in the Fit and Healthy Elderly. *Physical Therapy*, *70*(6), 340–347. https://doi.org/10.1093/ptj/70.6.340

Winter, D. A., Sidwall, H. G., & Hobson, D. A. (1974). Measurement and reduction of noise in kinematics of locomotion. *Journal of Biomechanics*, *7*(2), 157–159. https://doi.org/10.1016/0021-9290(74)90056-6

Wirth, R., Janczyk, M., & Kunde, W. (2018). Effect monitoring in dual-task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(4), 553–571. https://doi.org/10.1037/xlm0000474

14 List of Publications

Munzert, J., Müller, J., Joch, M.,Reiser, M. (2018). Specificity of Postural Control: Comparing Expert and Intermediate Dancers. *Journal of Motor Behavior*, 51(3), 1-13. https://doi.org/10.1080/00222895.2018.1468310?urlappend=%3Futm_source%3Dresearchgate.net%26utm_medium%3Darticle

15 Declaration of Independence

I hereby declare that I have prepared the thesis at hand independently and without undue aid or the use of any resources other than indicated within the thesis. All parts of my thesis taken either verbatim or analogously from the published or unpublished works of or based on oral communications with others are indicated as such. Regarding all aspects of my scientific enquiries as they appear in my thesis, I have upheld the tenets of good scientific practice as laid out in the "Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" and complied with the precept of ethics, data protection and animal welfare. I declare that I have neither directly nor indirectly given monetary or any other valuable considerations to others in connection with the thesis at hand. I declare that I have not presented the thesis at hand, either in an identical or similar form, to an examination office or agency in Germany or any other country as part of any examination or degree. All materials from other sources as well as all works performed by others used or directly referenced within the thesis at hand have been indicated as such. In particular, all persons involved directly or indirectly in the development of the thesis at hand have been named. I agree with the screening of my thesis for plagiarism via offline or online detection-software.

I further declare that the AI tool ChatGPT was used for some paragraphs to paraphrase the written text. In addition, Grammarly software was used to check and correct spelling and grammar.

Location, Date　　　　　　　　　　Jelena Müller