# The Psychological Investigation of Subjective Time

-

# New Concepts and Contemporary Research Methods

Inauguraldissertation

zur Erlangung des Akademischen Grades

eines Dr. phil.

vorgelegt dem Fachbereich 02 Sozialwissenschaften, Medien und Sport

der Johannes Gutenberg-Universität Mainz

von

M.Sc.-Psych. Sven Thönes

aus Lünen

2017

Tag des Prüfungskolloquiums: 24.08.2017

## Abstract

In contrast to physical (objective) time, the subjective representation or perception of time is influenced by various factors. For example, in everyday life, humans perceive time to pass slower or faster depending on the situational context, and this impression may lead to an over- or underestimation of duration. Based on five sub-projects, the present work investigates several different phenomena in the field of subjective time. The first two sub-projects (meta-analytical reviews) focused on time perception in clinical populations. The results indicate that time passes less quickly for depressive patients. Judgments of duration, however, do not differ between patients and healthy control subjects. Patients with schizophrenia show a strong increase in the variability of their duration judgments, whereas mean duration estimates do not differ between patients and controls. In sub-projects three and five, I investigated whether subjective time can by influenced by (subtle) interpersonal cues. The results indicate rather weak effects of gaze direction on duration judgments and the spatialized representation of time (mental time lines). Sub-project four comprised two experiments on the effects of chronometric counting on the production of time intervals. Interestingly, while productions became less variable (more precise), the accuracy of time productions did not benefit from counting. Beyond the specific foci of the different sub-projects, the results from the present work approach some fundamental aspects of subjective time. Based on the meta-analytical review of time perception in depressive patients, factors that influence the subjective passage of time do not necessarily affect the perception of duration. Therefore, passage and duration need to be conceptualized as disentangled temporal dimensions. The concept of passing time may be more closely related to the concept of a mental time than to the concept of duration. The meta-analytical review on time perception in patients with schizophrenia and the experiments on the effects of counting on time productions show that measures of accuracy and measures of precision reflect independent aspects of subjective time, which need to be considered carefully in models of time perception and by future research.

## TABLE OF CONTENTS

# 1. GENERAL INTRODUCTION

In the first part of the General introduction section (1.1), I will introduce the major terms and definitions that provide the conceptual frame for the empirical investigation of subjective time. Within this frame, the second part (1.2) presents an outline of the Main section that includes two meta-analytical reviews and three experimental studies. The third part (1.3) provides an overview of the empirical research methods that are used to investigate subjective time.

## 1.1 Major concepts and definitions

Time is regarded as a fundamental property of our world. The history of philosophy reveals a controversy about the nature of time that dates back thousands of years. Ever since, a major concern has been the debate on whether time is an absolute entity that exists independently or whether time depends on other entities and properties, such as movement, space, or a (conscious) observer.

In this regard, the concept of time by Aurelius Augustinus represents an early and radically subjectivistic view. In book XI of his confessions (Augustinus, ca. 400/ 2000), he investigates the concept of time and argues that time tends to nonexistence on an objective level. On a subjective level, however, time is constructed by the human mind and comes into being as the subject remembers the past, experiences the present, and anticipates the future.

Contrary to Augustinus, especially in the modern era, objectivistic theories of time have been formulated. Probably, the most influential concept of objective and absolute time has been brought forward by Isaac Newton (Newton, 1687/ 1972). As the fundamental element of his natural philosophy, Newton constitutes time (and space) as independent of all other properties. In the Newtonian concept, time is objectively real and absolute, and its passage is perfectly continuous and independent of the perceiving subject.

In modern cognitive science and psychology, an objective (observer-independent) reality of time is generally accepted. It is regarded as common sense that conscious beings came into existence after the universe had evolved through time for billions of years. Notwithstanding, the main focus on time in cognitive science and psychology lies on its subjective element, which I will refer to as *subjective time*. The present work argues that *subjective time* encompasses the processing of (basic) temporal information (I will refer to as *temporal processing),* the *perception of time* in terms of *duration* (sometimes referred to as timing) and *passage* (flow)*,* and the (spatialized) representation of time, which is reflected in the *mental time line* (see Figure 1). Based on these considerations, cognitive science and psychology represent an intermediate position between the two radical views of subjectivism on the one hand, and objectivism on the other hand. Over the last decades, empirical psychology has developed various methods and models, such as the concept of an internal clock (Treisman, 1963), in order to investigate the relationship between objective time and *subjective time*.

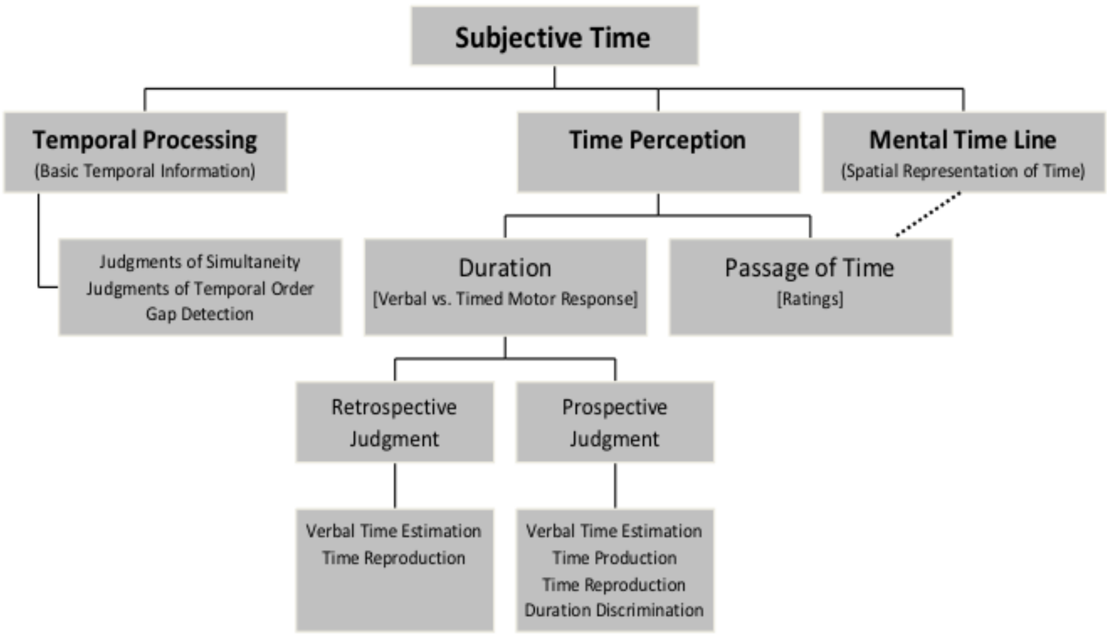


**Figure 1.** Conceptual scheme of subjective time.

## 1.2 Outline of the present work

Briefly within this General introduction section, and in more detail in the different parts of the Main section, I will introduce the current cognitive models of *subjective time* as well as the psychological methods that are used to investigate *subjective time* and its influencing factors. By means of two meta-analytical reviews of the existing literature, the first (2.1) and second (2.2) part of the Main section will focus on *time perception* and *temporal processing* in clinical populations (depressed and schizophrenic patients) in comparison to healthy subjects. The third part of the Main section (2.3) will present the results from three experiments on the effects of gaze orientation on *time perception* (*judgments of duration*) in interpersonal situations. In these parts, models and research methods of *subjective time* will be illustrated in detail. In the fourth part of the Main section (2.4), two experiments on the effects of chronometric counting on *time perception* are presented. While part one to four focus on *time perception* and *temporal processing*, the fifth part (2.5) approaches the concept of the *mental time line* and its potential interpersonal embodiment. In a final General discussion section (3), the meta-analytical and experimental results will be related to the current state of the art in *subjective time* and its general challenges.

## 1.3 Empirical research methods in subjective time

As an empirical science, psychology is faced with the general question of how *subjective time* can be measured. The basic methodological challenge is related to the fact that, for example, the perception of time is an abstract experience. What can be measured is behavior that is related to this experience. Thus, in a psychological experiment on *subjective time*, the subject is required to show behavior that reflects his or her temporal experience.

### 1.3.1 Temporal processing

On a basic sensory level, humans (as well as other animals) continuously process temporal information from the environment. In this regard, time is not perceived in terms of a conscious temporal experience, but the *temporal processing* ability, for example, the temporal acuity of the visual and the auditory system, is a necessary prerequisite for successful guidance of behavior and survival. The *temporal processing* ability of a subject can be measured, for example, by means of an auditory *gap detection* task. Such a task requires the subject to detect short gaps of silence in a continuous auditory signal, e.g. white noise (Plomp, 1964). The smallest gap that can be detected reliably (e.g., with 75% correct) serves as a measure of performance and indexes the temporal acuity of the auditory system (e.g., Todd, Michie, Budd, Rock, & Jablensky, 2000). Usually, healthy subjects are able to detect auditory gaps of 10 to 20 ms duration (Boehnke & Phillips, 2005). A second task that is used to measure temporal processing abilities in (human) subjects, is the *temporal-order judgment* (TOJ) task, where two different stimuli (e.g., two visual stimuli) are presented successively, and the subject has to indicate which stimulus was presented first. The inter-onset interval (IOI) between the two stimuli at which a certain performance level is reached, e.g., 75% correct responses, serves as a measure of temporal resolution of the sensory system (Hirsh, 1959). Healthy subjects are able to reliably detect the order of two visual stimuli at an IOI of about 50 ms (Capa, Duval, Blaison, & Giersch, 2014; Schmidt, McFarland, Ahmed, McDonald, & Elliott, 2011). A similar task, the *judgment of simultaneity*, requires the subject to decide whether the onsets of two stimuli were synchronous or asynchronous. The subject's performance level is determined based on the IOI between the two stimuli leading to a certain proportion of "simultaneous" responses of, for example, 50%. This is termed the simultaneity threshold or the point of subjective simultaneity (Martin, Giersch, Huron, & van Wassenhove, 2013). The smaller the simultaneity threshold, the higher the temporal resolution of the sensory system.

### 1.3.2 Time perception

Beyond the mere sensory processing of temporal information, in *time perception*, a basic paradigmatic dissociation has been established between verbal statements (ratings) that refer to the experienced *passage of time* on the one hand, and measures that refer to *judgments of duration* on the other hand (e.g., Wearden, 2015).

Statements or ratings on the *passage of time* can be obtained simply by asking the subject how quickly time goes by for him or her in general or within a specific situation, for example during an experiment. Usually, 7 or 9 point Likert scales or visual analogue scales ranging between "very slow" and "very fast" are used to assess ratings of *time passage* (e.g., Bschor et al., 2004).

#### *1.3.2.1 Judgments of duration*

Methods for obtaining *judgments of duration* on the other hand are more diverse (see also Figure 1). A basic distinction has been made between *prospective* and *retrospective duration judgments* (Block & Zakay, 1997; Grondin, 2010). In a *prospective duration judgment* task, the subject is explicitly informed about the *time perception* task prior to the experiment. Thus, the subject will pay attention to time. In a *retrospective duration judgment* task, the subject is not informed prior to but only after the relevant time interval has passed. In such a *retrospective* task, the subject does not explicitly attend to time but has to reconstruct his or her temporal experience from memory in order to give a judgment of duration. Retrospective tasks are necessarily comprised of a single trial only. As the subject knows about the time perception task right after the first trial when an estimate has been given, the task would turn into a prospective (informed) time perception task in a second trial. In a psychological experiment, prospective and retrospective judgments might be affected differently indicating whether attentional (prospective) processes or memory-related (retrospective) processes of time perception (or both) are affected by the experimental manipulation.

Over the last decades, several different tasks have been developed in order to measure a subject's *time perception* in the sense of *duration judgments*. These tasks are mainly: (a) *verbal time estimation* (sometimes referred to as *'time estimation'*), where a time interval is presented, defined for instance by the IOI between two brief tones or flashes, and the subject gives an estimate in conventional time units like seconds, (b) *time production,* where the experimenter specifies a time interval in temporal units, and the subject produces this interval for example by pressing a button to mark the interval's beginning and end, (c) *time reproduction,* where a time interval is presented as in (a) and the subject produces a corresponding interval as in (b), and (d) *duration discrimination tasks*, where the subject has to compare two or more time intervals of almost equal length and is required to select the longer (or shorter) one. Several different versions of duration discrimination tasks are commonly used. For a detailed review of the different (duration discrimination) tasks, please refer to the Introduction section of part two of the Main section (part 2.2.2).

Only *verbal time estimation* and *time reproduction* can be constructed in a *retrospective* manner. All other tasks can be realized only *prospectively*, i.e., by informing the subject about the task at the beginning of the experiment. In the case of *time production*, a subject needs to know that a specific duration is to be produced in order to do so. Hence, *time production* does, in principle, not work retrospectively. *Duration discrimination tasks* cannot be constructed in a retrospective manner because they need to be comprised of several trials, for example in order to determine a difference limen. Retrospective verbal time estimates, on the other hand, are used frequently, for example by asking the subject how much time has elapsed since the beginning of the experiment. This procedure is most common in order to investigate the perception of longer durations up to hours. All other tasks are less applicable in this duration range and are used in the range of milliseconds (mainly *duration discrimination*) up to a few minutes (mainly prospective *verbal time estimation* and *time production*).

Importantly, it is assumed that different mechanisms are involved in *time perception* depending on the length of the time interval to be judged (Bangert, Reuter-Lorenz, & Seidler, 2011; Grondin, 2012; Lewis & Miall, 2003). Similar to basic *temporal processing*, judging intervals in the range of milliseconds is assumed to be based mainly on sensory mechanisms, while cognitive factors like attention and memory become more important in the interval range above one second (e.g., Grondin, 2010).

The behavior of a subject in a duration judgment task reveals different aspects of the subject's perception of time. In general, a subject's performance can be analyzed in terms of *accuracy* and in terms of *precision*. The term *accuracy* refers to the "constant error" (Fechner, 1860), that is, the deviation of *subjective time* (e.g., a verbal estimate, a production, or a reproduction) from the veridical value of the duration to be judged (objective time). Accordingly, the *accuracy* of duration judgments refers to systematic over- and underestimations of the perceived or produced durations. Beside these systematic shifts in duration perception, the data from many duration judgment tasks also provide information about the stability or continuity of *subjective time*. This stability of subjective time is termed *precision* and refers to the "variable error" in terms of Fechner (1860). It is reflected, for example, in the variability of a series of verbal time estimates across several presentations of the same physical duration. Especially part two of the Main section will elaborate on the two concepts of *accuracy* and *precision* and on the related time perception tasks that provide information about these two measures.

Another important issue in the context of empirical *time perception* research are the perceptual and cognitive demands that are related to specific tasks (Gil & Droit-Volet, 2011). For instance, while *time production* and *time reproduction* tasks require timed motor responses, *verbal time estimation* and *duration discrimination* do not. Moreover, memory processes are involved differently depending on the particular task. *Time production* and *verbal time estimation*, on the one hand, require the subject to refer to long-term memory representations of time in terms of chronometric units like seconds or minutes. In *time reproduction* and *duration discrimination*, on the other hand, the information necessary for doing the task is presented within a given trial, or within the experimental block, so that these tasks are likely to depend on working memory rather than on long-term memory.

#### *1.3.2.2 Models of time perception: The Scalar Expectancy Theory*

The forgoing considerations indicate that cognitive processes and *time perception* are closely connected. And accordingly, psychological models of *time perception* and *temporal processing*, such as the *Scalar Expectancy Theory* (SET; Gibbon, 1977; Gibbon, Church, & Meck, 1984; Meck, 1996), explicitly consider attention-, memory-, and decision-related processes. Basically, *SET* assumes an *internal clock* consisting of a *pacemaker* that emits pulses and an *accumulator* (or counter) that collects these pulses. This clock device is integrated into an information processing framework that encompasses attention, memory, and decision stages. According to *SET*, as soon as a subject begins to process an interval, an attentionally modulated *switch* between *pacemaker* and *accumulator* closes. Therefore, the clock pulses emitted by the pacemaker can reach the accumulator, which starts to collect the pulses. The more pulses being accumulated, the longer the perceived length of an interval. The amount of pulses collected in the accumulator is constantly compared to duration samples that are stored in the reference memory. In a *time production* task, for example, a subject will decide to mark the end of the interval to be produced (e.g., “one minute”) as soon as the amount of the pulses accumulated matches the memory representation of the interval to be produced (e.g., the sample of “one minute”). If a subject’s clock runs faster, more pulses get accumulated within a specified interval, and therefore the interval is perceived as longer/produced shorter compared to a subject with a slower clock speed. Accordingly, changes in clock speed affect the *accuracy* of duration judgments causing systematic shifts, such as over- or underestimation of duration. Beside the speed of the internal clock, the pulse-to-pulse variability (clock variability) may differ between subjects (or within subjects between different situations). An increase in clock variability would impair the *precision* of duration judgments (and temporal processing). In a time production task, for instance, an increase in clock variability would be reflected in a larger variability of interval productions across several trials.

According to *SET,* the relationship between objective time and *subjective time* is characterized by two basic properties (Grondin, 2010). First, the mean representation of duration is accurate, that is, mean duration judgments equal the veridical value of a given objective duration (constant error = 0). Second, the variability of represented duration (duration judgments) increases linearly with the mean objective duration, that is, the proportion between variability and mean is constant. In terms of classical psychophysics, this relationship between variability and mean corresponds to Weber’s law (Grondin, 2010; Killeen & Weiss, 1987).

Based on *SET* and extensive empirical research, several factors have been investigated that may influence the perception of time (for a recent review, see Matthews & Meck, 2016). Most importantly, *duration* judgments (and *passage of time* ratings) seem to be affected by the level of (bodily) arousal (Droit-Volet & Meck, 2007; Treisman, 1963) and by attentional resources (Zakay & Block, 1996). Accordingly, these two factors have been integrated into the pacemaker-accumulator models (see, for instance, the attentional-gate model; Block &

Zakay, 1996). Arousal is assumed to directly influence the frequency of the pacemaker. Higher levels of arousal are associated with an increased rate of pulse emission (Angrilli, Cherubini, Pavese, & Manfredini, 1997; Droit-Volet & Meck, 2007; Gil & Droit-Volet, 2012; Mella, Conty, & Pouthas, 2011). If more pulses are emitted by the pacemaker within a specified time interval, more pulses can get accumulated and the interval is perceived to be longer. Attention is considered to alter the accumulation process. For example, if attention is distracted from the time perception task, the accumulator is assumed to miss pulses from the pacemaker, thus producing shorter and more variable duration judgments (Block & Zakay, 1996; S. W. Brown, 1997; Zakay & Block, 1996). The basic model of the internal clock illustrating the role of arousal and attentional processes is presented in Figure 2.

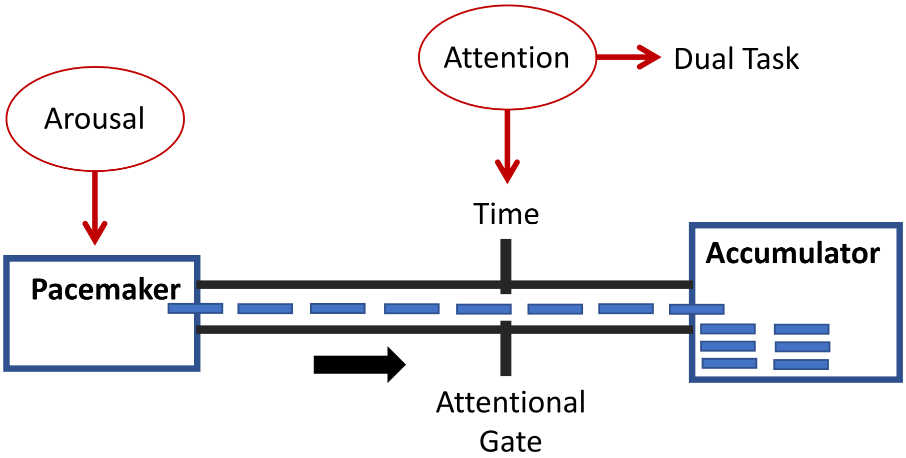


**Figure 2.** Pacemaker-accumulator model of the internal clock illustrating the role of arousal and attentional processes.

### 1.3.3 The mental time line

Beside basic *temporal processing* and *time perception* in the sense of *duration* and *passage,* an intriguing aspect of *subjective time* is that past, present, and future are spatially represented along a *mental time line* (Bonato, Zorzi, & Umilta, 2012; Stocker, Hartmann, Martarelli, & Mast, 2015). The strong intuition that past-related concepts are located to the left while future-related concepts are located to the right has been confirmed by several empirical studies from the past decade. For example, it has been shown (in cultures that read from left to right) that the left hand reacts faster to the past and the right hand faster to the future (Ouellet, Santiago, Funes, & Lupianez, 2010; Santiago, Lupianez, Perez, & Funes, 2007). Moreover, there is evidence that (Western) observers look more leftward/downward when processing the personal (episodic) past and more rightward/upward when processing the personal (episodic) future (Hartmann, Martarelli, Mast, & Stocker, 2014). These results suggest that time and space are closely interrelated in the human mind, mirroring Einstein's astonishing discoveries about the fundamental interweaving of physical time and space in general relativity (e.g., Bennett, Donahue, Schneider, & Voit, 2010).

# 2. MAIN SECTION

## 2.1 Time perception in depression: A meta-analysis[1]

### 2.1.1 Abstract

*Background:* Depressive patients frequently report to perceive time as going by very slowly. Potential effects of depression on duration judgments have been investigated mostly by means of four different time perception tasks: *verbal time estimation*, *time production, time reproduction,* and *duration discrimination*. Ratings of the subjective *flow (passage) of time* have also been obtained. *Methods:* By means of a classical random-effects meta-regression model and a robust variance estimation model, this meta-analysis aims at evaluating the inconsistent results from 16 previous studies on time perception in depression, representing data of 433 depressive patients and 485 healthy control subjects. *Results:* Depressive patients perceive time as going by less quickly relative to control subjects ($g$ = 0.66, $p$ = .033). However, the analyses showed no significant effects of depression in the four duration judgment (timing) tasks. There was a trend towards inferior time discrimination performance in depression ($g$ = 0.38, $p$ = .079). The meta-regression also showed no significant effects of interval duration. Thus, the lack of effects of depression on timing does not depend on interval duration. However, for time production, there was a tendency towards overproduction of short and underproduction of long durations in depressive patients compared to healthy controls. Taken together, there is no evidence for sizable effects of depression on duration judgments in controlled experimental settings. *Limitations:* Several aspects, such as influences of medication and the dopaminergic neurotransmitter system on time perception in depression, have not been investigated in sufficient detail yet and were therefore not addressed by this meta-analysis. Conclusions: Depression has medium effects on the subjective flow of time whereas duration judgments basically remain unaffected.

---

[1] This paper (pages 11 to 55) has been published in Journal of Affective Disorders together with PD Dr. Daniel Oberfeld: Thönes, S., & Oberfeld, D. (2015). Time perception in depression: A meta-analysis. *Journal of Affective Disorders, 175*, 359-372.

### 2.1.2 Introduction

Depressive patients frequently report to perceive time as passing by extremely slowly (Blewett, 1992; Ratcliffe, 2012; Straus, 1947). However, the question of whether time perception in the sense of judgments of defined time intervals is also affected by depression remains unresolved. We are faced with a large body of inconclusive and often contradictory findings. The present meta-analysis evaluates the existing literature on time perception in depression.

Over the last few decades, the potential effects of depression on time perception have been investigated empirically mostly by means of four different experimental tasks (see Msetfi, Murphy, and Kornbrot (2012) for a recent review). These tasks are (a) *verbal time estimation* (sometimes referred to as *'time estimation'*), where a time interval is presented, defined for instance by the inter-onset interval between two brief tones, and the subject gives an estimate in conventional time units like seconds (e.g., Bech, 1975; Bschor et al., 2004; Dilling & Rabin, 1967; Kitamura & Kumar, 1983), (b) *time production,* where the experimenter specifies a time interval in temporal units, and the subject produces this interval for example by pressing a button to mark the interval's beginning and end (e.g., Münzel, Gendner, Steinberg, & Raith, 1988; Tysk, 1984), (c) *time reproduction,* where a time interval is presented as in (a) and the subject produces a corresponding interval as in (b) (Mahlberg, Kienast, Bschor, & Adli, 2008; Mundt, Richter, van Hees, & Stumpf, 1998), and (d) *duration discrimination*, where typically two time intervals of almost equal length are presented successively, and the subject selects the longer interval (Msetfi et al., 2012; Rammsayer, 1990; Sevigny, Everett, & Grondin, 2003). For tasks (a) to (c), most studies focused on the mean duration of the time estimates, or on deviations of the estimates from the veridical values. Thus, in terms of Fechner (1860), the studies compared the "constant error" between depressive patients and controls. For duration discrimination (task (d)), performance is often characterized in terms of the duration difference limen, defined as for example the difference in duration between the two presented time intervals that results in 75% correct responses. It should be noted that for tasks (a) to (c) a corresponding measure of sensitivity is provided by the standard deviation of the estimates or productions across several trials (Treisman, 1963). This corresponds to the "variable error" in terms of Fechner (1860). However, only few studies (e.g., Oberfeld, Thönes, Palayoor, & Hecht, 2014) analyzed the variable error, and for this reason we restricted our meta-analysis to the mean duration of the time estimates (constant error) for tasks (a) to (c). Several studies additionally asked for ratings of the subjects' experience of the flow of time (task (e)), often by means of visual analogue scales

(VAS; e.g., Bschor et al., 2004; Mundt et al., 1998; Oberfeld et al., 2014) or questionnaires (Bech, 1975; Münzel et al., 1988). On visual analogue scales, the subjects are asked to mark a point on a line where the endpoints represent a very slow and very fast subjective flow of time. Notably, these ratings differ from tasks (a) to (d) because the subjective flow of time is assessed rather than the perception or production of defined time intervals.

Occasionally, effects of depression on other than the five tasks listed above have been studied. For instance, Bolbecker et al. (2011) measured the timing abilities of depressive patients by means of a paced finger tapping task, and Oberfeld et al. (2014) studied time-to-contact estimates for approaching visual objects (cf. Regan & Gray, 2000). However, these additional tasks have not been investigated in more than two primary studies each, and were therefore not included in our meta-analysis.

In order to predict in which way depression might influence the performance on the experimental tasks, it seems sensible to consider the influential cognitive *pacemaker-accumulator models* of interval timing (Gibbon et al., 1984; Treisman, 1963). These models assume an *internal clock* consisting of a pacemaker emitting pulses and an accumulator (or counter) collecting these pulses. In *Scalar Expectancy Theory (SET)*, which is one of the most prominent pacemaker-accumulator models (Gibbon et al., 1984; Meck, 1996), this clock device is integrated into an information processing framework that encompasses memory and decision stages. According to SET, as soon as a subject begins to process an interval, an attentionally modulated switch between pacemaker and accumulator closes. Therefore, the clock pulses emitted by the pacemaker can reach the accumulator, which starts to 'count' these pulses. The more pulses being accumulated, the longer the perceived length of an interval. This means that if the subject's clock runs faster, more pulses get accumulated within a specified interval, and therefore the interval is perceived as longer compared to a subject with a slower clock speed.

In terms of this model, the observation that depressive patients frequently report to perceive time as going by less quickly can be explained by a faster running clock in depressive patients than in non-depressive controls. This assumption leads to precise predictions of performance differences between depressives and healthy control subjects in some of the interval timing tasks introduced above (Msetfi et al., 2012). For example, if the *verbal estimation* of a presented time interval in time units like seconds or minutes is required, according to the notion of an accelerated internal clock, the depressives accumulate more pulses during the presentation of the to-be-judged time interval, and hence produce higher estimates of the duration of the interval compared to control subjects. The opposite relation is predicted for a *production* task where the task is to produce an interval specified in time units, for example by marking its beginning and end by finger taps. If the internal clock runs at a faster pace, then the depressive patients should produce shorter intervals than the control subjects. According to the internal clock model, the subject starts to accumulate clock pulses at the first tap, and produces the second tap as soon as the accumulated number of pulses reaches a value (stored in long term memory) corresponding to for example "2 s". Due to the faster-running clock, the depressive patients should decide to mark the end of the interval at an earlier point in time than the control subjects. In a *reproduction* task, subjects are required to reproduce a previously presented time interval, for example by pressing a button to mark the interval's beginning and end. In contrast to *production* tasks, the interval is not specified in terms of time units but it is presented explicitly before the subject is asked to reproduce it. Here, a faster accumulation of pulses should affect the representation of the interval to be timed as well as its reproduction. According to SET, the memory representation of the pulses accumulated during the presentation of the time interval, and the accumulation process during the production phase should be affected in the same way. Therefore, the clock speed should have no effect in a reproduction task, and therefore no differences between depressives and controls are to be expected. *Duration discrimination* tasks require the detection of small differences between two successively presented time intervals (two-interval task), or the comparison of a presented time interval with internal (memory) references, as for example in a temporal generalization task (cf. Grondin, 2010). At first glance, higher clock speed might provide higher temporal resolution capacity. If the clock period is not sufficiently smaller than the temporal interval that is to be judged (e.g., clock period of 1 s, interval duration 500 ms), then the performance in a duration discrimination task would indeed be impaired. Apart from this obvious relation, however, the effect of changes in the clock rate critically depends on the relation between the clock rate and the variance of the pulse counts. For example, it has been suggested that the internal clock is a Poisson process (Gibbon, 1992; Grondin, 2010), for which the mean count of pulses emitted during a given time interval is proportional to the variance of the pulse count. Thus, increasing the clock rate by, e.g., 50% would also increase the variability of the pulse count by 50%. For a duration discrimination task, the signal-detection theory measure of sensitivity, $d'$, can be defined as the expected difference between the pulse counts during the longer and the shorter temporal interval, divided by the standard deviation of the pulse count (Green & Swets, 1966). For a Poisson-process clock, $d'$ will increase with the square root of the clock rate. In other words, the discrimination performance should be superior with a higher clock rate. Thus, depressive patients should show higher sensitivity than controls.

If one considers the literature, for each of the different tasks, there is evidence for effects of depression on time perception, but also against such effects. Most studies did not focus on one task only but included a variety of tasks and interval durations (e.g., milliseconds, seconds, minutes). Verbal time estimation has been studied frequently. However, the results are not conclusive for this task: While some studies provide clear evidence for a systematic overestimation in depressive patients compared to healthy control subjects at several interval durations (e.g., Kitamura & Kumar, 1983; Kornbrot, Msetfi, & Grimwood, 2013; Wyrick & Wyrick, 1977), compatible with an increased clock speed in depression, others reported mixed results (Biermann et al., 2011; Bschor et al., 2004) or even opposite effects (underestimation in depressives relative to controls) at certain interval durations (Tysk, 1984). Similarly, the effects in production tasks are discussed controversially. Beside statistically significant results in favor of the 'faster clock in depressives-assumption' (underproduction in depressives) (Bschor et al., 2004; Mundt et al., 1998), several studies did not find differences in the produced interval lengths between depressives and controls (Kitamura & Kumar, 1983; Tysk, 1984). Moreover, Münzel et al. (1988) and Kornbrot et al. (2013) reported overproduction in a depressive group. In the case of interval discrimination, the results are also mixed (Gil & Droit-Volet, 2009; Msetfi et al., 2012; Sevigny et al., 2003), with tendencies towards lower discrimination thresholds, i.e., higher temporal sensitivity in control subjects (Rammsayer, 1990). With regard to time reproduction (Mahlberg et al., 2008; Mundt et al., 1998; Oberfeld et al., 2014) and time experience (passage) (Bech, 1975; Kitamura & Kumar, 1982; Münzel et al., 1988; Oberfeld et al., 2014; Wyrick & Wyrick, 1977), similar inconsistent results can be found.

Taken together, when qualitatively reviewing the existing literature on the topic, one has to conclude that the empirical evidence for or against effects of depression in the different interval timing tasks is mixed. It is therefore difficult to answer the question of whether the would-be common clinical phenomenon of a subjective slowing of the flow of time in depressives is accompanied by systematic effects on time perception in terms of duration judgments (interval timing). One of the potential origins of the indecisive body of literature is the limited statistical power of each single study. Owing to the usual problems in recruiting large numbers of patients, each study investigated only a rather small sample of depressive patients. The meta-analytic approach used in the present study allows overcoming this limitation by combining information from multiple studies.

Apart from issues of statistical power, the inconsistency of the empirical findings could also be due to methodological heterogeneity of the studies, owing for example to the use of different tasks (e.g., verbal estimation versus production) and/or different time intervals (e.g., milliseconds versus minutes). For example, the demands in terms of motor activity clearly differ between the tasks. While production and reproduction tasks require timed motor responses, time estimation, duration discrimination, and judgment of the flow of time (time experience) do not. Also, it has to be assumed that memory processes are involved differently depending on the particular task. Time production and time estimation require the subject to refer to (long term) memory representations of the intervals to be timed. In time reproduction and discrimination tasks, however, the information necessary for doing the task is presented within a given trial, or within the experimental block, so that these tasks are likely to depend on short-term memory or sensory memory rather than on long-term memory. Moreover, it is important to consider that production and estimation, for example, produce opposite effects if the internal clock is accelerated. Therefore, deviations of the estimates from the veridical values on the two different tasks should not be pooled directly. Another crucial aspect might be the interval duration. There is evidence that different timing mechanisms are involved depending on the length of the interval to be timed (Bangert et al., 2011; Grondin, 2012; Lewis & Miall, 2003). Processing of intervals in the range of milliseconds is assumed to be based mainly on sensory mechanisms while cognitive factors like attention and memory become more important in the interval range above one second (e.g., Grondin, 2010). Hence, possible effects of depression on interval timing might depend on the time intervals used.

To answer the question of whether the differences in time perception between depressive patients and controls depend on the task and on the interval duration, which might explain some portion of the seemingly contradictory findings, we analyzed the existing studies according to the experimental task and to the interval durations. We defined four interval duration ranges: ultra short: < 1 s, short: 1 s to 10 s, medium: 10 s to 10 min, and long > 10 min.

### 2.1.3 Method

#### *2.1.3.1 Search strategies and study selection*

We searched for relevant studies in Web of Science. The primary key words were ‘depression’ and ‘timing’ or ‘time’. Additional studies were identified by including the references listed in the studies found in Web of Science, and by considering the studies that cited the resulting body of literature, again using Web of Science. Moreover, based on an email list from the "International Conference on Timing and Time Perception" (Corfu, 2014), we sent calls for unpublished data on the topic to more than 100 researchers in the field of timing and time perception. Additionally, we contacted the authors of previous papers on depression and time perception as far as current email addresses were available. This iterative literature search strategy yielded 39 articles addressing time perception in depression, with 31 papers reporting empirical data (see Appendix A.1).

For the meta-analyses, we selected studies according to the following four criteria.

Criterion 1). The studies had to provide data from a group of depressive patients as well as a control group consisting of healthy adults only. Because healthy subjects also tend to produce systematic errors in interval timing tasks (e.g., Wearden & Lejeune, 2008), it is uninformative to simply compare judgments of depressive subjects, for example verbal time estimates, to the veridical values of the presented time intervals. Hence, for studies that focused on depressive subjects only (e.g., Mezey & Cohen, 1961), it is not possible to decide whether the reported deviations of the time estimates from the veridical values are specifically related to depression. Therefore, systematic comparisons between depressive subjects and healthy control subjects on the single study level are required.

Criterion 2). Depressives had to be either diagnosed by means of standard diagnostic criteria (DSM or ICD), or had to be assigned to a depressive group based on a depression inventory like Hamilton's Rating Scale for Depression (HRS) (Hamilton, 1960) or the Beck Depression Inventory (BDI) (Beck, Erbaugh, Ward, Mock, & Mendelsohn, 1961).

Criterion 3). The report of sample sizes, means, and standard deviations or $t/F$ values of the response measures had to be sufficiently detailed in order to compute effect sizes (Hedges' $g$) and their variance. If this was not the case, we contacted the authors of the original study for papers published after 1990. Only one author (Kornbrot et al., 2013) kindly provided additional statistic necessary for our meta-analysis.

Criterion 4). At least one of the five common tasks (a) to (e) listed above had to be used.

Only 15 of the 31 empirical studies met these four criteria and were considered for further analyses (Bschor et al., 2004; Gil & Droit-Volet, 2009; Kitamura & Kumar, 1982, 1983; Kornbrot et al., 2013; Mahlberg et al., 2008; Mioni, Stablum, & Grondin, in prep.; Msetfi et al., 2012; Mundt et al., 1998; Münzel et al., 1988; Oberfeld et al., 2014; Rammsayer, 1990; Sevigny et al., 2003; Tysk, 1984; Wyrick & Wyrick, 1977) (see also Appendix A.1). Data from Kitamura and Kumar (1982) and Kitamura and Kumar (1983) are based on the same sample. However, the two studies focused on different tasks (1982: time experience; 1983: time estimation and production) and for practical reasons they will be listed as two separate studies. Msetfi et al. (2012) reported two experiments investigating independent samples of subjects. Therefore, in the analyses, their first and second experiment were treated as two separate studies, resulting in a total of 16 independent studies entering the analyses.

#### *2.1.3.2 Description of studies*

The 16 studies included in this meta-analysis provided data from a total of 918 subjects (433 depressive patients and 485 healthy controls). The median publication year was 2003 (range: 1977 to 2014). Studies were conducted in Great Britain (*n* = 5), Germany (*n* = 6), France (*n* = 1), Italy (*n* = 1), Sweden (*n* = 1), USA (*n* = 1), and Canada (*n* = 1).

Most patients (52.12%) were diagnosed according to DMS-III, DSM-IV, or ICD-9 criteria. 36.03% of the subjects in the patient group were assigned based on a BDI score. A minority of depressive subjects (17.55%) was diagnosed / assigned based on a “Present State Examination” and a questionnaire on “Depressive Mood” (Kitamura & Kumar, 1982, 1983), or according to the Multiple Affect Adjective Check List (MAACL) (Wyrick & Wyrick, 1977). Some patients were under medication, others not. The exact number of medicated subjects was not reported in most of the primary studies, and only one study reported separate data for subjects on and off medication (Oberfeld et al., 2014).

Supplementary Table 1 provides an overview of study-specific criteria for group assignment, diagnostics, and potential (drug) treatment of patients. Also, data on age and gender are presented as far as reported by the primary studies.

#### *2.1.3.3 Preprocessing*

Some studies reported separate results for patient groups suffering from different subtypes of depression, for example major depression with melancholia, major depression without melancholia, neurotic depression, etc. (e.g., Kitamura & Kumar, 1982; Tysk, 1984). Due to the inhomogeneity of diagnostic criteria (DSM-III, DSM-IV, ICD-9, BDI score, etc.), the small number of studies focusing on particular subtypes of depression, and the aim of the present meta-analysis to answer the question whether depression as such influences time perception, we decided to aggregate data from the different patient groups within those studies. Therefore, if multiple patient groups had been tested, for each response measure we computed weighted averages of the means and standard deviations with weights proportional to the number of subjects in the respective subgroup.

To ensure the comparability of effect size measures from the different studies, further preprocessing was required. With regard to judgments of time experience, the questionnaire scale for assessing the flow of time in Münzel et al. (1988) (1 indicating fast and 5 indicating slow) had to be inverted in order to be comparable to the scale used by Kitamura and Kumar (1982) and Wyrick and Wyrick (1977) (1 indicating slow and 5 indicating fast). The VAS measure used by Mundt et al. (1998) was also inverted in order to be comparable to the measures used by Bschor et al. (2004) and Oberfeld et al. (2014). Hence, for all judgments of time experience, higher values indicated a quicker flow of time. In the case of discrimination tasks, Rammsayer (1990) and Msetfi et al. (2012) reported the mean duration discrimination threshold with smaller values indicating higher sensitivity while Sevigny et al. (2003) analyzed the percentage of correct responses with larger values indicating higher sensitivity. For Sevigny et al. (2003), we multiplied the effect size measure by −1 in order to analyze comparable measures for all discrimination tasks.

#### *2.1.3.4 Effect size estimates*

Based on the reported means, standard deviations, and sample sizes for depressive groups ($M_d$, $SD_d$, $n_d$) and control groups ($M_c$, $SD_c$, $n_c$), we calculated Hedges' *g* as an effect size index, which is an estimate of the standardized mean difference between the two groups. According to Hedges and Olkin (1985), *g* is defined as

$$g = \frac{M_d - M_c}{s}, \tag{1}$$

where *s* is the pooled sample standard deviation,

$$s = \sqrt{\frac{(n_d - 1)SD_d^2 + (n_c - 1)SD_c^2}{n_d + n_c - 2}}. \tag{2}$$

According to Hofmann, Sawyer, Witt, and Oh (2010), the magnitude of *g* may be interpreted based on the conventions for the common effect size estimator *d* (small: ≥ 0.2; medium: ≥ 0.5; large: ≥ 0.8) (Cohen, 1988).

For Wyrick and Wyrick (1977), because means and standard deviations were not reported, *g* was calculated based on the presented *F* values and sample sizes (Rosnow & Rosenthal, 1996),

$$g = \sqrt{F \frac{n_d + n_c}{n_d n_c} \frac{n_d + n_c}{n_d + n_c - 2}}. \tag{3}$$

Following the recommendation by Normand (1999), we did not consider the bias correction factor *J(m)* proposed by Hedges and Olkin (1985), which yields somewhat smaller effect size estimates.

The asymptotic variance of *g* is (Hedges & Olkin, 1985)

$$Var(g) = \frac{n_d + n_c}{n_d n_c} + \frac{g^2}{2(n_d + n_c)}. \tag{4}$$

Following the preprocessing explained above, one value of *g* (and *Var (g)*) was computed for each pair of means reported in the selected studies, that is, for each combination of sample, task, and interval duration (see Supplementary Table 2).

For some of the tasks, the *g* values were multiplied by −1 so that positive values of *g* reported in our analyses indicate overestimation in time estimation tasks (a), underproduction in time production tasks (b), over-reproduction in temporal reproduction tasks (c), lower discrimination performance in duration discrimination tasks (d), and a reduced speed of the flow of time in time experience tasks (e) in depressive subjects relative to control subjects. Thus, for all tasks, positive effect sizes represent the effects of depression that are typically expected in the literature.

It should be noted that for tasks (a) to (c), different response measures were reported by the primary studies. Performance was analyzed in terms of the mean time estimates, the deviation of the estimates from the veridical values (constant error), the constant error divided by the veridical value (relative error), or the ratio between the estimated value and the veridical value. As all of these measures are linear transformations of each other, the effect size estimates according to Eq. (1) and (3) are not affected by the choice of the response measure in the primary studies.

### 2.1.4 Results

As pointed out above, the effects of depression on the different time perception tasks (time estimation, production, reproduction, discrimination and experience) are not necessarily identical. For instance, the internal clock model predicts opposite effects of an accelerated or decelerated internal clock in a verbal estimation and a production task, and predicts no effects of depression in a time reproduction task. Therefore, we decided to differentiate between data provided by the five different tasks (time estimation vs. production vs. reproduction vs. discrimination vs. experience). In addition, we considered different interval duration ranges (ultra short: < 1 s, short: 1 s to 10 s, medium: 10 s to 10 min, long: > 10 min) as a possible source of variance in the interval timing tasks (not including time experience, for reasons explained above).

#### *2.1.4.1 Pooled effect size estimates per task*

The aim of the first step of analysis was to determine one pooled effect size estimator for each of the five tasks. We used two different meta-analytic approaches.

In the first approach, possible effects of depression on task performance were analyzed by fitting a classical random effects meta-regression model (van Houwelingen, Arends, & Stijnen, 2002). On the level of single studies, in cases where several effect sizes were available for the same sample, task and interval range, for example because different intervals within the same interval range had been studied, we first computed the arithmetic mean of the *g*-values, and then computed the asymptotic variance for the mean value according to Eq. (1) or Eq. (3) and Eq. (4), respectively. The column labeled *J* in Appendix A.2 specifies the number of pairs of means for depressives versus controls reported for the same sample, task and interval range, each corresponding to one effect size (*g*). If the sample sizes reported within a combination of sample, task and range differed between for example time intervals, owing most likely to dropout (e.g., Kitamura & Kumar, 1983; time estimation), we averaged the sample sizes when computing *Var* (*g*). These analysis steps provided one value of *g* and its variance for each combination of sample, task and interval range (see Appendix A.2).

The interval range was entered as an effect-coded rather than continuous covariate because we are not aware of any argument that the effects of depression should be linearly related to the interval duration. This corresponds to the analyses in the original studies where the interval duration was analyzed as a within- or between-subjects factor in ANOVAs (e.g., Kornbrot et al., 2013; Oberfeld et al., 2014).

Using the SAS PROC MIXED procedure (Littell, Milliken, Stroup, Wolfinger, & Schabenberger, 2006), for each task, the meta-regression model provided an intercept value that represents an effect size estimate of depression. It also provided a Type 3 test for the effect of interval duration. The degrees of freedom were calculated according to Kenward and Roger (1997). The PROC MIXED script used in our analysis is provided in Appendix B.1

As an alternative approach, we considered a recent robust variance estimation approach that provides an advanced meta-regression method to handle dependent effect sizes (Hedges, Tipton, & Johnson, 2010). In our meta-analysis, most primary studies reported several effect sizes for the same task, for example because different time intervals had been studied. These effect sizes cannot be assumed to be independent because they had been obtained from the same sample of subjects. However, the classical meta-regression model explained above does not explicitly account for these potential dependencies. To identify potential resulting biases in the pooled effect size estimates, we additionally calculated effect size estimates based on the robust variance estimation approach (Hedges et al., 2010) using the SPSS ROBUST macro by Tanner-Smith and Tipton (2014) and compared these results to those from the classical meta-regression model (van Houwelingen et al., 2002). The analyses using the robust variance estimation approach were based on the data as reported in Supplementary Table 2. Thus, the effect sizes for each combination of sample, task, and interval duration as reported by the primary studies were entered in the analysis, without first computing an average effect size per interval range as for the classical random-effects meta-regression model. The interval range was included as an effect-coded covariate, just as in the classical meta-regression model. The method by Hedges et al. (2010) requires the specification of the (common) correlation ρ between the effect sizes in the different conditions. As no evidence concerning the to be expected correlation was available, we set ρ to 0.8, following Hedges et al. (2010). All analyses were repeated with ρ set to 0.0, but this resulted in only very minor changes in the estimated effects sizes, variances, and *p*-values.

For each of the five tasks, Table 1 presents the pooled effect size estimates and *t*-statistics based on a) the classical meta-regression model and b) the robust variance estimation approach.

Yielding similar or even identical (time experience) results, both regression models did not indicate statistically significant effects of depression on task performance in any of the interval timing tasks. However, a significant effect of medium size in time experience tasks indicated a reduced speed of the passage of time in depressive subjects. In addition, the random-effects meta-regression indicated a marginally significant, small detrimental effect of depression on discrimination performance, but this effect was clearly non-significant in the robust variance estimation analysis.

**Table 1**

Pooled effect size estimates per task. The table shows results from the random-effects meta-regression model and the robust variance estimation approach.

| Task | *N* ind. samples | Classical random-effects meta-regression | | | | | | | | | | Robust variance estimation meta-analysis | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *N* effect sizes | θ | $CI_L$ | $CI_U$ | *t* | *df* | *p* | $\tau^2$ | $SE_{\tau^2}$ | $p_{\tau^2}$ | *N* effect sizes | θ | $CI_L$ | $CI_U$ | *t* | *df* | *p* |
| estimation | 7 | 19 | 0.16 | –0.22 | 0.54 | 0.93 | 13.7 | .370 | 0.24 | 0.12 | .028 | 53 | 0.15 | –0.41 | 0.71 | 0.86 | 3 | .454 |
| production | 8 | 16 | 0.04 | –0.28 | 0.37 | 0.29 | 11.8 | .778 | 0.10 | 0.08 | .218 | 28 | 0.03 | –0.48 | 0.54 | 0.87 | 4 | .866 |
| reproduction | 4 | 8 | 0.14 | –0.30 | 0.57 | 0.81 | 5 | .455 | 0.12 | 0.13 | .184 | 13 | 0.15 | –3.44 | 3.74 | 0.52 | 1 | .695 |
| discrimination | 5 | 8 | 0.38 | –0.06 | 0.82 | 2.11 | 6 | .079 | 0.15 | 0.12 | .107 | 9 | 0.41 | –0.19 | 1.02 | 2.18 | 3 | .117 |
| experience | 6 | 6 | 0.66* | 0.08 | 1.24 | 2.93 | 5 | .033 | 0.22 | 0.19 | .261 | 13 | 0.66* | 0.08 | 1.23 | 2.29 | 5 | .033 |

*Note.* "*N* ind. samples": number of single studies/independent samples (i.e., the sample size on level 1). "*N* effect sizes": number of single effect sizes that were entered into the corresponding model (i.e., the sample size on level 2). θ: pooled effect size estimate. $CI_L$ and $CI_U$ are the lower and upper bounds of the 95% confidence interval, respectively, and *t*, *df*, and *p* refer to a test of θ against 0. $\tau^2$: estimate of the inter-study variance. $SE_{\tau^2}$ : standard error of $\tau^2$. $p_{\tau^2}$ refers to a test of $\tau^2$ against 0. * indicates statistically significant effects ($p < .05$).

#### *2.1.4.2 Meta-regression analysis on the effect of interval range for each task*

The aim of the second step of the analysis was to determine for each interval timing task (estimation, production, reproduction, and discrimination) whether the effect sizes differed between the time interval ranges (ultra short, short, medium, long). Time experience measures were not included in this step because, as pointed out in the Introduction section, the concept of interval duration does not apply to this type of task. It has to be noted that for some tasks data are available for a few interval ranges only (see Table 2). The tasks which were studied most frequently are verbal estimation and time production, in particular in the interval ranges short and medium.

The effect of interval range on the effect size (i.e., on the difference between depressives and controls) was analyzed by fitting the random effects meta-regression model introduced above (van Houwelingen et al., 2002) per task, again using the SAS PROC MIXED procedure, and again based on the data presented in Appendix A.2, that is, averaged effect sizes and variances per sample, task and interval range. Because SAS PROC MIXED computes Type 3 tests only for dummy-coded, but not for effect-coded covariates, a slightly different SAS syntax was used (see Appendix B.2). The meta-regression models also provided least-squares means as estimators of effect size per interval range and task. These estimates represent predicted population marginal means, based on the estimated fixed-effects parameters (Littell et al., 2006). Results are presented in Figure 1 and Table 3.

Note that although the robust variance estimation approach (Hedges et al., 2010) can be used for meta-regression, at present it is not possible to compute Type 3 tests for the effects of a categorical covariate within this procedure (Elizabeth Tipton, personal communication, November 2014).

**Table 2**

List of studies covering the different combinations of task and interval range.

| | ultra short (< 1s) | short (1s – 10s) | medium (10s – 10min) | long (> 10min) |
|---|---|---|---|---|
| **estimation** | *Oberfeld, et al. (2014)* | *Wyrick & Wyrick (1977)*<br>*Kitamura & Kumar (1983)*<br>*Tysk (1984)*<br>*Bschor, et al. (2004)*<br>*Kornbrot et al. (2013)*<br>*Oberfeld et al. (2014)* | *Wyrick & Wyrick (1977)*<br>*Kitamura & Kumar (1983)*<br>*Tysk (1984)*<br>*Münzel, et al. (1988)*<br>*Bschor, et al. (2004)*<br>*Kornbrot et al. (2013)*<br>*Oberfeld et al. (2014)* | *Kitamura & Kumar (1983)*<br>*Münzel, et al. (1988)*<br>*Bschor, et al. (2004)*<br>*Oberfeld, et al. (2014)*<br>*Wyrick & Wyrick (1977)* |
| **production** | *Oberfeld, et al. (2014)*<br>*Mioni et al (in prep)* | *Tysk (1984)*<br>*Münzel, et al. (1988)*<br>*Mioni et al (in prep)*<br>*Mundt, et al. (1998)*<br>*Bschor, et al. (2004)*<br>*Oberfeld, et al. (2014)*<br>*Kornbrot et al. (2013)* | *Tysk (1984)*<br>*Münzel, et al. (1988)*<br>*Mundt, et al. (1998)*<br>*Bschor, et al. (2004)*<br>*Oberfeld, et al. (2014)*<br>*Kornbrot et al. (2013)* | *Kitamura & Kumar (1983)* |
| **reproduction** | *Oberfeld, et al. (2014)*<br>*Mioni et al (in prep)* | *Mahlberg, et al. (2008)*<br>*Oberfeld, et al. (2014)*<br>*Mioni et al (in prep)* | *Mundt, et al. (1998)*<br>*Mahlberg, et al. (2008)*<br>*Oberfeld, et al. (2014)* | |
| **discrimination** | *Rammsayer (1990)*<br>*Sevigny, et al. (2003)*<br>*Gil & Droit-Volet (2009)*<br>*Msetfi, et al. (2012) exp.1*<br>*Msetfi, et al. (2012) exp.2* | *Sevigny, et al. (2003)*<br>*Msetfi, et al. (2012) exp.1*<br>*Msetfi, et al. (2012) exp.2* | | |

**Table 3**

Estimated effect sizes ($\theta_{LSM}$) per combination of task and interval range, as provided by least-squares means computed in the classical random effects meta-regression model.

| Task | Interval range | *N* studies | $\theta_{LSM}$ | $CI_L$ | $CI_U$ | *t* | *df* | *p* |
|---|---|---|---|---|---|---|---|---|
| **estimation** | | | | | | | | |
| | *ultra short* | *1* | *–0.12* | *–1.35* | *1.11* | *0.21* | *14.0* | *.834* |
| | short | 6 | 0.24 | –0.29 | 0.77 | 0.98 | 13.3 | .344 |
| | medium | 7 | 0.04 | –0.42 | 0.50 | 0.87 | 12.8 | .871 |
| | long | 5 | 0.50 | –0.05 | 1.05 | 1.97 | 13.6 | .070† |
| **production** | | | | | | | | |
| | ultra short | 2 | –0.45 | –1.18 | 0.27 | 1.36 | 12.0 | .200 |
| | short | 7 | –0.24 | –0.60 | 0.12 | 1.47 | 10.2 | .173 |
| | medium | 6 | 0.29 | –0.10 | 0.68 | 1.69 | 9.77 | .123 |
| | *long* | *1* | *0.57* | *–0.39* | *1.53* | *1.31* | *11.2* | *.217* |
| **reproduction** | | | | | | | | |
| | ultra short | 2 | 0.15 | –0.72 | 1.03 | 0.45 | 5 | .669 |
| | short | 3 | 0.50 | –0.20 | 1.19 | 1.83 | 5 | .127 |
| | medium | 3 | –0.24 | –0.91 | 0.43 | 0.93 | 4.8 | .398 |
| **discrimination** | | | | | | | | |
| | ultra short | 5 | 0.15 | –0.38 | 0.67 | 0.68 | 6 | .521 |
| | short | 3 | 0.61 | –0.09 | 1.32 | 2.12 | 6 | .078† |

*Note*. † indicates statistically marginal significant results ($p < .1$). Categories including data from one single study only are indicated by italics.

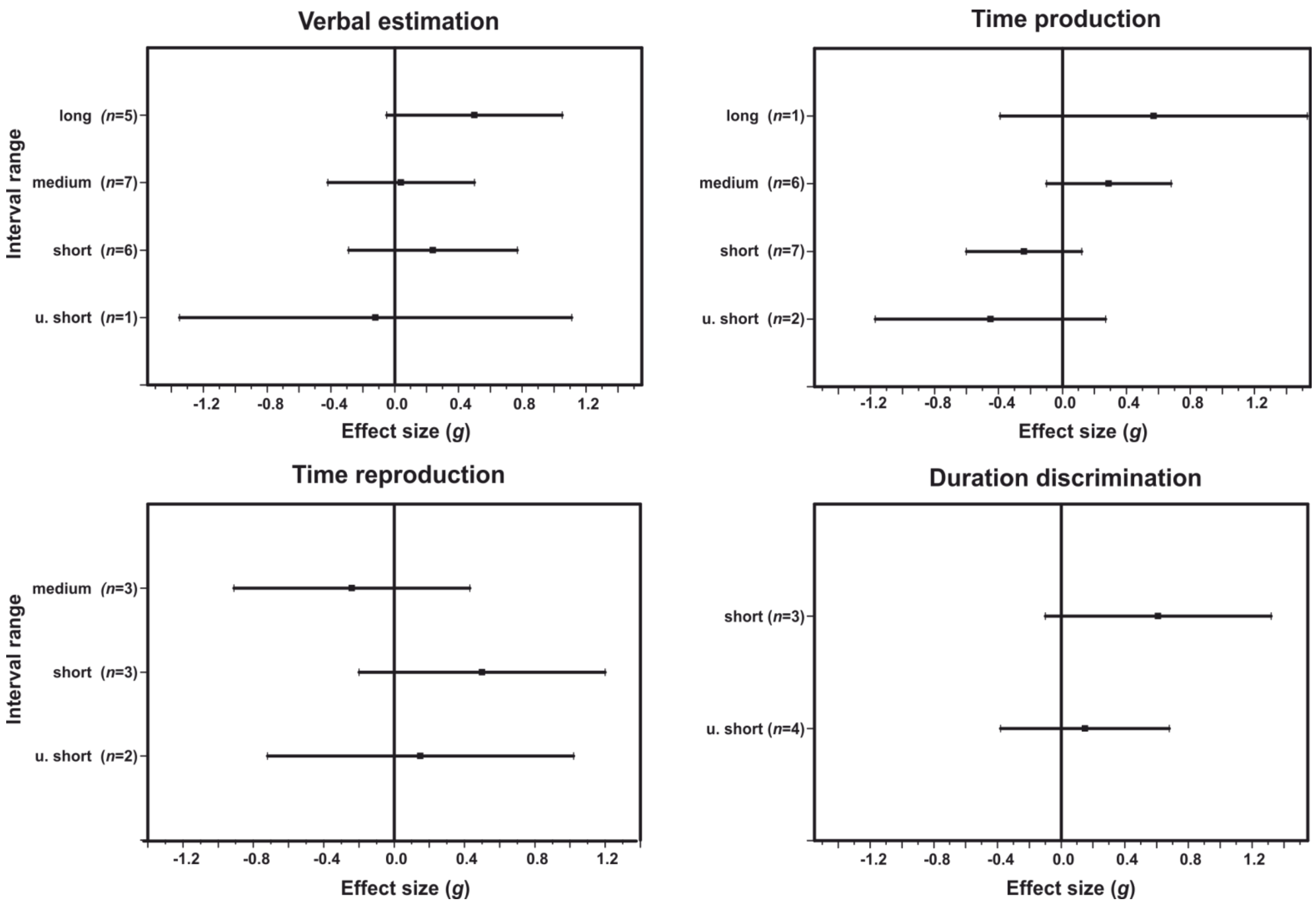


**Figure 1**

Pooled effect sizes per combination of task and time interval range. Effect size estimates and corresponding 95% confidence intervals as provided by least-squares means (marginal means) in the meta-regression analysis. Each data point represents one interval range, with the number of studies indicated by the values in parentheses. Each panel represents one task.

For no task, the Type 3 test indicated a significant influence of the covariate interval range on the effect of depression on task performance (time estimation: $F(3,13.5) = 0.78$, $p = .525$; time production: $F(3,11.4) = 2.85$, $p = .085$; temporal reproduction: $F(2,5) = 2.06$, $p = .223$; duration discrimination: $F(1,6) = 1.68$, $p = .243$). However, for time production, there was a tendency towards overproduction of short and underproduction of long durations in depressive patients.

The least-squares means (Table 3) showed marginally significant effects of medium size for verbal time estimation at long intervals (depressives overestimated time intervals) and duration discrimination at short intervals (depressives discriminated less precisely). For time production and time reproduction, no significant or marginally significant results were obtained for any interval range.

In summary, the meta-regression showed no strong effects of interval range. Thus, the lack of significant effects of depression in Step 1 cannot be attributed to effects of depression on time perception differing between interval durations ranging from the sub-second range to minutes.

#### *2.1.4.3 Outlier-corrected results*

In the third step, for each task, we used regression diagnostics to identify outlying data. Following the recommendations by Viechtbauer and Cheung (2010), we analyzed the externally studentized residuals (also called the studentized deleted residuals) and the DFFITS index proposed by Belsley, Kuh, and Welsch (1980) as a measure of the influence of an observation. Following Belsley et al. (1980), externally studentized residuals with an absolute value exceeding 1.96, or an absolute DFFITS value exceeding $2\sqrt{p/n}$, where $n$ is the number of effect sizes analyzed in the model, and $p$ is the number of levels of the covariate (interval range), were defined as outliers. For the time estimation task, the data in the long interval range provided by Oberfeld et al. (2014) ($g = -0.33$) and Wyrick and Wyrick (1977) ($g = 2.11$) were identified as outliers. Excluding these data points from the analyses yielded a pooled effect size estimate for the time estimation task of $g = 0.11$ ($p = .334$), which is slightly smaller than the value based on the analysis including the two outliers (Table 1). With the two outliers excluded, the effect of interval range remained statistically insignificant, $F(3,11.9) = 0.78$, $p = .530$, indicating no differences in comparison to the analysis that included the outliers. The least-squares mean in the long interval range dropped from $g = 0.50$ (cf. Table 3) to $g = 0.28$ ($p = .175$).

For the time production task, one effect size of 0.85 for the medium interval range reported by Bschor et al. (2004) was identified as an outlier. Its exclusion resulted in an even reduced pooled effect size estimate ($g$ = 0.017, $p$ = .91), and no significant effect of interval range, $F(3,9.55)$ = 2.34, $p$ = .14.

For time discrimination, one effect size of 0.95 for the ultra-short interval range reported by Rammsayer (1990) was identified as an outlier. Its exclusion resulted in a reduced pooled effect size estimate ($g$ = 0.18, $p$ = .29), and no significant effect of interval range, $F(1, 2.14)$ = 8.55, $p$ = .17. The outlier-corrected predicted marginal means ($\theta_{LSM}$) were −0.12 ($SE$ = 0.13; $p$ = .500) and 0.48 ($SE$ = .16; $p$ = .172) for the ultra short and short interval range, respectively.

For the time reproduction task and the subjective flow of time, no outliers were identified.

#### 2.1.5 Discussion

The objective of this meta-analysis was to evaluate whether time perception in terms of judgments of temporal intervals on the one hand and judgments of the subjective flow of time on the other hand is altered in depression. In a first step, pooled effect size estimates were computed for each of the five different types of tasks that have been used in the literature to investigate time perception in depressives: verbal time estimation, time production, time reproduction, duration discrimination, and time experience (subjective flow of time). In a second step of analysis, we investigated whether possible effects of depression on time perception depend on the interval durations (e.g., sub-second versus minute range) that had to be judged by the subjects. Using a meta-regression approach, we considered interval duration as a covariate, which might influence the effect of depression on time perception in the interval timing tasks (not time experience) and could explain some of the between study variance.

Our analyses provided no evidence for effects of depression on time interval judgments. Verbal estimation, production, reproduction, and discrimination of time intervals did not significantly differ between depressives and control subject. However, effects of depression on judgments of the flow of time were statistically significant and medium in size ($g$ = 0.66). Thus, the results confirm the notion of a reduced speed of the flow of time in depressive subjects and, at the same time, suggest that this impression is not tantamount to a change in the ability to judge the duration of time intervals.

The meta-regression using interval range as a covariate (Step 2) showed that differences in interval durations cannot explain the heterogeneity of the single study results or the lack of significant results in the time perception tasks in this meta-analysis. For none of the four tasks, interval range had a significant influence on potential effects of depression on time perception. Only in the production task, there was a tendency towards overproduction of short and underproduction of long durations in depressive patients compared to a healthy control group. And, in the cases of long range verbal estimation and short range duration discrimination, least-squares means indicated medium and marginally significant effects of depression. Here, depressive subjects tended to overestimate the time intervals, and to discriminate time intervals less precisely compared to control subjects. However, even these marginal effects would be further reduced when taking corrections for multiple testing and our exclusion of outliers into account.

Partly, the data on time perception in depression are in line with the predictions of the *pacemaker-accumulator models* (Gibbon et al., 1984; Treisman, 1963). Judging time as going by less quickly is compatible with a faster running internal clock in depressive subjects. As discussed above, the null effects of depression in reproduction and duration discrimination tasks are also predicted by the internal clock model. However, the lack of significant group differences in the context of verbal estimation and production tasks is not in favor of this assumption. Thus, to some extent, the inconsistencies in the empirical findings remain and even meta-analyzed data do not fully support the hypothesis of a faster running internal clock in depressive patients.

Moreover, it has to be noted that even the few statistically significant and marginally significant results have to be viewed cautiously due to several reasons. First, we used the uncorrected (slightly biased) effect size estimator *g*. According to Hedges and Olkin (1985), *g* is biased towards overestimation of effects especially when effects and sample sizes are small. At least the problem of small effect sizes has to be considered in our analysis, indicating that the effect of depression on time perception might be even smaller in the depressive population than in our aggregated sample. Second, the analyses included multiple testing because we computed pooled effect sizes separately for each of the five tasks. On an appropriately adjusted $\alpha$-level (Hochberg, 1988), no effect would have reached statistical significance. Third, publication bias may have caused an overestimation of effect size in our analysis. We did not attempt to provide quantitative estimates of publication bias or to correct for its estimated influence (cf. Sutton & Higgins, 2008) because given the small number of primary studies in our view there is insufficient information for doing these analyses. However, we sent calls for unpublished data on the topic to more than 100 researches in the field of timing and time perception and received only one additional set of data (Mioni et al., in prep.). For this reason, we do not assume a substantial amount of unpublished data on the topic.

Beside these limitations suggesting a potential overestimation of effect sizes in the analyses, the fact that some of the studies included depressive patients who were on medication or in psychotherapy might have led to an underestimation of effect sizes. The small number of studies that were included in the analysis, and the failure of most studies to provide detailed information concerning the number of subjects under medication, did not allow controlling for possible effects of medication and psychotherapy. Half of the studies (see Supplementary Table 1) reported that at least some of their patients were medicated or in psychotherapy. Needless to say, medication and psychotherapy are applied in order to reduce depressive symptoms. Therefore, possible effects of depression on time perception might also decline in subjects that are under medication or in psychotherapy. Hence, the inclusion of depressive patients that received some sort of therapy probably led to an underestimation of the pooled effect sizes reported in this meta-analysis. Unfortunately, it was not possible to include medication status as a covariate, because only our own study (Oberfeld et al., 2014) analyzed the effects of depression on time perception separately for patients on and off medication. In the latter study, there was no significant effect of medication status. In addition, for the studies that administered a depression inventory like the BDI (e.g., Gil & Droit-Volet, 2009; Kornbrot et al., 2013; Msetfi et al., 2012; Oberfeld et al., 2014; Sevigny et al., 2003), the BDI scores provide information about the severity of depression on the day of testing. In our own study (Oberfeld et al., 2014), there were no significant correlations between the BDI score and the different time perception measures. Thus, individual differences in the severity of depression, which might have been due to medication or psychotherapy, do not appear to explain the weak effects of depression on time perception.

Based on the foregoing arguments, we conclude that the pooled effect sizes reported in this meta-analysis are more likely to overestimate than to underestimate the effects of depression on time perception on the population level.

A general issue for this meta-analysis was the heterogeneity of diagnostic criteria (DSM, ICD, BDI score etc.) and their modifications over time (e.g., DSM-III vs. DSM-IV). This factor might have led to inconsistencies in group assignment between the different studies. For instance, some subjects that were assigned to the control group in one study might have been classified as depressives in another study that used different criteria for group assignment. In order to minimize this potential problem, future research should consider group assignments based on BDI scores or using the BDI score as a continuous variable. This would provide a consistent quantitative measure of the severity of depression, facilitating the comparison between studies.

With regard to future research, our review of the literature identified several additionally aspects that have not yet been investigated sufficiently and which therefore were not evaluated in our meta-analysis. For example, some of the mechanisms that are considered to be responsible for alterations in clock rate not only in depressive subjects are changes in neurotransmitter systems. In the context of interval timing in the seconds-to-minutes range, the internal clock appears to depend on the level of dopaminergic activity (Meck, 1996, 2005; Rammsayer, 1990). Evidence from animal as well as human research suggests that decreased dopaminergic activity is related to a deceleration of clock speed (underestimation of time intervals) while increased dopamine levels usually cause opposite effects (e.g., Cheng, Ali, & Meck, 2007; Jones & Jahanshahi, 2009; Rammsayer, 1993; Wiener, Lee, Lohoff, & Coslett, 2014). In depressive patients, abnormalities in neurotransmitter systems are observed frequently (Leonard, 2014; Werner & Covenas, 2010). In particular, dopamine levels are typically decreased in depression, but with large inter-individual variability (Ebert & Lammers, 1997; Kapur & Mann, 1992; Yadid & Friedman, 2008). Thus, effects of depression on time perception might be mediated by dopamine levels and therefore vary between different patients or subtypes of depression. It would be interesting to consider direct measures of dopamine level in future studies in order to shed more light on the role of alterations in neurotransmitter systems in time perception in clinical populations.

In this context, additional research on patients in manic and mixed states (bipolar disorders) is also needed. For example, in contrast to major depression, which is often associated with slowed thinking (e.g.,Marazziti, Consoli, Picchetti, Carlini, & Faravelli, 2010), patients in manic and mixed states typically report racing thoughts (sometimes referred to as tachypsychia) (Benazzi & Akiskal, 2003; Geller et al., 2002). These opposing symptoms might be related differently to alterations in clock speed. Bschor et al. (2004), for example, provide evidence for an accelerated passage of subjective time in manic patients compared to a control group and a major depressive group. In order to reduce noise in the data from depressive patients, future studies should distinguish between the different subgroups of depression according to recent classifications. Moreover, ratings on the subjective flow of time can be improved by additionally assessing boredom ratings. This might help (depressive) subjects in understanding questions on the passage of time correctly, and can prevent them from confusing the concepts of time passage and boredom.

Another interesting aspect would be to focus on possible influences of the modality in which the stimuli to be timed are presented. In the primary studies analyzed here, temporal intervals were presented either visually or acoustically. The question whether the processing of visual stimuli is affected more strongly by depression than the processing of acoustic stimuli or vice versa remains unaddressed so far. Future research might consider modality as a covariate of time perception in clinical populations. In healthy subjects, there is evidence for differences in the temporal processing of visual vs. acoustic stimuli. Compared to visually marked time intervals, intervals of the same duration presented in the auditory modality are perceived as being longer (Goldstone & Lhamon, 1974; Wearden, Edwards, Fakhri, & Percival, 1998). Moreover, sensitivity to time is higher when intervals are presented acoustically, i.e., temporal judgments are less variable than for visually presented time intervals (Grondin, 2010; Grondin & McAuley, 2009). Accordingly, effects of depression on time perception might be easier to detect when the tasks involve auditory rather than visual stimuli.

Additionally, for time estimation tasks, it is crucial to differentiate between prospective and retrospective judgments (cf. Grondin, 2010), which has not been done systematically in the context of depression and time perception. In the prospective paradigm, the subject is informed about the task to estimate a presented time interval. Here, the subject explicitly focuses attention on the temporal task. In the retrospective paradigm, the subject is uninformed about the time estimation task. For example, a participant is asked at the end of the experiment how much time had elapsed since the beginning of the experiment, without having been informed previously that such an estimate would be required. In this case, the subject did not focus attention on time. In contrast to prospective judgments, the task to give a retrospective estimate is less attention-related and more memory demanding (e.g., S. W. Brown, 2008; Grondin, 2010). There is evidence indicating systematical differences between prospective and retrospective estimation paradigms (for a meta-analytic review see Block & Zakay, 1997), with prospective judgments being longer and less variable than retrospective judgments. Regarding possible effects of depression on time estimation, prospective and retrospective judgments might be affected differently indicating whether attentional (prospective) processes or/and memory-related (retrospective) processes of temporal information processing are altered in depression. Due to the fact that only few studies could be included in this meta-analysis, further covariates as for example the estimation paradigm (prospective vs. retrospective) could not be considered in the analysis. Inspection of the studies, however, indicated differences between results from one study including retrospective judgments (Münzel et al., 1988) and another study including prospective judgments (Bschor et al., 2004), with a trend towards larger effects in the context of retrospective (memory-related) judgments (more overestimation in depressives). However, in three out of four studies that actually tested both types of judgments within the same sample (Kitamura & Kumar, 1983; Oberfeld et al., 2014; Tysk, 1984; Wyrick & Wyrick, 1977), no systematic differences between retrospective and prospective estimates are evident. Only Wyrick and Wyrick (1977) reported larger effects for retrospective than for prospective judgments.

Taken together, the results of our meta-analysis indicate that judgments of time intervals are not affected systematically by depression. However, the notion of a reduced speed of the flow of time in depression has been confirmed. This also emphasizes the importance of a clear distinction between judgments on the flow of time and estimates of precisely defined time intervals (for a discussion see Oberfeld et al., 2014). Our review also shows that several aspects have not yet been investigated in sufficient detail in the context of time perception in depression and might be addressed by future research. These aspects include the role of the dopaminergic neurotransmitter system, influences of different subtypes of depression, potential influences of stimulus modality, and specific task-related characteristics like prospective versus retrospective time estimation. The effect sizes provided by our meta-analyses may be used for selecting appropriate sample sizes in future experiments.

### 2.1.6 Acknowledgements

We are grateful to Heiko Hecht for helpful comments on an earlier version of this manuscript, to Giovanna Mioni for contributing yet unpublished data to our analysis (Mioni et al., in prep.), and to Diana Kornbrot for providing the data from one of her studies (Kornbrot et al., 2013). We thank Emily Tanner-Smith and Elizabeth Tipton for the helpful correspondence concerning the robust variance estimation method.

### 2.1.7 Appendices

#### *2.1.7.1 Appendix A.1*

Results of the literature search. Studies were considered for further analyses only if all of the four inclusion criteria (see Introduction) were met. Studies included in the meta-analysis are indicated by bold font.

| **Study** | **Empirical study?** | **Patient and control group?** | **Sufficient statistical information?** | **Comment** | **Relevant tasks?** | **Inclusion?** |
|---|---|---|---|---|---|---|
| *Straus (1947)* | no | | | | | no |
| *Dubois (1954)* | no | | | | | no |
| *Mezey and Cohen (1961)* | yes | no (dep. before and after recovery) | yes | | yes | no |
| *Bojanovsky (1965)* | yes | no (psychogenic vs. endo. dep.) | no | *M* and imprecise *p* only | yes | no |
| *Dilling and Rabin (1967)* | yes | yes | no | no *SD/SE* or *t/F* values | yes | no |
| *Lehmann (1967)* | no | | | | | no |
| *Bojanovsky and Tölle (1973)* | yes | no (neurotic vs. endo. dep.) | no | *M* and imprecise *p* only | yes | no |
| *Grinker, Glucksman, Hirsch, and Viseltear (1973)* | yes | yes | no | no numerical *M/SD,* only figure, no *t/F* values, imprecise *p* only | yes | no |
| *Bech (1975)* | yes | yes | no | *M* & range, no *SD* and imprecise *p* only | yes | no |
| ***Wyrick and Wyrick (1977)*** | yes | yes | yes | | yes | yes |
| ***Kitamura and Kumar (1982)*** | yes | yes | yes | | yes | yes |

| Study | Empirical study? | Patient and control group? | Sufficient statistical information? | Comment | Relevant tasks? | Inclusion? |
|---|---|---|---|---|---|---|
| ***Kitamura and Kumar (1983)*** | yes | yes | yes | | yes | yes |
| ***Tysk (1984)*** | yes | yes | yes | | yes | yes |
| *Richter and Benzenhofer (1985)* | yes | no (case study) | yes | | yes | no |
| *Tysk (1985)* | yes | no (dep. longitudinal) | no | *r* and imprecise *p* only | yes | no |
| *Hawkins, French, Crawford, and Enzle (1988)* | yes | no | yes | | yes | no |
| ***Münzel, et al. (1988)*** | yes | yes | yes | | yes | yes |
| *Kuhs, Hermann, Kammer, and Tolle (1989)* | yes | yes | no | *M* and imprecise *p* only | yes | no |
| ***Rammsayer (1990)*** | yes | yes | yes | | yes | yes |
| *Watt (1991)* | yes | no | yes | | yes | no |
| *Blewett (1992)* | yes | no (depressives only) | yes | | yes | no |
| *Nosachev (1992)* | yes | yes | no | measure of dispersion reported, but not specified whether *SE*, *SD* etc. | yes | no |
| ***Mundt, et al. (1998)*** | yes | yes | yes | | yes | yes |
| *Lemke, Koethe, and Schleidt (1999)* | yes | yes | yes | | no (movement analysis) | no |
| ***Sevigny, et al. (2003)*** | yes | yes | yes | | yes | yes |
| ***Bschor, et al. (2004)*** | yes | yes | yes | | yes | yes |
| *Meck (2005)* | no (review) | | | | | no |

| Study | Empirical study? | Patient and control group? | Sufficient statistical information? | Comment | Relevant tasks? | Inclusion? |
|---|---|---|---|---|---|---|
| *Grondin, Pouthas, Samson, and Roy (2006)* | no (review) | | | | | no |
| ***Mahlberg, et al. (2008)*** | yes | yes | yes | | yes | yes |
| ***Gil & Droit-Volet (2009)*** | yes | yes | yes | | yes | yes |
| *Biermann, et al. (2011)* | yes | no | yes | | yes | no |
| *Bolbecker et al. (2011)* | yes | yes | yes | | no (f. tapping) | no |
| *Gallagher (2012)* | no (review) | | | | | no |
| ***Msetfi, et al. (2012)*** | yes | yes | yes | | yes | yes |
| *Ratcliffe (2012)* | no (review) | | | | | no |
| *Droit-Volet (2013)* | no (review) | | | | | |
| ***Kornbrot, et al. (2013)*** | yes | yes | yes | Author provided the raw data | yes | yes |
| ***Oberfeld, et al. (2014)*** | yes | yes | yes | | yes | yes |
| ***Mioni et al. (in prep.)*** | yes | yes | yes | | yes | yes |

*Note.* dep. = depression/depressive patients; f. tapping = finger tapping task; endo. = endogenous; *SD* = standard deviation; *SE* = standard error.

### 2.1.7.2 Appendix A.2

Averaged effect sizes (*g*) and effect size variances (*Var g*), number of reported means (*J*), and corresponding 95% confidence intervals per combination of sample, task and interval range. *J* denotes the number of effects sizes on which the average *g* is based on.

| **Study** | **Task** | **Interval range** | ***J*** | ***g*** | ***Var(g)*** | **$CI_L$** | **$CI_U$** |
|---|---|---|---|---|---|---|---|
| *Wyrick & Wyrick (1977)* | est | short | 2 | 0.14 | 0.12 | –0.37 | 0.64 |
| | | medium | 4 | 0.42 | 0.24 | –0.09 | 0.93 |
| | | long | 2 | 2.11 | 0.03 | 1.48 | 2.75 |
| *Kitamura & Kumar (1983)* | est | short | 6 | 0.53 | 0.09 | –0.06 | 1.13 |
| | | medium | 12 | 0.53 | 0.09 | –0.06 | 1.13 |
| | | long | 6 | 0.20 | 0.09 | –0.39 | 0.79 |
| | pro | long | 3 | 0.57 | 0.09 | –0.02 | 1.17 |
| *Tysk (1984)* | est | short | 1 | 0.44 | 0.05 | –0.02 | 0.89 |
| | | medium | 3 | −0.03 | 0.05 | –0.48 | 0.42 |
| | pro | short | 1 | 0.27 | 0.05 | –0.72 | 0.19 |
| | | medium | 2 | −0.19 | 0.05 | –0.26 | 0.64 |
| *Münzel, et al. (1988)* | est | medium | 3 | 0.38 | 0.08 | –0.19 | 0.95 |
| | | long | 1 | 0.00 | 0.08 | –0.57 | 0.57 |
| | pro | short | 3 | –0.08 | 0.08 | –0.65 | 0.48 |
| | | medium | 1 | 0.37 | 0.08 | –0.20 | 0.94 |
| *Rammsayer (1990)* | dis | u short | 1 | 0.95 | 0.05 | 0.52 | 1.37 |
| *Mundt, et al. (1998)* | pro | short | 1 | –0.13 | 0.09 | –0.72 | 0.47 |
| | | medium | 3 | 0.74 | 0.10 | 0.13 | 1.35 |
| | rep | medium | 4 | −0.48 | 0.09 | –1.08 | 0.12 |

| **Study** | **Task** | **Interval range** | ***J*** | ***g*** | ***Var(g)*** | **$CI_L$** | **$CI_U$** |
|---|---|---|---|---|---|---|---|
| *Sevigny, et al. (2003)* | dis | u short | 2 | 0.25 | 0.12 | –0.42 | 0.92 |
| | | short | 1 | 0.75 | 0.12 | 0.06 | 1.44 |
| *Bschor, et al. (2004)* | est | short | 1 | 0.00 | 0.06 | -0.50 | 0.49 |
| | | medium | 2 | –0.28 | 0.06 | –0.77 | 0.22 |
| | | long | 1 | 0.59 | 0.07 | 0.09 | 1.10 |
| | pro | short | 1 | –0.39 | 0.06 | –0.89 | 0.11 |
| | | medium | 2 | 0.84 | 0.07 | 0.32 | 1.35 |
| *Mahlberg, et al. (2008)* | rep | short | 2 | 1.03 | 0.08 | 0.49 | 1.58 |
| | | medium | 1 | 0.09 | 0.07 | –0.43 | 0.60 |
| *Gil & Droit-Volet (2009)* | dis | u short | 1 | −0.17 | 0.08 | –0.71 | 0.37 |
| *Msetfi, et al. (2012) exp. 1* | dis | u short | 1 | −0.28 | 0.11 | –0.93 | 0.38 |
| | | short | 1 | 0.84 | 0.12 | 0.16 | 1.52 |
| *Msetfi, et al. (2012) exp. 2* | dis | u short | 1 | −0.14 | 0.05 | –0.59 | 0.30 |
| | | short | 1 | 0.34 | 0.05 | –0.10 | 0.79 |
| *Kornbrot, et al. (2013)* | est | short | 1 | –0.40 | 0.10 | –1.01 | 0.21 |
| | | medium | 4 | –0.68 | 0.10 | –1.30 | –0.06 |
| | pro | short | 1 | –0.71 | 0.10 | –1.33 | –0.08 |
| | | medium | 4 | –0.67 | 0.10 | –1.29 | –0.05 |

| **Study** | **Task** | **Interval range** | ***J*** | ***g*** | ***Var(g)*** | **$CI_L$** | **$CI_U$** |
|---|---|---|---|---|---|---|---|
| *Oberfeld, et al. (2014)* | est | u short | 1 | −0.12 | 0.09 | –0.71 | 0.47 |
| | | short | 1 | 0.27 | 0.09 | –0.32 | 0.87 |
| | | medium | 1 | −0.12 | 0.09 | –0.71 | 0.47 |
| | | long | 1 | –0.33 | 0.09 | –0.93 | 0.26 |
| | pro | u short | 1 | 0.01 | 0.09 | –0.58 | 0.60 |
| | | short | 1 | 0.20 | 0.09 | –0.39 | 0.80 |
| | | medium | 1 | 0.21 | 0.09 | –0.38 | 0.80 |
| | rep | u short | 1 | −0.10 | 0.09 | –0.69 | 0.50 |
| | | short | 1 | −0.09 | 0.09 | –0.68 | 0.51 |
| | | medium | 1 | −0.37 | 0.09 | –0.97 | 0.23 |
| *Mioni et al. (in prep)* | pro | u short | 1 | –1.09 | 0.16 | –1.88 | –0.30 |
| | | short | 2 | –0.33 | 0.14 | –1.08 | 0.41 |
| | rep | u short | 1 | 0.47 | 0.15 | –0.28 | 1.22 |
| | | short | 2 | 0.50 | 0.15 | –0.25 | 1.25 |

### *2.1.7.3 Appendix B.1*

SAS syntax for the classical random-effects meta-regression model used to compute the pooled effect size estimates per task.

This meta-regression model was applied separately for each task. In the example, the data for the time production task are analyzed. The data set `time_depression_TimeProduction` contains one effect size estimate (variable `g_i`) for each combination of study and interval range (variable `intervalRange`). Each of these combinations was treated as a separate study, indicated by a unique value of the variable `studyNo`. The interval range was analyzed as a categorical covariate, using effect-coding via the indicator variables `e1` to `e3`, so that the estimated intercept corresponds to the pooled effect size estimate (θ). SAS syntax based on van Houwelingen et al. (2002).

```
data time_depression_TimeProductionEC;
      set time_depression_TimeProduction;
      if intervalRange="large" then
            do;
                  e1=1;e2=0;e3=0;
            end;
      else if intervalRange="medium" then
            do;
                  e1=0;e2=1;e3=0;
            end;
      else if intervalRange="short" then
            do;
                  e1=0;e2=0;e3=1;
            end;
      else if intervalRange="ultra_sh" then
            do;
                  e1=-1;e2=-1;e3=-1;
            end;
      run;

proc mixed data= time_depression_TimeProductionEC order=data method=REML covtest NOBOUND; /*Option COVTEST provides a
test for heterogeneity. Option 'NOBOUND' is used to prevent a non positive definite estimated R matrix, which often
happens if the between-study variance is close to 0.*/
      class studyNo; /*Classification variable studyNo contains the study number*/
```

```
      model g_i= e1 e2 e3 /outp = time_depression_TProdout s ddfm=KR CL; /*g_i is the effect size estimate. The model
contains an intercept and the three indicator variables representing the effect-coded covariate (interval range).
Degrees of freedom computed according to Kenward and Roger (1997). Option 's': print fixed effects estimates. Option
'outp': save estimates in new data set. Option 'CL': print confidence limits for the estimated parameters*/
            random int /subject=studyNo G s; /*Random intercept model, intercept allowed to vary between studies.
Option 's': print random effects estimates. Option 'G': print G matrix */
            repeated /group=studyNo type=VC R ; /*Each study has its own variance. Covariance matrix R has type
"variance components" (Wolfinger, 1993). Option 'R': print estimated R matrix */
            parms (2) (.0919) (.0691) (.1006) (.0848) (.0966) (.0914) (.0526) (.0647) (.1012) (.1441) (.0838) (.0918)
            (.0914) (.0528) (.1627) (.0909) /eqcons=2 to 17; /*List of parameters: contains a starting value for the
            between study variance (first entry) and the variances (Var(g_i)) for each study (see Appendix A.2), which
            should be kept fixed (Option 'eqcons')*/
run;
```

In the SAS output, the row "Intercept" in Table "Solution for Fixed Effects" provides the estimate of the population effect size θ and its standard error. Row "Intercept" in Table "Covariance Parameter Estimates" provides the estimate of the inter-study variance $\tau^2$ and a test for homogeneity of the single-study means.

#### 2.1.7.4 Appendix B.2

SAS syntax for the random effects meta-regression model used to compute Type 3 tests for the effects of the covariate interval range.

This analysis uses the same model as in Appendix B.1, but with dummy-coded rather than effect-coded covariate.

```
proc mixed data= time_depression_TimeProductionEC order=data method=REML covtest NOBOUND; /*Option COVTEST provides a
test for heterogeneity. Option 'NOBOUND' is used to prevent a non positive definite estimated R matrix, which often
happens if the between-study variance is close to 0.*/
      class studyNo intervalRange; /*Classification variable studyNo contains the study number, variable intervalRange
specifies the interval range*/
      model g_i= intervalRange /outp = time_depression_TProdout s ddfm=KR CL; /*g_i is the effect size estimate. The
model contains an intercept and the dummy-coded variable representing the covariate (interval range). Degrees of
freedom computed according to Kenward and Roger (1997). Option 's': print fixed effects estimates. Option 'outp': save
estimates in new data set. Option 'CL': print confidence limits for the estimated parameters*/
            random int /subject=studyNo G s; /*Random intercept model, intercept allowed to vary between studies.
Option 's': print random effects estimates. Option 'G': print G matrix */
            repeated /group=studyNo type=VC R ; /*Each study has its own variance. Covariance matrix R has type
"variance components" (Wolfinger, 1993). Option 'R': print estimated R matrix */
            parms (2) (.0919) (.0691) (.1006) (.0848) (.0966) (.0914) (.0526) (.0647) (.1012) (.1441) (.0838) (.0918)
            (.0914) (.0528) (.1627) (.0909) /eqcons=2 to 17; /*List of parameters: contains a starting value for the
            between study variance (first entry) and the variances (Var(g_i)) for each study (see Appendix A.2), which
            should be kept fixed (Option 'eqcons')*/
            lsmeans intervalRange / adjdfe=row; /*This provides least-squares estimates (marginal means) of the effect
size per level of the classification variable intervalRange, plus confidence intervals. Option 'adjdfe': use Kenward &
Roger degrees-of-freedom specific for each level of intervalRange*/
run;
```

In the SAS output, the table "Type 3 Tests of Fixed Effects" displays the test for an effect of the covariate (interval range). Row "Intercept" in table "Covariance Parameter Estimates" provides the estimate of the inter-study variance $\tau^2$ and a test for homogeneity of the single-study means.

### 2.1.8 Supplementary Material

#### *2.1.8.1 Supplementary Table 1*

Diagnostic criteria, treatment, gender, and age of depressive and control subjects as well as covered tasks and time interval ranges for each study included in the meta-analysis.

| **Study and country** | **Diagnostics** | **Treatment** | **Depressives** | | | | | **Controls** | | | | | **Task** | **I. R.** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | **Age** | | **Gender** | | | **Age** | | **Gender** | | | | |
| | | | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | | |
| *Wyrick & Wyrick (1977) (US)* | MAACL | varying dosages of amitriptyline | 35.3 | | 15 | 15 | 30 | 33.6 | | 15 | 15 | 30 | Est(p,r) | Sho<br>Med<br>Lon |
| | | | | | | | | | | | | | Exp | |
| *Kitamura and Kumar (1982) (UK)* | Depressive mood; PSE (Catego Computer System) | "Variety of treatment" | 42.4 | | 13 | 10 | 23 | | | "Matched" | | | Exp | |
| *Kitamura and Kumar (1983) (UK)* | *see (1982)* | | | | | | | | | | | | Est(p,r) | Sho<br>Med<br>Lon |
| | | | | | | | | | | | | | Pro | Lon |
| *Tysk (1984) (Sweden)* | DSM-III (maj dep) | "Psychotropic drugs" | | | | | 56 | | | | | 60 | Est(p,r) | Sho<br>Med |
| | | | | | | | | | | | | | Pro | Sho<br>Med |

| **Study and country** | **Diagnostics** | **Treatment** | **Depressives** | | | | | **Controls** | | | | | **Task** | **I. R.** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | **Age** | | **Gender** | | | **Age** | | **Gender** | | | | |
| | | | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | | |
| *Münzel, et al. (1988) (Germany)* | ICD-9 (unipol, bipol, neurotic) | "Drugs" | 44.3 | 13.0 | 11 | 36 | 47 | 44.0 | 13.0 | 4 | 12 | 16 | Est(r) | Med<br>Lon |
| | | | | | | | | | | | | | Pro | Sho<br>Med |
| | | | | | | | | | | | | | Exp | |
| *Rammsayer (1990) (Germany)* | DSM-III (maj dep) | | 57.25 | 11.7 | 15 | 18 | 33 | 31.50 | 10.95 | 36 | 44 | 80 | Dis | U sh |
| *Mundt, et al. (1998) (Germany)* | ICD-9;<br>DSM-III (maj dep) | | 43.20 | 9.38 | 24 | 16 | 40 | 40.20 | 12.50 | 8 | 7 | 15 | Pro | Sho<br>Med |
| | | | | | | | | | | | | | Rep | Med |
| | | | | | | | | | | | | | Exp | |
| *Sevigny, et al. (2003) (Canada)* | Dep: BDI score > 14<br>Con: BDI score < 8 | none | 22.3 | | 2 | 13 | 15 | 22.1 | | 5 | 15 | 20 | Dis | U sh<br>Sho |
| *Bschor, et al. (2004) (Germany)* | DSM-4;<br>HDRS score > 14 | "psychotropic medication" | 52.8 | 15.3 | 6 | 26 | 32 | 45.7 | 17.8 | 13 | 18 | 31 | Est(p) | Sho<br>Med<br>Lon |
| | | | | | | | | | | | | | Pro | Sho<br>Med |
| | | | | | | | | | | | | | Exp | |

| Study and country | Diagnostics | Treatment | Depressives | | | | | Controls | | | | | Task | I. R. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Age | | Gender | | | Age | | Gender | | | | |
| | | | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | | |
| *Mahlberg, et al. (2008) (Germany)* | DSM-4;<br>HDRS score > 14 | "psychotropic medication" | 51.7 | 15.7 | 6 | 22 | 28 | 46.1 | 18.0 | 14 | 16 | 30 | Rep | Sho<br>Med |
| *Gil and Droit-Volet (2009) (France)* | Dep: BDI score ≥ 9<br>Con: BDI score 0 to 3 | none | 27 | 4.5 | 6 | 13 | 19 | 26 | 3 | 24 | 19 | 43 | Dis | U sh |
| *Msetfi, et al. (2012) exp. 1 (UK)* | Dep: BDI score ≥ 9<br>Con: BDI score < 9 | none | "matched" | | 9 | 9 | 18 | "matched" | | 9 | 9 | 18 | Dis | U sh<br>Sho |
| *Msetfi, et al. (2012) exp. 2 (UK)* | Dep: BDI score ≥ 9<br>Con: BDI score < 9 | none | "matched" | | 22 | 20 | 42 | "matched" | | 20 | 17 | 37 | Dis | U sh<br>Sho |
| *Kornbrot et al. (2013) (UK)* | Dep: BDI score ≥ 7<br>Con: BDI score < 7 | none | 19.90 | | | | 21 | 19.90 | | | | 21 | Est(p)<br><br>Pro | Sho<br>Med<br>Sho<br>Med |

| Study and country | Diagnostics | Treatment | Depressives | | | | | Controls | | | | | Task | I. R. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Age | | Gender | | | Age | | Gender | | | | |
| | | | *M* | *SD* | *N* Ma | *N* Fe | *N* | *M* | *SD* | *N* Ma | *N* Fe | *N* | | |
| *Oberfeld, et al. (2014) (Germany)* | DSM-IV (maj dep) | partly on medication | 35.23 | 10.9 | 7 | 15 | 22 | 25.03 | 4.58 | 11 | 11 | 22 | Est(p,r) | U sh<br>Sho<br>Med |
| | | | | | | | | | | | | | Pro | U sh<br>Sho<br>Med |
| | | | | | | | | | | | | | Rep | U sh<br>Sho<br>Med |
| | | | | | | | | | | | | | Exp | |
| *Mioni et al. (in prep) (Italy)* | BDI | partly medicated;<br>psycho-therapy | 49.85 | 8.08 | 2 | 10 | 12 | 45.82 | 13.65 | 5 | 12 | 17 | Pro | U sh<br>sho |
| | | | | | | | | | | | | | Rep | U sh<br>sho |

*Note.* Empty cells represent missing information in primary studies. Kitamura and Kumar (1982) and Kitamura and Kumar (1983) used the same sample of subjects. Data in quotation marks refer to statements from the primary studies. Multiple Affect Adjective Check List (MAACL); Present State Examination (PSE); Hamilton's Rating Scale for Depression (HRS); Beck Depression Inventory II (BDI); Est = Time estimation; (r) = retrospective; (p) = prospective; (p,r) = prospective and retrospective; Pro = Time production; Rep = Time reproduction; Dis = Duration discrimination; Exp = Time experience (unrelated to interval ranges); I. R. = Interval range; lon = long; med = medium; sho = short; u sh = ultra short; maj dep = major depression; unipol = unipolar depression; bipol = bipolar depression; neurotic = neurotic depression: Ma = male; Fe = Female; *M* = Mean; *SD* = Standard deviation; *N* = sample size.

### 2.1.8.2 *Supplementary Table 2*

Effect sizes (*g*)*,* effect size variances (*Var g*), and corresponding 95% confidence intervals for each pair of means reported by the primary studies.

| **Study** | **Task** | **Interval range** | **Interval duration [s]** | ***g*** | ***Var(g)*** | **$CI_L$** | **$CI_U$** | **Note (discrimination and experience measures)** |
|---|---|---|---|---|---|---|---|---|
| *Wyrick & Wyrick (1977)* | Est | short | 5 | 0.05 | 0.07 | –0.45 | 0.56 | |
| | | | 10 | 0.22 | 0.07 | –0.28 | 0.73 | |
| | | medium | 20 | 0.20 | 0.07 | –0.30 | 0.71 | |
| | | | 80 | 0.23 | 0.07 | –0.27 | 0.74 | |
| | | | 160 | 0.70 | 0.07 | 0.18 | 1.22 | |
| | | | 240 | 0.54 | 0.07 | 0.03 | 1.06 | |
| | | long | 900 | 2.13 | 0.10 | 1.50 | 2.77 | |
| | | | 1800 | 2.09 | 0.10 | 1.47 | 2.72 | |
| | Exp | | | 1.46 | 0.08 | –2.03 | –0.89 | Questionnaire: How quickly does time pass in different situations? (scale: 1 *slowly* – 7 *quickly*) |
| *Kitamura & Kumar (1982)* | Exp | | | 0.87 | 0.10 | –1.47 | –0.26 | Questionnaire: How quickly does time pass in different everyday life situations? (scale: 1 *slowly* – 5 *quickly*) |
| | | | | 0.27 | 0.09 | –0.85 | 0.31 | |
| | | | | 0.00 | 0.09 | –0.58 | 0.58 | |

| Study | Task | Interval range | Interval duration [s] | *g* | *Var(g)* | $CI_L$ | $CI_U$ | Note (discrimination and experience measures) |
|---|---|---|---|---|---|---|---|---|
| *Kitamura & Kumar (1983)* | Est | short | 5 | 0.36 | 0.09 | –0.23 | 0.94 | |
| | | | 5 | 0.90 | 0.10 | 0.29 | 1.51 | |
| | | | 5 | 0.39 | 0.09 | –0.20 | 0.99 | |
| | | | 10 | 0.48 | 0.09 | –0.11 | 1.07 | |
| | | | 10 | 0.74 | 0.09 | 0.15 | 1.34 | |
| | | | 10 | 0.33 | 0.09 | –0.26 | 0.93 | |
| | | medium | 20 | 0.60 | 0.09 | 0.01 | 1.20 | |
| | | | 20 | 0.87 | 0.10 | 0.27 | 1.48 | |
| | | | 20 | 0.40 | 0.09 | –0.20 | 1.00 | |
| | | | 80 | 0.55 | 0.09 | –0.05 | 1.14 | |
| | | | 80 | 0.83 | 0.09 | 0.22 | 1.43 | |
| | | | 80 | 0.60 | 0.10 | 0.00 | 1.21 | |
| | | | 160 | 0.59 | 0.09 | –0.01 | 1.19 | |
| | | | 160 | 0.68 | 0.09 | 0.09 | 1.28 | |
| | | | 160 | 0.44 | 0.09 | –0.16 | 1.04 | |
| | | | 240 | 0.17 | 0.09 | –0.42 | 0.76 | |

| **Study** | **Task** | **Interval range** | **Interval duration [s]** | ***g*** | ***Var(g)*** | **$CI_L$** | **$CI_U$** | **Note (discrimination and experience measures)** |
|---|---|---|---|---|---|---|---|---|
| *Kitamura & Kumar (1983)* | Est | medium | 240 | 0.48 | 0.09 | –0.11 | 1.06 | |
| | | | 240 | 0.16 | 0.09 | –0.43 | 0.76 | |
| | | long | 900 | 0.07 | 0.09 | –0.51 | 0.66 | |
| | | | 900 | 0.15 | 0.09 | –0.43 | 0.74 | |
| | | | 900 | 0.18 | 0.09 | –0.42 | 0.78 | |
| | | | 1800 | 0.43 | 0.09 | –0.17 | 1.02 | |
| | | | 1800 | 0.13 | 0.09 | –0.45 | 0.71 | |
| | | | 1800 | 0.22 | 0.09 | –0.37 | 0.81 | |
| | Pro | long | 30 | 0.86 | 0.10 | –1.47 | –0.26 | |
| | | | 30 | 0.36 | 0.09 | –0.94 | 0.22 | |
| | | | 30 | 0.49 | 0.09 | –1.09 | 0.11 | |
| *Tysk (1984)* | Est | short | 7.5 | 0.44 | 0.05 | –0.02 | 0.89 | |
| | | medium | 17.5 | 0.10 | 0.05 | –0.34 | 0.55 | |
| | | | 27.5 | 0.41 | 0.05 | –0.05 | 0.86 | |
| | | | 450 | –0.61 | 0.06 | –1.07 | –0.14 | |
| | Pro | short | 10 | –0.27 | 0.05 | –0.19 | 0.72 | |
| *Tysk (1984)* | Pro | medium | 20 | –0.39 | 0.05 | –0.06 | 0.84 | |
| | | | 30 | 0.78 | 0.06 | –1.24 | –0.32 | |

| Study | Task | Interval range | Interval duration [s] | *g* | *Var(g)* | $CI_L$ | $CI_U$ | Note (discrimination and experience measures) |
|---|---|---|---|---|---|---|---|---|
| *Münzel, et al. (1988)* | Est | medium | 240 | 0.38 | 0.08 | –0.19 | 0.95 | |
| | | | 240 | 0.24 | 0.08 | –0.33 | 0.81 | |
| | | | 240 | 0.53 | 0.09 | –0.05 | 1.10 | |
| | | long | 1800 | 0.00 | 0.08 | –0.57 | 0.57 | |
| | Pro | medium | 240 | 0.37 | 0.08 | –0.94 | 0.20 | |
| | | short | 1 | –0.53 | 0.09 | –0.05 | 1.10 | |
| | | | 5 | 0.08 | 0.08 | –0.65 | 0.48 | |
| | | | 10 | 0.20 | 0.08 | –0.76 | 0.37 | |
| | Exp | | | 0.93 | 0.09 | –1.52 | –0.34 | Questionnaire: How quickly does time pass in different everyday life situations? (scale: 1 *slowly* – 5 *quickly*) |
| | | | | –0.52 | 0.10 | –0.1 | 1.14 | |
| | | | | –0.17 | 0.10 | –0.44 | 0.79 | |
| | | | | –0.20 | 0.10 | –0.41 | 0.82 | |
| | | | | 0.70 | 0.10 | –1.32 | –0.07 | |
| *Rammsayer (1990)* | dis | ultra sh | 0.05 | 0.95 | 0.05 | 0.52 | 1.37 | Two–interval task (First or second interval longer?). Performance measure: difference limen (DL) corresponding to 70.7% correct (adaptive procedure) |

| Study | Task | Interval range | Interval duration [s] | *g* | *Var(g)* | $CI_L$ | $CI_U$ | Note (discrimination and experience measures) |
|---|---|---|---|---|---|---|---|---|
| *Mundt, et al. (1998)* | Pro | short | 10 | –0.13 | 0.09 | –0.47 | 0.72 | |
| | | medium | 30 | 0.87 | 0.10 | –1.48 | –0.25 | |
| | | | 60 | 0.74 | 0.10 | –1.35 | –0.13 | |
| | | | 120 | 0.61 | 0.10 | –1.21 | 0.00 | |
| | Rep | medium | 10 | 0.18 | 0.09 | –0.41 | 0.78 | |
| | | | 30 | –0.92 | 0.10 | –1.54 | –0.30 | |
| | | | 60 | –0.55 | 0.09 | –1.15 | 0.05 | |
| | | | 120 | –0.63 | 0.10 | –1.23 | –0.02 | |
| | Exp | | | 0.83 | 0.10 | –1.44 | –0.22 | Visual analogue scale: How quickly did the last hour pass (slowly – fast)? |
| | | | | 0.95 | 0.10 | –1.57 | –0.33 | |
| *Sevigny, et al. (2003)* | Dis | ultra sh | 0.1 | 0.55 | 0.12 | –0.13 | 1.24 | One interval task (short or long interval presented?) Performance measure: %correct |
| | | | 0.5 | –0.06 | 0.12 | –0.73 | 0.61 | |
| | | short | 1.2 | 0.75 | 0.12 | 0.06 | 1.44 | |

| Study | Task | Interval range | Interval duration [s] | *g* | *Var(g)* | $CI_L$ | $CI_U$ | Note (discrimination and experience measures) |
|---|---|---|---|---|---|---|---|---|
| *Bschor, et al. (2004)* | Est | short | 8 | 0.00 | 0.06 | –0.50 | 0.49 | |
| | | medium | 43 | –0.15 | 0.06 | –0.65 | 0.34 | |
| | | | 109 | –0.40 | 0.06 | –0.90 | 0.10 | |
| | | long | 760.2 | 0.59 | 0.07 | 0.09 | 1.10 | |
| | Pro | short | 7 | –0.39 | 0.06 | –0.11 | 0.89 | |
| | | medium | 35 | 0.67 | 0.07 | –1.18 | –0.16 | |
| | | | 90 | 1.00 | 0.07 | –1.53 | –0.48 | |
| | exp | | | 0.98 | 0.07 | –1.51 | –0.46 | Visual analogue scale: How do you experience the flow of time today (slow – fast)? |
| *Mahlberg, et al. (2008)* | Rep | short | 1 | 1.09 | 0.08 | 0.53 | 1.64 | |
| | | | 6 | 0.98 | 0.08 | 0.44 | 1.53 | |
| | | medium | 37 | 0.09 | 0.07 | –0.43 | 0.60 | |
| *Gil & Droit–Volet* | Dis | ultra sh | 1 | –0.17 | 0.08 | –0.71 | 0.37 | Temporal bisection task;<br>anchor durations (a.d.): 400 and 1600 ms<br>(Is the interval presented more similar to the short or to the long a.d.?) Performance measure: weber fractions |
| *Msetfi, et al. (2012) Exp. 1* | Dis | short | 1 | 0.84 | 0.12 | 0.16 | 1.52 | Two–interval task (First or second interval longer?). Performance measure: DL corresponding to 75% correct (adaptive procedure) |
| | | ultra sh | 0.05 | –0.28 | 0.11 | –0.93 | 0.38 | |
| *Msetfi, et al. (2012) Exp. 2* | Dis | short | 1 | 0.34 | 0.05 | –0.10 | 0.79 | See *Msetfi et al. (2012) Exp. 1* |
| | | ultra sh | 0.05 | –0.14 | 0.05 | –0.59 | 0.30 | |

| **Study** | **Task** | **Interval range** | **Interval duration [s]** | ***g*** | ***Var(g)*** | **$CI_L$** | **$CI_U$** | **Note (discrimination and experience measures)** |
|---|---|---|---|---|---|---|---|---|
| *Oberfeld, et al. (2014)* | Est | ultra sh | 0.5 | –0.12 | 0.09 | –0.71 | 0.47 | |
| | | short | 2 | 0.27 | 0.09 | –0.32 | 0.87 | |
| | | medium | 60 | –0.12 | 0.09 | –0.71 | 0.47 | |
| | | long | 1800 | –0.33 | 0.09 | –0.93 | 0.26 | |
| | Pro | ultra sh | 0.5 | 0.01 | 0.09 | –0.60 | 0.58 | |
| | | short | 2 | 0.20 | 0.09 | –0.80 | 0.39 | |
| | | medium | 60 | 0.21 | 0.09 | –0.80 | 0.38 | |
| | Rep | ultra sh | 0.5 | –0.10 | 0.09 | –0.69 | 0.50 | |
| | | short | 2 | –0.09 | 0.09 | –0.68 | 0.51 | |
| | | medium | 60 | –0.37 | 0.09 | –0.97 | 0.23 | |
| | Exp | | | 0.05 | 0.09 | –0.64 | 0.54 | Visual analogue scale: How do you experience the flow of time today (slow – fast)? |

| Study | Task | Interval range | Interval duration [s] | *g* | *Var(g)* | $CI_L$ | $CI_U$ | Note (discrimination and experience measures) |
|---|---|---|---|---|---|---|---|---|
| *Kornbrot, et al. (2013)* | Est | short | 4.33 | –0.40 | 0.10 | –1.01 | 0.21 | |
| | | medium | 10.86 | –0.78 | 0.10 | –1.40 | –0.15 | |
| | | | 19.21 | –0.86 | 0.10 | –1.50 | –0.23 | |
| | | | 33.78 | –0.55 | 0.10 | –1.16 | 0.07 | |
| | | | 60.92 | –0.52 | 0.10 | –1.14 | 0.09 | |
| | Pro | short | 4.33 | –0.71 | 0.10 | 0.08 | 1.33 | |
| | | medium | 10.86 | –0.66 | 0.10 | 0.04 | 1.28 | |
| | | | 19.21 | –0.76 | 0.10 | 0.14 | 1.39 | |
| | | | 33.78 | –0.67 | 0.10 | 0.05 | 1.29 | |
| | | | 60.92 | –0.59 | 0.10 | –0.03 | 1.21 | |
| *Mioni, et al. (in prep)* | Pro | ultra sh | 0.5 | –1.09 | 0.16 | 0.30 | 1.88 | |
| | | short | 1 | –0.69 | 0.15 | –0.07 | 1.45 | |
| | | | 1.5 | 0.02 | 0.14 | –0.76 | 0.72 | |
| | Rep | ultra sh | 0.5 | 0.47 | 0.15 | –0.28 | 1.22 | |
| | | short | 1 | 0.77 | 0.15 | 0.01 | 1.54 | |
| | | | 1.5 | 0.24 | 0.14 | –0.51 | 0.98 | |

## 2.2 Meta-analysis of time perception and temporal processing in schizophrenia: Differential effects on precision and accuracy[2]

### 2.2.1 Abstract

Numerous studies have reported that time perception and temporal processing are impaired in schizophrenia. In a meta-analytical review, we differentiate between time perception (judgments of time intervals) and basic temporal processing (e.g., judgments of temporal order) as well as between effects on accuracy (deviation of estimates from the veridical value) and precision (variability of judgments). In a meta-regression approach, we also included the specific tasks and the different time interval ranges as covariates. The meta-analysis comprises 68 studies of the past 65 years, corresponding to 938 schizophrenic and 1041 healthy control participants. Independent of tasks and interval durations, our results demonstrate that time perception and basic temporal processing are less precise (more variable) in patients (Hedges' $g > 1.00$), whereas effects of schizophrenia on accuracy of time perception are rather small and task-dependent ($g$ between 0.04 for time reproduction and 0.53 for time production). Our review also shows that several aspects, e.g., potential influences of medication, have not yet been investigated in sufficient detail. In conclusion, the results are in accordance with theoretical assumptions and the notion of a more variable internal clock in schizophrenic patients, but not with a strong effect of schizophrenia on clock speed.

### 2.2.2 Introduction

Over the last decades, numerous studies have reported that the perception of time and the processing of temporal information is distorted in clinical disorders such as depression (Bschor et al., 2004; Kornbrot et al., 2013; Thoenes & Oberfeld, 2015; Wyrick & Wyrick, 1977), Parkinson's disease (Allman & Meck, 2012; Malapani, Deweer, & Gibbon, 2002; Meck, 1996), and schizophrenia (Martin et al., 2014; Rammsayer, 1990; Roy, Grondin, & Roy, 2012). However, especially in the case of schizophrenia, empirical studies largely differ in the tasks and methods used, and the outcomes of the studies do not always agree. The

[2] A paper based on this manuscript (pages 56 to 102) has been published together with PD Dr. Daniel Oberfeld: Thönes, S. & Oberfeld, D. (2017). Meta-analysis of time perception and temporal processing in schizophrenia: Differential effects on precision and accuracy. *Clinical Psychology Review, 54*, 44-64.

present study provides a meta-analytical review of the literature on time perception and temporal processing in schizophrenia from the past 65 years.

With regard to the conceptual and methodological heterogeneity of the literature, it is important to distinguish between different aspects of temporal information processing, and different aspects of human performance in the relevant tasks. These distinctions have not yet been addressed in a systematic review of studies on time perception and temporal processing in schizophrenia.

First, we suggest that *time perception* in the sense of explicit judgments of the durations of events, or the production of time intervals should be distinguished from tasks like judging the simultaneity of two events or the order of two stimuli, which we refer to as *temporal processing*. The latter tasks represent lower level processing of temporal stimulus features, and index for example the temporal acuity of the visual or auditory system, but without the necessity of explicit judgments of duration.

Second, the participants' performance can be analyzed in terms of *accuracy*, which indexes the (signed) deviation of a judgment from the veridical value, and in terms of *precision,* which refers to the *variability* of the judgments (cf. Grondin, 2010). According to the scalar expectancy theory (SET; sometimes referred to as scalar timing theory), which represents the most influential theory of time perception and temporal processing, humans are able to estimate time on average accurately and precisely (Gibbon, 1977; Gibbon et al., 1984). However, several factors, such as the level of bodily arousal, can lead to systematic deviations of the time estimates from the veridical value, and to less precise (i.e., more variable) temporal judgments (Droit-Volet & Meck, 2007). In the context of clinical disorders, both measures of temporal performance, accuracy and precision, are considered to be altered in schizophrenia (Allman & Meck, 2012; Bolbecker et al., 2014).

Third, apart from these two major distinctions, different tasks involve different components of human information processing and behavior. For example, some tasks that were used to study time perception in schizophrenia require timed motor responses, while other tasks require a perceptual judgment but no temporally precise motor response. A recent meta-analysis on time perception in schizophrenic patients has considered the latter differentiation (Ciullo, Spalletta, Caltagirone, Jorge, & Piras, 2015). However, the authors did not address the importance of the distinction between time perception and temporal processing, nor the distinction between accuracy and precision. Also, several relevant studies were not included in their meta-analysis.

Before describing our meta-analytic strategy, we now discuss the different tasks used in the relevant literature, and the distinction between accuracy and precision.

#### *2.2.2.1 Tasks used to study time perception and basic temporal processing*

Concerning the first distinction, tasks used to study *time perception* in general and in the context of schizophrenia encompass the well-established cases of *a) verbal time estimation, b) time production, c) time reproduction*, and *d) duration discrimination,* as well as *e) rhythm production tasks* (e.g., Vorberg & Wing, 1996).

In *verbal time estimation*, a time interval is presented, defined for instance by the inter-onset interval (IOI) between two brief tones or light flashes, and the participant gives an estimate of this time interval in conventional chronometric units like seconds or minutes (Broadhurst, 1969; Carlson & Feinberg, 1968; Clausen, 1950; Densen, 1977; Dilling & Rabin, 1967; Johnson & Petzel, 1971; Orme, 1966; Pearl & Berg, 1963; Roy et al., 2012; Rutschmann, 1973; Tracy et al., 1998). Such a task is most frequently used for *prospective time estimation,* where the participant is aware that time intervals are to be judged. It can also be used in a *retrospective* manner, however. For instance, the participant could be asked to estimate the duration that has elapsed since the beginning of the experiment, without having been informed at the beginning of the experimental session that such a time estimation will be required (Oyanadel & Buela-Casal, 2014; Rabin, 1957; Tysk, 1983a; Wahl & Sieg, 1980). In this case, usually the perception of longer time intervals (in the range of several minutes to hours) is investigated, compared to intervals in the second or minute range that are typically used in prospective time estimation. The verbal estimation task provides information about both measures of temporal performance, accuracy and precision. The signed deviation of the mean verbal judgment from the veridical duration (*signed error*) measures the accuracy, while the variability of the verbal estimates (e.g., the standard deviation of the estimates across 10 presentations of the same temporal interval; often termed *variable error*) is a measure of precision.

In a *time production task* (*b)*, a time interval is defined in terms of conventional chronometric units, i.e. "2.0 s", and the participant is required to produce the interval, for example by giving two motor responses marking its beginning and end (Carlson & Feinberg, 1968; Clausen, 1950; Johnson & Petzel, 1971; Nosachev, 1992; Oyanadel & Buela-Casal, 2014; Tysk, 1983b; van der Veen, Roder, & Smits, 2013; Wahl & Sieg, 1980). Sometimes, several repeated productions rather than just a single production are required on a trial (Turgeon, Giersch, Delevoye-Turrell, & Wing, 2012). The production task also provides information about both accuracy and precision, similar to the verbal estimation task. An important difference to the verbal estimation task is, however, that the production task requires timed motor actions. For this reason, the behavioral results will not only be affected by changes in the cognitive representation of time intervals or the "clock mechanism", but also by factors influencing the motor system (for a discussion see Oberfeld et al., 2014).

In a *time reproduction task (c)*, a time interval is presented as in *a)* and the participant reproduces the interval as in *b)* (Carlson & Feinberg, 1968; Clausen, 1950; Roy et al., 2012; Tracy et al., 1998). Thus, the reproduction task combines the perception and the (motor) production of a time interval, and again provides information on accuracy as well as on precision.

In the case of *duration discrimination (d)*, often a *two-interval task* is used where two time intervals are presented successively and the participant has to decide which interval was longer and which intervals was shorter (Rammsayer, 1990; Todd et al., 2000; Todd, Michie, & Jablensky, 2003; Ulferts, Meyer-Lindenberg, & Gallhofer, 1999; Volz et al., 2001). Based on fitting a psychometric function to the data, or on an adaptive procedure (e.g., Levitt, 1971), the two-interval duration discrimination task provides an estimate of the *duration difference limen*, which is the duration difference between the two stimuli at which the participant is able to identify the longer /shorter interval with, for example, 75% correct responses. The two-interval discrimination tasks measures precision, but does not provide information about accuracy (i.e., about a potential over- or underestimation of time intervals). In *one-interval discrimination tasks*, only a single time interval is presented per trial and has to be compared to a so-called *standard interval* that has either been learnt explicitly prior to the discrimination task (Davalos, Rojas, & Tregellas, 2011; Lhamon & Goldstone, 1973; Waters & Jablensky, 2009) or implicitly during the task (Lhamon & Goldstone, 1956). In the latter case, the participant develops an internal representation of an intermediate standard duration based on the processing of different comparison durations that are slightly longer or shorter than the intermediate standard duration (Nachmias, 2006; Oberfeld, 2014). If the standard interval is presented on each trial before the to-be-judged time interval, this is termed a *reminder task* (e.g., Lapid, Ulrich, & Rammsayer, 2008). A specific one-interval duration discrimination procedure that has been used frequently in the time perception literature in general and also in schizophrenic samples is the *temporal bisection task* (Bolbecker et al., 2014; Carroll, Boggs, O'Donnell, Shekhar, & Hetrick, 2008; Carroll, O'Donnell, Shekhar, & Hetrick, 2009b; Davalos, Kisley, & Ross, 2002; Elvevag et al., 2003; Lee et al., 2009; Lee, Dixon, Spence, & Woodruff, 2006; Penney, Meck, Roberts, Gibbon, & Erlenmeyer-Kimling, 2005). Here, the participant first learns a short and a long *anchor duration* of, for example, 1.0 and 2.0 seconds, respectively. Subsequently, intermediate durations ranging from 1.0 s to 2.0 s are presented and the participant is asked to categorize these as being either more similar to the short anchor duration or to the long anchor duration. In these one-interval tasks, a psychometric function (cf. Treutwein & Strasburger, 1999) is fitted to the data and provides a measure of sensitivity in the duration discrimination task. For example, half the difference between the 75%- and the 25% points on the psychometric functions is often used as a measure of the duration discrimination limen. The difference limen (DL) is a measure of precision and is correlated to measures of precision obtained for example in the time production task described above, and of course to a DL measured in a two-interval discrimination task. Notably, the one-interval tasks also provide information about the average perceived duration of the stimuli, in terms of the 50%- point on the psychometric function, which in the case of the bisection task is often termed the "bisection point" (BP) (Allan & Gibbon, 1991). The signed deviation of the BP from the veridical value (i.e., the arithmetic or geometric mean of the presented time intervals) can be viewed as a measure of accuracy. The sensitivity (precision) in duration discrimination can also be studied by means of a *temporal deviant detection task,* where not only one or two time intervals are presented, but a rhythmic sequence. In an isochronous rhythmic sequence (constructed as a sequence of identical IOIs), occasional phase shifts (temporal deviants) are presented. These are tones or other events with onsets occurring earlier or later than implied by the isochronous rhythm. The task is to detect such a phase shift (Bourdet, Brochard, Rouillon, & Drake, 2003; Davalos, Kisley, & Freedman, 2005; Turgeon et al., 2012). Based on the participant's detection performance, a difference limen for temporal deviant detection can be determined.

Several variants of *rhythm production tasks (e)* have been used to study potential effects of schizophrenia on time perception. In a *continuation tapping task*, the trial starts with the presentation of an isochronous rhythmic sequence and the participants are asked to tap in synchrony with it (Carroll, O'Donnell, Shekhar, & Hetrick, 2009a; Papageorgiou et al., 2013). After some seconds, the sequence stops and the participants are required to continue their tapping at exactly the same rate. The variability of the produced IOIs provides information about the precision of the internal representation of the to-be-produced IOIs, but also about motor variability. Only few studies also provided information about the mean of the produced IOIs, which could be viewed as a measure of accuracy. The continuation tapping task is closely related to the reproduction task, with the difference that in a reproduction task only one time interval is presented and is reproduced once, while in the continuation tapping task first several identical time intervals are presented in a rhythmic sequence and then the participant is required to reproduce these time intervals several times. In a *synchronization tapping task,* the participants provide motor responses (e.g., by tapping with the index finger on a response key) in synchrony to a continually presented rhythmic sequence (Jirsa, Libiger, Mohr, Radil, & Indra, 1996). The deviation of the taps from the onsets of the stimuli in the rhythmic sequence is analyzed and its variability provides information about precision. This task requires anticipatory motor responses (e.g., Fraisse, 1982): for tapping in synchrony with an isochronous rhythm, the participant needs to time his or her next tap so that the time interval between the preceding rhythmic event and the tap corresponds to the IOIs presented in the rhythmic sequence. For this reason, the synchronization tapping task could be viewed as a rhythmic variant of a reproduction task. However, due to the continuous presentation of the rhythm, the participant receives immediate information about the deviation of his or her tap from the rhythmic event. Thus, the synchronization tapping could be described as a closed-loop task, while the reproduction task is an open-loop task. It has been suggested that synchronization tapping involves automatic and unconscious phase corrections that are different from time judgments obtained in other paradigms (e.g., Repp, 2000). For these reasons, synchronization tapping involves rather different processes than the other time-perception tasks. A third variant of finger tapping is the *spontaneous tapping task* (Delevoye-Turrell, Wilquin, & Giersch, 2012). Here, the participant is asked to tap an isochronous rhythm at a self-selected, subjectively preferred rate. Spontaneous tapping is similar to time production, with several successive productions of the same time interval, except that no interval duration is specified by the experimenter. The variability of the produced time intervals is a measure of precision, while the task provides no information about accuracy because a "veridical value" is not defined here.

Laboratory tasks that have been used to investigate basic *temporal processing* not involving judgments of duration (i.e., not measuring time perception) in schizophrenic patients are *a) judgments of simultaneity, b) temporal-order judgments (TOJ), and c) gap detection.*

In *a)*, two stimuli are presented either successively or simultaneously and the participant has to decide whether the onsets of the two stimuli were synchronous or asynchronous. The participant’s performance level is determined based on the IOI between the two stimuli leading to a certain proportion of “simultaneous” responses of, for example, 50%. This is termed the simultaneity threshold or the point of perceived simultaneity. The smaller the simultaneity threshold, the higher the precision of the participant’s judgments, indexing higher temporal resolution of the sensory system (Braus, 2002; Capa et al., 2014; Foucher, Lacambre, Pham, Giersch, & Elliott, 2007; Lalanne, van Assche, & Giersch, 2010; Martin et al., 2013; Schmidt et al., 2011).

In a *temporal-order judgment (TOJ) task b),* two different stimuli (e.g., two tones differing in frequency) are presented successively, and the participant has to indicate which stimulus was presented first (Braus, 2002; Capa et al., 2014). The IOI between the two stimuli at which a certain performance level is reached, e.g., 75% correct responses, serves as a measure of temporal resolution (precision). The smaller the IOI that is sufficient to discriminate the temporal order of two stimuli, the higher is the temporal resolution of the sensory system (Hirsh, 1959).

A *gap detection task* (*c)* requires the detection of short gaps of silence in a continuous auditory signal, e.g. white noise (Plomp, 1964). The smallest gap that can be detected reliably (e.g., with 75% correct) serves as the performance measure (e.g., Todd et al., 2000). The performance in tasks *a)* to c*)* provides measures of the temporal acuity (precision) of the information processing system.

#### *2.2.2.2 Accuracy versus precision*

As discussed above, an important distinction concerns the information provided by the dependent measures. Broadly, the participants' performance in time perception tasks can be analyzed in terms of *accuracy* and in terms of *precision.* The term *accuracy* refers to the deviation of a temporal judgment (or production) from the veridical value. As usual, we analyze accuracy in terms of the signed error, which denotes the signed deviation of the mean verbal judgment from the veridical duration. For example, if schizophrenic patients and healthy controls are required to estimate the duration of a visual stimulus presented for 3.0 s in a verbal estimation task, the average estimate of the patients might be 2.5 s and the average estimate of the controls might be 2.9 s. In this case, the patients show a more negative signed error than the controls, i.e., a stronger underestimation compared to the veridical value than the controls. Thus, the accuracy is affected by systematic shifts in the perceived or produced durations. Note that in some of the studies the accuracy was assessed in terms of the deviation of the estimates from the veridical values divided by the veridical value (relative error), or the ratio between the estimated value and the veridical value. As all of these measures are linear transformations of each other, the effect size estimates are not affected by the choice of the response measure in the primary studies (cf. Thoenes & Oberfeld, 2015).

A complementary measure of accuracy is the *absolute error*, which is the average deviation of the estimates from the veridical value, regardless of the direction of the deviation. Unfortunately, the absolute error was reported only in one study (Tracy et al., 1998). For this reason, we were not able to include this measure of accuracy in our review.

Apart from the systematic deviations of the estimates from the veridical value (*accuracy: signed error*), the data of many tasks also provide information about the *variability* of the estimates across presentations of the same stimulus. This is termed *precision*, or, in terms of Fechner (1860), the "variable error" (VE). The variable error in a verbal estimation task (for example the standard deviation of the estimates across 10 presentations of the same temporal interval) and the difference limen estimated for example from psychometric functions both measure precision, and are closely related (Treisman, 1963). Note that accuracy and precision provide independent information about performance on a task. For instance, a schizophrenic patient and a healthy control participant might both produce an average verbal estimate of 2.9 s for a visual stimulus with a duration of 3.0 s. However, across 30 trials, the variability of the estimates might be higher for the patient (e.g., standard deviation of 60 ms) than for the control participant (e.g., standard deviation of 40 ms). In this case, both participants show equal accuracy, but the controls show higher precision.

In tasks measuring temporal processing but not time perception, the dependent measures like the duration of the just-detectable gap in an auditory stimulus can be viewed as a measure of precision, while the aspect of accuracy plays no role.

#### *2.2.2.3 Task-dependent demands*

With regard to the third distinction, different time perception tasks involve different components of human information processing and behavior. While time production, time reproduction, and finger tapping tasks require timed motor responses, verbal time estimation and duration discrimination do not. Also, it is likely that memory processes are involved differently depending on the particular task. Time production and time estimation require the participant to refer to long-term memory representations of time in terms of chronometric units like seconds or minutes. In time reproduction, duration discrimination, and finger tapping, however, the information necessary for doing the task is presented within a given trial, or within the experimental block, so that these tasks are likely to depend on short-term memory or sensory memory rather than on long-term memory.

#### *2.2.2.4 Theoretical assumptions concerning time perception and temporal processing in schizophrenia*

When reviewing the literature on time perception and temporal processing in schizophrenia qualitatively, the results obtained by means of similar and different tasks are inconclusive so far. However, based on theoretical assumptions, the performance of schizophrenic patients should differ systematically from the performance of healthy control participants. It has been suggested that the clinical symptoms in schizophrenia, such as delusions and hallucinations, may arise from a deficit in the temporal coordination of information processing (Allman & Meck, 2012; Andreasen et al., 1999; Ciullo et al., 2015; Densen, 1977). According to Andreasen et al. (1999), mistimed information transfer in schizophrenic patients may lead to incorrect connections of thoughts and actions, and to misinterpretations of external and internal processes. Accordingly, the *precision* of basic temporal processing and of judgments of duration (time perception) should be impaired and systematically related to the severity of the typical symptoms in schizophrenia (Bolbecker et al., 2014; Rammsayer, 1990). Moreover, in comparison to control participants, patients with schizophrenia exhibit lower activity in brain areas that are involved in temporal processing and duration judgments (Allman & Meck, 2012). In particular, the cerebellum has been proposed to be affected by schizophrenia, which is involved in precise (motor) timing of short durations in particular (Andreasen & Pierson, 2008). Accordingly, schizophrenic patients may show impaired precision of time perception and temporal processing, whereas mean duration judgments (accuracy) remain unaffected.

In order to predict in which specific way schizophrenia may influence the performance on the experimental tasks, it seems sensible to consider the influential *pacemaker-accumulator models* of interval timing (Gibbon et al., 1984; Treisman, 1963). Such models assume an *internal clock* consisting of a pacemaker emitting pulses and an accumulator (or counter) collecting these pulses. The perceived length of a time interval is assumed to depend on the number of accumulated pulses. This means that if a participant's clock runs faster, more pulses get accumulated within a specified interval, and therefore the interval is perceived as longer compared to a participant with a slower clock speed. Accordingly, clock speed affects the mean estimates, that is, the *accuracy* of duration judgments in an experimental task. Importantly, depending on the specific task used, different patterns of results are to be expected. In the *verbal estimation* task, an accelerated internal clock causes the accumulation of more pulses during the presentation of the to-be-judged time interval. Hence, if the clock of schizophrenic patients was accelerated (i.e., "ticking faster"), schizophrenic patients should *overestimate* the duration of the time interval compared to control subjects. However, the opposite relation is predicted for a *production* task. According to the internal clock model, a participant starts to accumulate clock pulses at the start signal, and produces the end of the interval as soon as the accumulated number of pulses reaches a value (stored in long term memory) corresponding to for example "2 s". If the internal clock of schizophrenic patients runs at a faster pace, then they should produce *shorter* intervals than the control subjects, that is, the patients should *underproduce* the time interval compared to healthy controls. In a *reproduction* task, a time interval is not specified in terms of time units but it is presented explicitly before the subject is asked to reproduce it. Here, a faster (or slower) accumulation of pulses should affect the representation of the interval to be timed as well as its reproduction. According to the clock models, the accumulation of pulses during the presentation of the time interval, and the accumulation process during the production phase should be affected in the same way. Therefore, clock speed should have no effect on the reproduced duration, so that no differences between patients and controls are to be expected (Carlson & Feinberg, 1968). Beside the speed of the internal clock, the pulse-to-pulse variability (clock variability) might be increased in schizophrenic patients, which appears plausible based on the theoretical assumptions as discussed above. An increase in clock variability would impair the precision of duration judgments and temporal processing. In a *duration discrimination* task, for example, an increase in clock variability would be reflected in larger difference limens in schizophrenic patients as compared to controls.

#### *2.2.2.5 Structure of the present study*

In the present study, we adopted a meta-analytical approach that investigated whether the assumed and reported effects of schizophrenia are substantial by considering the three important distinctions/aspects of, a) time perception versus temporal processing, b) measures of accuracy versus precision, and c) the specific task used in the study (for example, tasks involving or not involving timed motor responses). Therefore, in a first step, we meta-analyzed the effects of schizophrenia on *accuracy in time perception tasks*. In a second step, we focused on effects on *precision in time perception tasks*. And in a third step, we investigated effects on *precision* in tasks addressing *basic temporal processing* (measured in temporal simultaneity judgments and temporal-order judgments), where information about the participants' performance is provided in terms of precision only. Regarding aspect c) discussed above, in each step, we analyzed potential task-specific effects by defining the specific task as a covariate in the meta-analytical model. Thus, if task-specific effects do exist, the analyses allow for attributing such effects to specific task demands, as for example motor or memory demands. We also investigated whether potential effects of schizophrenia are only substantial for specific interval durations. If effects of schizophrenia on time perception and temporal processing do exist irrespective of the temporal tasks, and time intervals used, this would be compatible with a general timing deficit in schizophrenia (Andreasen et al., 1999).

### 2.2.3 Method

#### *2.2.3.1 Search strategies and study selection*

We searched for relevant studies in Web of Science and Google Scholar. The primary key words were '*schizophrenia*' in conjunction with '*timing'* or '*time*' or '*temporal*'. Additional studies were identified by including the references listed in the studies found in Web of Science and Google Scholar, and by considering the studies that cited the resulting body of literature. Moreover, based on an email list from the "International Conference on Timing and Time Perception" (Corfu, 2014), we sent calls for unpublished data on the topic to more than 100 researchers in the field of timing and time perception. This iterative literature search strategy yielded 68 papers (including four conference contributions) that address time perception or temporal processing in schizophrenia, with 60 papers reporting empirical data.

For the meta-analyses, we selected studies according to the following four criteria.

Criterion 1). The studies had to provide data from a group of schizophrenic patients (or from individuals at high risk of schizophrenia in the case of Penney et al. (2005)) as well as from a control group consisting of healthy adults only. Because healthy participants also tend to produce systematic errors in time perception and temporal processing tasks (e.g., Wearden & Lejeune, 2008), it is uninformative to simply compare judgments of schizophrenic participants, for example verbal time estimates, to the veridical values of the presented time intervals. Hence, for studies that focused on schizophrenic participants only (e.g., Clausen, 1950; Yang et al., 2004), it is not possible to decide whether the reported deviations of the time estimates from the veridical values are specifically related to schizophrenia. Therefore, systematic comparisons between schizophrenic participants and healthy control participants on the single study level are required.

Criterion 2). The report of sample sizes, means, and standard deviations of the response measures, or $t$ or $F$-values had to be sufficiently detailed in order to compute effect sizes (Hedges' $g$) and their variance. If this was not the case, we contacted the authors of the study for papers published after the year 2000 and asked for additional information.

Criterion 3). At least one of the common time perception or temporal processing tasks as listed above had to be used.

Criterion 4). The reported response measures and analyses had to describe the participants' performance in terms of accuracy, precision, or both (separately). Unfortunately, this was not the case for all studies, because many papers did not use the established psychophysical tasks or data analyses. For example, as described above, one-interval discrimination tasks provide information about both precision (in terms of the difference limen) and about accuracy (in terms of the 50% point on the psychometric function) if psychometric functions are fitted to the data. However, several studies analyzed only the proportion correct (e.g., Davalos, Kisley, Polk, & Ross, 2003; Davalos et al., 2002), which is affected by both accuracy and precision. For example, a low proportion correct in the one-interval discrimination task could be due to an imprecise perception of the temporal intervals (precision), but also to a systematic over- or underestimation (accuracy). As a second example, schizophrenic patients might be less precise in their perception of the time interval than control subjects, but at the same time show a smaller tendency towards over- or underestimation (i.e., smaller signed error). In this case, the effects of schizophrenia on precision and accuracy could cancel, resulting in no difference in proportion correct between the two groups.

Only 28 of the 60 empirical studies met these four criteria and were considered for further analyses (see Appendix Table A.1). Lhamon and Goldstone (1973) reported two experiments investigating two independent samples of participants. Therefore, in the analyses, their first and second experiment were treated as two separate studies, resulting in a total of 29 independent studies entering the analyses.

#### *2.2.3.2 Description of the studies*

The 29 studies included in this meta-analysis provided data from a total of 1979 participants (938 schizophrenic and 1041 healthy control participants). Overall, there were 701 male and 237 female participants in the schizophrenic samples, and 611 male and 430 female participants in the control samples. The mean age of schizophrenic participants was 35.70 years (*SD* = 6.30 years), and 33.45 years (*SD* = 6.33 years) for healthy control participants (mean age weighted with regard to the sample size). Three studies did not provide sufficient information about their participants' age (Bolbecker et al., 2014; Lhamon & Goldstone, 1973; Wahl & Sieg, 1980). The median publication year of the studies included in the analyses was 2003 (range: 1956 to 2014).

In 21 studies (72 %), patients were diagnosed according to DSM-II, DMS-III, DSM-IV, ICD-9, or ICD-10 criteria. In one study (Braus, 2002), participants were assigned to the patient group based on the Brief Psychiatric Rating Scale (Overall & Gorham, 1962). A minority of 7 studies (24 %) did not specify the diagnostic criteria used. These studies usually recruited hospitalized schizophrenic patients that were currently treated in psychiatric clinics. The "schizophrenic" sample in the study by Penney et al. (2005) was comprised of participants at genetically high risk for schizophrenia (no acute schizophrenic patients).

In most studies, some patients were under medication, others not. The exact number of medicated participants was not reported in most of the studies, and only one study reported separate data for participants on and off medication (Braus, 2002). Three studies reported that none of their participants received medication (Broadhurst, 1969; Johnson & Petzel, 1971; Penney et al., 2005). In most studies, patients were (partially) treated with neuroleptics (antipsychotic medication).

Appendix Table A.2 provides an overview of study-specific diagnostics and data on age and gender as far as reported by the studies. Also, covered tasks, interval durations, and dependent measures for each study are described. The transcription of the data from the studies was double-checked by independent observers.

#### *2.2.3.3 Preprocessing and effect size estimates*

Two studies (Braus, 2002; Schmidt et al., 2011) reported data for two samples of schizophrenic patients (medicated vs. unmedicated and first episode vs. chronic schizophrenia). For both studies, we averaged the reported data (means and standard deviations; weighted with regard to the sample size) across the two patient samples.

Based on the reported means, standard deviations, and sample sizes for schizophrenic groups ($M_s$, $SD_s$, $n_s$) and control groups ($M_c$, $SD_c$, $n_c$), we calculated Hedges' $g$ as an effect size index, which is an estimate of the standardized mean difference between the two populations. According to Hedges and Olkin (1985), $g$ is defined as shown in Equation Eq.B.1, where $s$ is the pooled sample standard deviation (see Eq.B.2). According to Hofmann et al. (2010), the magnitude of $g$ may be interpreted based on the conventions for the common effect size estimator $d$ (small: $\geq 0.2$; medium: $\geq 0.5$; large: $\geq 0.8$) (Cohen, 1988).

For seven studies (Capa et al., 2014; Carroll et al., 2009a; Elvevag et al., 2003; Johnson & Petzel, 1971; Lalanne et al., 2010; Lhamon & Goldstone, 1973; Turgeon et al., 2012), because means and standard deviations were not reported in sufficient detail, $g$ was calculated based on the presented *F* values and sample sizes (see Eq.B.3, Rosnow & Rosenthal, 1996). For Schmidt et al. (2011), *g* was calculated based on the reported *t* value and the sample sizes (see Eq.B.4, Rosnow & Rosenthal, 1996). When *g* was computed from *F* or *t* values, the sign of *g* was determined based on the reported means for the patient group and the control group.

Based on *g* and the reported sample sizes, we then determined the asymptotic variance of *g.* According to Hedges and Olkin (1985), the asymptotic variance of *g*, denoted *Var(g),* is calculated as shown in Eq.B.5.

Following the preprocessing explained above, one value of *g* (and *Var(g)*) was computed for each pair of means reported in the selected studies. Because finger tapping tasks were rare in the resulting body of studies that entered the analyses, we considered continuation tapping (Carroll et al., 2009a) as time reproduction, as explained above. None of the analyzed studies used synchronization or spontaneous tapping.

As pointed out above, it is particularly important to consider that according to the internal-clock model (Treisman, 1963), an underproduction of duration in time production task goes along with an overestimation of duration in verbal estimation tasks. For this reason, for measures of accuracy (signed error, bisection point), positive values of *g* reported in our analyses always indicate an overestimation of duration in patients relative to control participants in time estimation and bisection tasks, under-production in patients relative to control participants in time production tasks, and under-reproduction in patients relative to control participants in time reproduction tasks. This important aspect was addressed in many studies (e.g., Wahl & Sieg, 1980), but not in the meta-analysis by Ciullo et al. (2015). For measures of precision (difference limen, Weber ratio, coefficient of variation, *d′*, etc.), negative values of *g* reported in our analyses always indicate lower precision in patients relative to control participants.

Most studies reported several pairs of means for the same sample and task. Usually, this was due to testing multiple interval durations or different modalities. In our analysis, we averaged across these multiple values for *g*, and then computed *Var*(*g*) according to Eq. B.5, resulting in one value of *g* and *Var*(*g*) for each combination of sample and task (see Appendix Table A.3). The effect size estimates from the different studies were aggregated according to the discussed differentiations between time perception vs. temporal processing and measures of accuracy vs. precision. Potential effects of schizophrenia on 1) accuracy in time perception tasks, 2) precision in time perception tasks, and 3) precision in temporal processing tasks were analyzed separately by means of random effects meta-regression models (van Houwelingen et al., 2002). In each analysis, the task (as used by each study) was entered as an effect-coded covariate in order to investigate potential task-dependent effects, which might be related to different memory- or motor-related demands (see above). Using the SAS PROC MIXED procedure (Littell et al., 2006) for each analysis, the meta-regression model provided an estimate of the pooled effect size and its confidence interval (fixed effect). The degrees of freedom were calculated according to Kenward and Roger (1997). The model also provided a Type 3 test for the influence of task (the covariate) on the effect of schizophrenia (fixed effect), and least-squares means as estimators of effect size for each task (level of the covariate). These estimates represent predicted population marginal means, based on the estimated fixed-effects parameters (Littell et al., 2006). Finally, the analysis provides an estimate of the between-study variance (van Houwelingen et al., 2002).

In an additional step of analysis, we investigated whether differences between schizophrenic and control participants in time perception may depend on the interval durations that have been used in the studies. Based on the preprocessed data, we grouped the interval durations to four different interval ranges: *ultra-short* (< 1 s), *short* (1 – 10 s), *medium* (10 s – 10 min), and *long* (> 10 min) (for study-specific interval durations see also Appendix Table A.2). Here, we averaged across different tasks used, which provided one value of *g* and - according to Eq.B.5 - one *Var*(*g*) for each combination of sample and interval range (see Appendix Table A.4). We fitted two meta-regression models, one for accuracy in time perception, and one for precision in time perception. In both models, interval range was entered as an effect-coded covariate. Note that interval duration is not defined in tasks measuring temporal processing.

Each analysis was repeated once under consideration of potentially outlying data points. We used regression diagnostics to identify outlying data. Following the recommendations by Viechtbauer and Cheung (2010), we analyzed the externally studentized residuals (also called the studentized deleted residuals) and the DFFITS index proposed by Belsley et al. (1980) as a measure of the influence of a single observation. Following Belsley et al. (1980), externally studentized residuals with an absolute value exceeding 1.96, or with an absolute DFFITS value exceeding $2\sqrt{p/n}$, where *n* is the number of effect sizes analyzed in the model, and *p* is the number of levels of the covariate (task), were defined as outliers.

### 2.2.4 Results

#### *2.2.4.1 Accuracy of time perception*

The single-study effect sizes, task-specific pooled effect sizes, and the overall pooled effect size for accuracy in time perception are displayed in Figure 1. As shown in Table 1, the accuracy of time perception (signed error) of schizophrenic patients did not differ significantly from healthy controls. The lack of a significant effect was particularly obvious when excluding two studies (Carroll et al., 2009a; Lhamon & Goldstone, 1956) identified as outliers according to the criteria stated above.

The Type 3 test indicated a marginally significant influence of the covariate task on the effect of schizophrenia, $F(3,19) = 2.53$, $p = .088$. Based on the outlier-corrected data, this effect reached statistical significance, $F(3,17) = 3.93$, $p = .027$. Least-squares means (as presented in Figure 1 and Table 2; reported for outlier-corrected data) indicated overestimation of duration (underproduction) in time production tasks in schizophrenic participants as compared to control participants. In verbal time estimation, there was also a tendency towards overestimation in schizophrenic participants as compared to control participants, but this effect was only marginally significant. In temporal bisection tasks, there was a non-significant tendency towards the reversed effect. Here, relative to healthy controls, schizophrenic patients tended to classify the duration of a presented comparison interval as *short* more often than as *long*, which could be viewed as an *underestimation* of duration. The estimated effect size for time reproduction was statistically insignificant and close to zero. Note that after outlier correction only one study remained for time reproduction.

In the analysis including time interval range as covariate, the effect of schizophrenia on the accuracy of time perception did not significantly depend on interval range, $F(3,19) = 1.46$, $p = .256$ (outlier corrected: $F(3,18) = 2.61$, $p = .083$). The outlier-corrected least-squares means (Table 3) showed a significant effect of schizophrenia at medium intervals (schizophrenic patients overestimated time intervals), but not at other interval ranges.

#### *2.2.4.2 Precision of time perception*

As seen in Figure 1, the precision of duration judgments was significantly lower in schizophrenic participants than in healthy controls. As shown in Table 1, this effect remained significant after excluding one outlier (Bolbecker et al., 2014), and was large according to the classification of Cohen (1988).

The influence of task (duration discrimination vs. time reproduction) on the effect of schizophrenia on precision in time perception was not significant, $F(1,13.2) = 0.45$, $p = .512$ ($F(1,10.2) = 0.44$, $p = .520$, for outlier-corrected data). Here was also no significant effect of task when differentiating between the different types of duration discrimination tasks (two interval, temporal bisection, one interval reminder, temporal deviant detection) in an additional analysis, $F(5,9.73) = 0.18$, $p = .964$ ($F(5,7.4) = 0.55$, $p = .737$, for outlier-corrected data). As shown in Table 2, least-squares means indicated large effects for each discrimination task, ranging between $\theta = -1.00$ and $-1.99$. Note that for temporal deviant detection, the effect size estimate was based on a single study only.

Regarding the second covariate, the effect of schizophrenia on the precision of time perception did not depend on interval range, $F(1,17) = 0.02$, $p = .887$ (outlier corrected: $F(1,13.5) = 1.37$, $p = .263$). As indicated by the LSMs (Table 3), a significantly negative pooled effect size was obtained at both ranges that had been tested by the studies, short as well as ultra-short, indicating more variable (less precise) duration judgments in schizophrenic patients relative to healthy controls, as in the main analysis.

#### *2.2.4.3 Precision of temporal processing*

As seen in Table 1, the effect of schizophrenia on precision was also significant for temporal processing tasks, and it was similar in size to the effect of schizophrenia on precision in time perception tasks. Compared to the analyses on time perception, the effect sizes for precision of temporal processing were more homogeneous and no outliers were detected.

There was no significant effect of the covariate task, $F(1,2.12) = 0.05$, $p = .850$. For both tasks (judgments of simultaneity and temporal-order judgments), least-squares means (Table 2) were close to the overall estimated effect of schizophrenia on precision of temporal processing ($\theta = -1.05$) in the main analysis.

**Table 1**

Pooled effect size estimates. The table shows the results from the meta-regression models for the influence of schizophrenia on accuracy and precision in time perception (TP) and temporal processing tasks (TPR).

| | *N* ind. samples | *N* effect sizes | θ | $CI_L$ | $CI_U$ | *t* | *df* | *p* | $\tau^2$ | $SE_{\tau^2}$ | $p_{\tau^2}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Accuracy TP** | 17 | 23 | 0.29 | –0.04 | 0.63 | 1.83 | 19 | .084 | 0.32 | 0.13 | .014 |
| **outlier-corr.** | 15 | 21 | 0.04 | –0.37 | 0.45 | 0.20 | 16 | .844 | 0.21 | 0.10 | .037 |
| **Precision TP** | 16 | 16 | **–1.52*** | –2.31 | –0.72 | 4.10 | 13.5 | .001 | 1.80 | 0.73 | .014 |
| **outlier-corr.** | 15 | 15 | **–1.15*** | –1.56 | –0.74 | 6.26 | 10.7 | <.001 | 0.34 | 0.19 | .070 |
| **Precision TPR** | 5 | 6 | **–1.05*** | –1.81 | –0.28 | 5.58 | 2.12 | .003 | 0.08 | 0.17 | .641 |

*Note*. "*N* ind. samples": number of studies/independent samples (i.e., the sample size on level 1). "*N* effect sizes": number of single effect sizes that were entered into the corresponding model (i.e., the sample size on level 2). θ: pooled effect size estimate. $CI_L$ and $CI_U$ are the lower and upper bounds of the 95% confidence interval, respectively, and *t, df,* and *p* refer to a test of θ against 0. $\tau^2$: estimate of the inter-study variance. $SE_{\tau^2}$ : standard error of $\tau^2$. $p_{\tau^2}$ refers to a test of $p_{\tau^2}$ against 0. There were no outlying data points detected for precision TPR. Bold font and * indicates statistically significant results ($p < .05$).

**Table 2**

Estimated effect sizes ($\theta_{LSM}$) for each task for measures of accuracy and precision, as provided by least–squares means computed in the outlier–corrected meta–regression models.

| | Task | Subtasks | *N* studies | $\theta_{LSM}$ | $CI_L$ | $CI_U$ | *t* | *df* | *p* |
|---|---|---|---|---|---|---|---|---|---|
| **Accuracy TP** | Verbal estimation | | 8 | 0.38 | −0.06 | 0.82 | 1.82 | 17.0 | .087 |
| | Time production | | 6 | **0.53*** | 0.03 | 1.03 | 2.25 | 16.3 | .039 |
| | *Time reproduction* | | *1* | *-0.08* | *−1.58* | *1.43* | *0.11* | *17.0* | *.917* |
| | Duration discrimi. (Temp. bisection) | | 6 | -0.50 | -0.99 | 0.00 | 2.12 | 15.9 | .050 |
| **Precision TP** | Time reproduction | | 2 | –0.82 | –1.84 | 0.19 | 1.80 | 10.1 | .102 |
| | Duration discrimi. | | 13 | **–1.15*** | –1.56 | –0.74 | 6.26 | 10.7 | <.001 |
| | | Two interval | 4 | **–1.22*** | –2.07 | –0.38 | 3.33 | 8.1 | .010 |
| | | Temp. bisection | 6 | **–1.00*** | –1.68 | –0.31 | 3.49 | 6.7 | .010 |
| | | One interval reminder | 2 | –1.15 | –2.32 | 0.03 | 2.29 | 7.2 | .055 |
| | | *Temp. deviant detection* | *1* | ***–1.99**** | *–3.68* | *–0.30* | *2.67* | *9.0* | *.026* |
| **Precision TPR** | Judg. of simul. | | 4 | **–1.01*** | –1.87 | –0.14 | 4.40 | 2.3 | .036 |
| | Temp. order. judg. | | 2 | –1.09 | –2.35 | 0.18 | 3.66 | 2.0 | .067 |

*Note*. Categories including data from one single study only are indicated by *italics*. Temporal bisection, one-interval reminder, and temporal deviant detection represent one-interval duration discrimination tasks. Bold font and * indicates statistically significant results ($p < .05$).

**Table 3**

Estimated effect sizes ($\theta_{LSM}$) for each interval range for measures of accuracy and precision, as provided by least–squares means computed in the outlier–corrected meta–regression models.

| | Interval range | *N* studies | $\theta_{LSM}$ | $CI_L$ | $CI_U$ | *t* | *df* | *p* |
|---|---|---|---|---|---|---|---|---|
| **Accuracy** | | | | | | | | |
| | ultra-short | 7 | –0.37 | –0.85 | 0.11 | –1.61 | 18.0 | .125 |
| | short | 6 | 0.10 | –0.41 | 0.61 | 0.41 | 18.0 | .687 |
| | medium | 6 | **0.54*** | 0.04 | 1.05 | 2.25 | 18.0 | .037 |
| | long | 3 | 0.27 | –0.44 | 0.98 | 0.79 | 17.9 | .439 |
| **Precision** | | | | | | | | |
| | ultra-short | 10 | **–0.85*** | –1.25 | –0.45 | –4.56 | 13.6 | <.001 |
| | short | 9 | **–1.17*** | –1.59 | –0.74 | –5.87 | 13.5 | <.001 |

*Note*. ultra-short < 1 s; short 1 – 10 s; medium 10 s – 10 min, large > 10 min; Bold font and * indicates statistically significant results ($p < .05$).

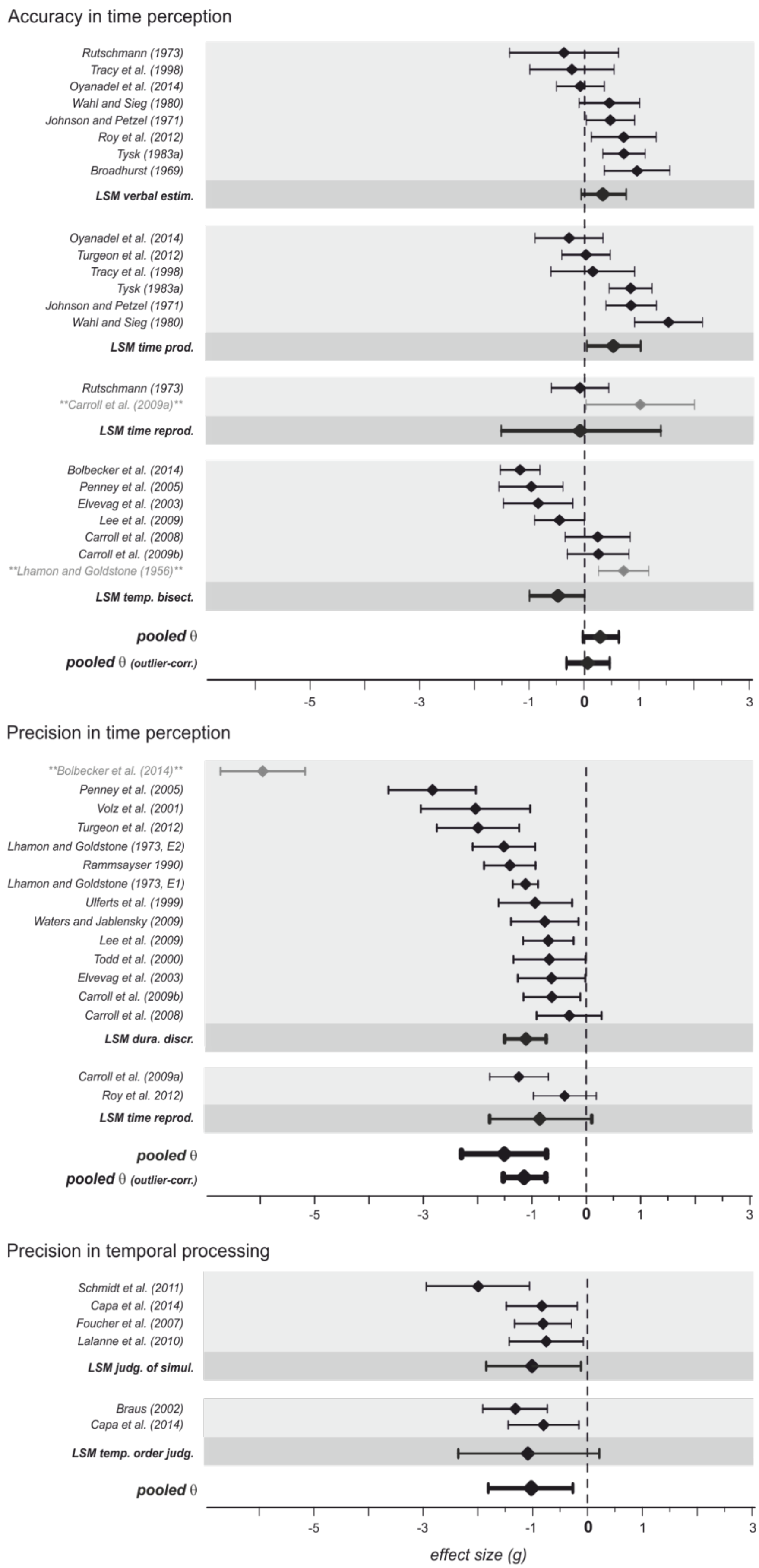
Accuracy in time perception
Rutschmann (1973)
Tracy et al. (1998)
Oyanadel et al. (2014)
Wahl and Sieg (1980)
Johnson and Petzel (1971)
Roy et al. (2012)
Tysk (1983a)
Broadhurst (1969)
LSM verbal estim.
Oyanadel et al. (2014)
Turgeon et al. (2012)
Tracy et al. (1998)
Tysk (1983a)
Johnson and Petzel (1971)
Wahl and Sieg (1980)
LSM time prod.
Rutschmann (1973)
**Carroll et al. (2009a)**
LSM time reprod.
Bolbecker et al. (2014)
Penney et al. (2005)
Elvevag et al. (2003)
Lee et al. (2009)
Carroll et al. (2008)
Carroll et al. (2009b)
**Lhamon and Goldstone (1956)**
LSM temp. bisect.
pooled θ
pooled θ (outlier-corr.)
-5
-3
-1
0
1
3
Precision in time perception
**Bolbecker et al. (2014)**
Penney et al. (2005)
Volz et al. (2001)
Turgeon et al. (2012)
Lhamon and Goldstone (1973, E2)
Rammsayser 1990)
Lhamon and Goldstone (1973, E1)
Ulferts et al. (1999)
Waters and Jablensky (2009)
Lee et al. (2009)
Todd et al. (2000)
Elvevag et al. (2003)
Carroll et al. (2009b)
Carroll et al. (2008)
LSM dura. discr.
Carroll et al. (2009a)
Roy et al. 2012)
LSM time reprod.
pooled θ
pooled θ (outlier-corr.)
-5
-3
-1
0
1
3
Precision in temporal processing
Schmidt et al. (2011)
Capa et al. (2014)
Foucher et al. (2007)
Lalanne et al. (2010)
LSM judg. of simul.
Braus (2002)
Capa et al. (2014)
LSM temp. order judg.
pooled θ
-5
-3
-1
0
1
3
effect size (g)

**Figure 1**

Forest plot showing effect size estimates and corresponding 95% confidence intervals grouped by the three analyzed measures (accuracy in time perception, precision in time perception, and precision in temporal processing). Rows denoted by author and year denote studies, "LSM" denotes least-squares means (marginal means), and "pooled θ" represents pooled θ estimates according to the random-effects meta-regression analyses. Studies identified as outliers are indicated by grey ink and **.

*Note.* For accuracy, positive values of *g* indicate an overestimation of duration in patients relative to control participants in verbal estimation (verbal estim.) and bisection tasks (temp. bisect.), under-production in patients relative to control participants in time production tasks (time prod.), and under-reproduction in patients relative to control participants in time reproduction tasks (time reprod.). For precision, negative values of *g* indicate lower precision in patients relative to control participants. Dura. discr.: duration discrimination (one- and two-interval tasks). Judg. of simul.: judgment of simultaneity. Temp. order judg.: temporal-order judgments.

### 2.2.5 Discussion

Based on a total of 938 schizophrenic and 1041 healthy control participants (29 original studies), we investigated the effects of schizophrenia on the accuracy and precision of time perception and temporal processing.

#### *2.2.5.1 Effects on accuracy of time perception*

Overall, the meta-analysis showed no significant effect on the accuracy of time perception. On average, the signed error of duration judgments did not differ significantly between patients and controls. However, the effect of schizophrenia on the accuracy of time perception depended on the task, at least for outlier-corrected data. Schizophrenic patients overestimated duration in verbal time estimation relative to controls, and showed a tendency to underproduce durations in time production tasks. As explained above, these pattern of effects indicates an accelerated internal clock in patients (Treisman, 1963). A different pattern was observed in temporal bisection tasks. Here, relative to healthy controls, schizophrenic patients tended to underestimate duration. Patients judged the comparison durations as short more often that controls did. Even though the effect in bisection tasks remained statistically insignificant, it is not in accordance with the results obtained for verbal estimation and production tasks. It has been reported earlier that temporal bisection and verbal time estimation / time production tasks may provide diverging results (e.g., Wearden, 2008). In fact, in a bisection task, an alternation of clock speed should affect the memory representations of the anchor durations during the learning phase and the perception of the intervals during the classification phase in a similar way, so that no systematic shift of the bisection point is to be expected. For this reason, the shift of the bisection point could represent a different type of response bias that is not directly linked to differences in clock speed. As the results from verbal time estimation tasks und time production tasks are quite consistent, we conclude that effects on the accuracy of time perception do not depend on motor demands. In time reproduction tasks, patients did not differ from controls. This result was expected, because according to the internal clock model changes in clock speed should not affect the accuracy in time reproduction, as explained above.

Across tasks, the estimated effect of schizophrenia on accuracy did not differ significantly between the four duration ranges (ultra-short < 1 s, short 1 – 10 s, medium 10 s – 10 min, large > 10 min), although descriptively, the overestimation in patients was most pronounced at medium intervals.

*2.2.5.2 Effects on precision of time perception and temporal processing*

In contrast to the ambiguous and rather small effects of schizophrenia on accuracy, the precision of time perception is clearly impaired in schizophrenic patients. Relative to control participants, patients' judgments were significantly more variable. This effect is large across all tasks and duration ranges that have been tested by the studies. As the effect of schizophrenia on precision did not differ between discrimination tasks and reproduction tasks, it cannot be attributed to task-specific (motor) demands. Moreover, the effect size estimates for precision of temporal processing are comparable to those determined for precision of time perception. Hence, irrespective of the specific temporal task and interval duration used, the results from our meta-analysis indicate a general timing deficit in schizophrenia regarding the variability of temporal judgments. In terms of the pacemaker-accumulator model, the internal clock of patients is ticking with larger variability. These results support the notion of generally mistimed information processing in schizophrenic patients that may lead to incorrect connections of thoughts and actions (Andreasen et al., 1999). The results are also in accordance with observations that schizophrenic patients exhibit lower activity in brain areas, such as the striatum, the supplementary motor area, and the insular, that are generally involved in temporal processing and duration judgments (Allman & Meck, 2012; Coull, Cheng, & Meck, 2011; Davalos et al., 2011), although we are not aware of any strong arguments that this should affect the precision rather than the accuracy.

A very robust finding is that the mismatch negativity in EEG responses (e.g., Näätänen, Paavilainen, Rinne, & Alho, 2007) is reduced in schizophrenia (e.g., Umbricht & Krljes, 2005). This has been attributed to changes in the NMDA-type neurotransmitter system (Gil-Da-Costa, Stoner, Fung, & Albright, 2013; Javitt, Steinschneider, Schroeder, & Arezzo, 1996). However, a reduced MMN in schizophrenia is found not only for responses to duration differences, but to a large variety of stimulus differences, although for duration deviants the effects of schizophrenia is particularly strong (Umbricht & Krljes, 2005). Thus, at present it seems unclear whether schizophrenia-related characteristics of the NMDA system play a major role for the reduced precision in time perception and temporal processing.

*2.2.5.3 Potential effects of medication*

It has to be considered that most of the studies included schizophrenic patients who were on medication or in psychotherapy. This might have led to an underestimation of effect sizes. The failure of many studies to provide detailed information concerning the number of subjects under medication, did not allow controlling for possible effects of medication and psychotherapy. As medication and psychotherapy are applied in order to reduce schizophrenic symptoms, possible effects of schizophrenia on time perception and temporal processing might also decline in subjects that are under medication or in psychotherapy. Hence, the inclusion of schizophrenic patients that received some sort of therapy probably led to an underestimation of the pooled effect sizes reported in this meta-analysis. Unfortunately, it was not possible to include medication status as a covariate, because only one study (Braus, 2002) analyzed the effects of schizophrenia on time perception separately for patients on and off medication. The results from this particular study did not indicate differences in performance between patients on and off medication. Penney et al. (2005) investigated time perception abilities in a sample of individuals at high risk of schizophrenia. As none of the participants was actually suffering from schizophrenia, all participants were off medication. Indeed, compared to the other studies that usually tested (partially) medicated patients, the results by Penney et al. (2005) indicated larger effects of schizophrenia on the accuracy and precision in temporal bisection tasks. Participants at genetically high risk of schizophrenia significantly underestimated the duration of the comparison intervals and showed strongly increased variability in their duration judgments. These results support the notion that due to medication artifacts, the effects of schizophrenia might be even larger than determined in our analyses.

Moreover, it has to be noted that medicated schizophrenic patients often receive dopamine antagonists, such as haloperidol (Allman & Meck, 2012). Thus, dopaminergic activity in schizophrenic patients most likely exhibits substantial intra- and inter-individual differences. Evidence from animal as well as human research suggests that decreased dopaminergic activity is related to a deceleration of clock speed (underestimation of time intervals) while increased dopamine levels usually cause opposite effects (e.g., Cheng et al., 2007; Jones & Jahanshahi, 2009; Meck, 1996; Rammsayer, 1993; Wiener et al., 2014). Accordingly, the mixed effects reported for schizophrenia on the accuracy of time perception might be mediated and explained by fluctuating dopamine levels and therefore strongly vary within and between patients. It would be interesting to consider direct measures of dopamine level in future studies in order to shed more light on the role of alterations in neurotransmitter systems in time perception in clinical populations.

#### *2.2.5.4 Limitations and recommendations for future research*

A general issue for this meta-analysis was the heterogeneity of diagnostic criteria (DSM, ICD, Brief Psychiatric Rating Scale, etc.) and their modifications over time (e.g., DSM-II vs. DSM-III vs. DSM-IV). This factor might have led to inconsistencies in group assignment between the different studies. For instance, some subjects that were assigned to the control group in one study might have been classified as schizophrenic patients in another study that used different criteria for group assignment.

Moreover, it was not possible to systematically differentiate between different subtypes of schizophrenia. Only one study (Schmidt et al., 2011) analyzed data separately for chronic and first episode schizophrenics, for example, reporting no differences between subgroups in the impaired precision of temporal processing.

Another interesting methodological aspect concerns verbal time estimation tasks. As pointed out above, it is crucial to differentiate between prospective and retrospective judgments (cf. Grondin, 2010). This has not been done systematically in the context of schizophrenia yet. In the prospective paradigm, the participant is informed about the temporal task and focuses attention on time. In the retrospective paradigm, the subject (being uninformed about the temporal task) does not focus attention on time. In contrast to prospective judgments, the task to give a retrospective estimate is less attention-related and more memory demanding (e.g., S. W. Brown, 2008; Grondin, 2010). There is evidence for systematic differences between prospective and retrospective time estimation (for a meta-analytic review see Block & Zakay, 1997), with prospective judgments being longer and less variable than retrospective judgments. Regarding possible effects of schizophrenia on time estimation, prospective and retrospective judgments might be affected differently, indicating whether attentional (prospective) processes or/and memory-related (retrospective) processes of time perception are altered in schizophrenia. Due to the fact that only few studies with verbal estimation tasks could be included in this meta-analysis, additional covariates, as for example the estimation paradigm (prospective vs. retrospective), could not be considered in the analysis. Inspection of the effect sizes (see Figure 1) did not indicate substantial differences between results from studies including retrospective judgments (Oyanadel & Buela-Casal, 2014; Roy et al., 2012) and those including prospective judgments (Broadhurst, 1969; Johnson & Petzel, 1971; Rutschmann, 1973; Tracy et al., 1998). However, the two studies that actually tested both types of judgments within the same sample (Tysk, 1983b; Wahl & Sieg, 1980) reported somewhat larger effects (more overestimation in patients) in prospective relative to retrospective tasks. This pattern of results might indicate less memory- but rather attention-related impairments of time perception in schizophrenic patients – a result that corresponds to the larger within-subject variability (impaired precision) in patients. In order to systematically address this issue, future studies need to directly compare patients' performance in prospective versus retrospective time estimation tasks.

Unfortunately, our review revealed significant shortcomings in experimental design and statistical analysis in a relatively large number of original studies. Future studies should use the established psychophysical tasks for measuring time perception or temporal processing (cf. Grondin, 2008; Grondin, 2010), and analyze the data with adequate and established methods. For instance, data from one-interval discrimination tasks provide information about precision (in terms of the difference limen that can be estimated from the psychometric function), but also about accuracy (in terms of the 50% point on the psychometric function). This information is lost if the data are analyzed in terms of percent-correct rather than by fitting a psychometric function (cf. Treutwein & Strasburger, 1999), as explained in the introduction. As shown in Appendix Table A.1, of a total of 4 studies using one-interval duration discrimination tasks could not be included in the meta-analysis due to this issue. Another less severe issue concerns the analysis of data from two-interval discrimination tasks using a transformed up-down adaptive procedure (Levitt, 1971). Here, in most studies on effects of schizophrenia, the duration difference limen was computed as the average of the duration differences between comparison and standard interval across for example the last 20 trials. The correct analysis is to compute the average duration differences between comparison and standard interval across for example the last 6 so-called reversals, which are trials on which the direction of the adaptive track changes (Levitt, 1971). If instead the average across trials is used, this results in a general overestimation of the duration difference limen.

#### *2.2.5.5 Conclusion*

Taken together, the results of our meta-analysis indicate that the precision of time perception and temporal processing is substantially impaired in schizophrenic patients. Thus, patients have a lower sensitivity in judging time intervals, and in more basic temporal tasks like temporal-order judgments. These effects of schizophrenia on precision do not significantly depend on interval durations used, and do not differ substantially between tasks that are comprised of purely perceptual judgments and those requiring timed motor responses. In contrast, the accuracy of time perception in the sense of a systematic deviation of time estimates form the veridical value is not significantly affected by schizophrenia across the different tasks and duration ranges. However, it may depend on specific task-related cognitive demands (e.g., memory), because a small to moderate effect was observed in verbal estimation and time production. Notwithstanding this qualification, the effect of schizophrenia on the accuracy of time perception is substantially smaller than the effect on precision.

Our review also shows that several aspects have not been investigated in sufficient detail yet and might be addressed by future research. These aspects include the role of specific task characteristics like prospective versus retrospective time estimation as well as potentially mediating effects of medication and neurotransmitter levels. The effect sizes provided by our meta-analysis may be used for selecting appropriate sample sizes in future experiments. These studies should carefully differentiate between measures of mean duration estimates (accuracy) on the one hand and measures of variability (precision) on the other hand, and use established experimental task and analysis procedures.

### 2.2.6 Acknowledgments

We are grateful to Sarah Cardoso and Nadia Kirsch for their help with the literature search, and to Marius Frenken for double-checking the transfer of the statistical information from the studies to our analyses.

## 2.2.7 Appendices

### *2.2.7.1 Table A.1*

Results of the literature search in alphabetical order. Studies were considered for further analyses only if all four inclusion criteria were met.

| Study | Empirical study? | Relevant tasks? | Patient and control group? | Data interpretable in terms of accuracy and / or precision? | Sufficient statistical information? | Authors' response in case of data request | Inclusion? |
|---|---|---|---|---|---|---|---|
| Allman and Meck (2012) | no (review) | | | | | | no |
| Bolbecker, et al. (2014) | yes | yes (duration discr.; temp. bisection) | yes | yes (A and P) | yes | | yes |
| Bonnot and Georgieff (2000) | no (review) | | | | | | no |
| Bonnot et al. (2011) | no (review) | | | | | | no |
| Bourdet, et al. (2003) | yes | yes (duration discr.; deviant detection) | yes | no (duration DL defined by an unspecific performance level, range between 70-90% correct) | yes | | no |
| Braus (2002) | yes | yes (ToJ) | yes (two groups of schizophrenic patients: medicated vs. unmedicated vs. control) | yes (P) | yes | | yes |
| Broadhurst (1969) | yes | yes (verbal time estimation) | yes | yes (A) | yes | | yes |

| Study | Empirical study? | Relevant tasks? | Patient and control group? | Data interpretable in terms of accuracy and / or precision? | Sufficient statistical information? | Authors' response in case of data request | Inclusion? |
|---|---|---|---|---|---|---|---|
| Capa, et al. (2014) | yes | yes (ToJ; JoS) | yes | yes (P) | yes (*F*-values and *N*s are reported) | | yes |
| Carlson and Feinberg (1968) | yes | yes (verbal time estimation; time production; time reproduction) | yes | no (Linear regression analysis of time estimates/productions across a wider range of intervals, thus not possible to distinguish between duration ranges to time intervals. No distinct measures of accuracy and precision.) | no (no tests of intercept and slope measures between groups) | | no |
| Carroll, et al. (2008) | yes | yes (duration discr.; temp. bisection) | yes | yes (A and P) | yes | | yes |
| Carroll, et al. (2009a) | yes | yes (continuation tapping) | yes | yes (A and P) | yes | | yes |
| Carroll, et al. (2009b) | yes | yes (temp. bisection) | yes | yes (A and P) | yes | | yes |
| Clausen (1950) | yes | yes (verbal time estimation; time production; time reproduction) | no (no healthy control group; reported data: before and after removal of frontal lobes) | yes (A) | no (no *SD*s or *F* / *t*-values reported) | | no |
| Davalos, et al. (2002) | yes | yes (duration discri.; one interval reminder) | yes | no (only errors rates for each comparison interval reported) | yes | data requested; no response | no |

| Study | Empirical study? | Relevant tasks? | Patient and control group? | Data interpretable in terms of accuracy and / or precision? | Sufficient statistical information? | Authors’ response in case of data request | Inclusion? |
|---|---|---|---|---|---|---|---|
| Davalos, Kisley, and Ross (2003) | yes | yes (duration discri.; one interval reminder) | yes | no (only error rates reported) | yes | data requested; no response | no |
| Davalos, et al. (2005) | yes | yes (duration discri.;one interval reminder) | yes | no (only error rates reported) | yes | data requested; no response | no |
| Davalos, et al. (2011) | yes | yes (duration discri.; one interval reminder) | yes | no (only error rates reported) | yes | data requested; no response | no |
| Delevoye, et al. (2013) | yes | no (spontaneous finger tapping at preferred tempo) | | | | | no |
| Densen (1977) | yes | yes (verbal time estimation) | yes | yes (A) | no (no *SD*s or *F* / *t*-values reported) | | no |
| Dilling and Rabin (1967) | yes | yes (verbal time estimation) | yes | no (only frequency distributions of over- and underestimation reported) | no (only medians and sum of ranks reported) | | no |
| Droit-Volet (2013) | no (review) | | | | | | no |
| Elvevag, et al. (2003) | yes | yes (duration discri.; temp. bisection) | yes | yes (A and P) | yes | | yes |
| Elvevag, Brown, McCormack, Vousden, and Goldberg (2004) | yes | yes (duration discri.; absolute identification task with 7 durations) | yes | no (only mean error rates reported) | yes | data requested but not received | no |
| Foucher, et al. (2007) | yes | yes (JoS) | yes | yes (P) | yes | | yes |

| Study | Empirical study? | Relevant tasks? | Patient and control group? | Data interpretable in terms of accuracy and / or precision? | Sufficient statistical information? | Authors' response in case of data request | Inclusion? |
|---|---|---|---|---|---|---|---|
| Franck, Posada, Pichon, and Haggard (2005) | yes | no ("timing judgment by Haggard") | | | | | no |
| Grondin et al. (2006) | no (review) | | | | | | no |
| Jirsa, et al. (1996) | yes | no (40 Hz quanta synchronization finger tapping) | | | | | no |
| Johnson and Petzel (1971) | yes | yes (verbal time estimation; time production) | yes | yes (A) | yes | | yes |
| Lalanne, et al. (2010) | yes | yes (JoS) | yes | yes (P) | yes | | yes |
| Lee, et al. (2006) | yes | yes (duration discri.; temp. bisection) | no (schizotypy as continuous variable) | | | | no |
| Lee, et al. (2009) | yes | yes (duration discri.; temp. bisection) | yes | yes (A and P) | yes | | yes |
| Lhamon and Goldstone (1956) | yes | yes (duration discri.; temp. bisection) | yes | yes (A and P) | yes | | yes |
| Lhamon and Goldstone (1973) Exp. 1 and 2 | yes | yes (duration discri.; temp. bisection; one interval reminder) | yes | yes (P) | yes | | yes |

| Study | Empirical study? | Relevant tasks? | Patient and control group? | Data interpretable in terms of accuracy and / or precision? | Sufficient statistical information? | Authors' response in case of data request | Inclusion? |
|---|---|---|---|---|---|---|---|
| Martin, et al. (2013) | yes | yes (JoS) | yes | no (data were only partially reported in terms of P and A) | yes | data requested and received, but only data from 50% of the subjects were analyzed in a way appropriate for our analyses | no |
| Martin, et al. (2014) | no (review) | | | | | | no |
| Martinez-Cascales, de la Fuente, Santiago, and Santiago (2013) | yes (conference contribution) | | | | no (missing statistics) | | no |
| Meck (2005) | no (review) | | | | | | no |
| Mishara and Gallistel (2005) | yes (conference contribution) | no | | | | | no |
| Nenadic et al. (2000) | yes (conference contribution) | yes (verbal time estimation) | yes | no | no (no statistics, no behavioral results) | | no |
| Nichols and Park (2011) | yes (conference contribution) | | | | no | | no |
| Nosachev (1991) | yes | yes (time production) | yes | yes (A) | no (distribution measure unclear) | | no |
| Nosachev (1992) | yes | yes (time production) | yes | yes (A) | no (statistical measures not explained in the paper) | | no |

| Study | Empirical study? | Relevant tasks? | Patient and control group? | Data interpretable in terms of accuracy and / or precision? | Sufficient statistical information? | Authors' response in case of data request | Inclusion? |
|---|---|---|---|---|---|---|---|
| Omre (1966) | yes | yes (verbal time estimation) | yes | yes (A) | no (no means or *F*/ *t*-values reported) | | no |
| Oyanadel and Buela-Casal (2014) | yes | yes (verbal time estimation) | yes | yes (A) | yes | | yes |
| Papageorgiou, et al. (2013) | yes | yes (duration discri.; one interval reminder) | yes | yes (A and P) | no (no *SD*s or *F* / *t*-values reported) | data requested but not received | no |
| Parsons et al. (2013) | yes | no (flicker fusion) | | | | | no |
| Pearl and Berg (1963) | yes | yes (verbal time estimation) | no | yes (A) | yes | | no |
| Penney, et al. (2005) | yes | yes (duration discr.; temp. bisection) | group of subjects at high genetic risk for schizophrenia vs control group | yes (A and P) | yes | | yes |
| Peterburs, Nitsch, Miltner, and Straube (2013) | yes | no (time to contact task estimation) | | | | | no |
| Rabin (1957) | yes | yes (verbal time estimation) | yes | no (frequency distributions of categories of overestimation, under estimation and correct estimation) | no | | yo |
| Rammsayer (1990) | yes | yes (duration discri.; two interval) | yes | yes (A and P) | yes | | yes |
| Ratcliffe (2012) | no (review) | | | | | | no |

| Study | Empirical study? | Relevant tasks? | Patient and control group? | Data interpretable in terms of accuracy and / or precision? | Sufficient statistical information? | Authors’ response in case of data request | Inclusion? |
|---|---|---|---|---|---|---|---|
| Roy, et al. (2012) | yes | yes (verbal time estimation, time reproduction) | yes | yes (A and P) | yes | | yes |
| Rutschmann, et al. (1973) | yes | yes (verbal time estimation, time reproduction) | yes | yes (A) | yes | | yes |
| Schmidt, et al. (2011) | yes | yes (JoS) | yes (data from chronic vs. first episode schizophrenic patients vs. control group) | yes (P) | yes | | yes |
| Teixeira et al. (2013) | no (review) | | | | | | no |
| Todd, et al. (2000) | yes | yes (duration discri.; two interval) | yes | yes (P) | yes | | yes |
| Todd, et al. (2003) | yes | yes (duration discri.; two interval) | yes | yes (P) | no | data requested, but not received | no |
| Tracy, et al. (1998) | yes | yes (time reproduction) | yes | yes (A) | yes | | yes |
| Turgeon, et al (2012) | yes | yes ( duration discri.; deviant detection; time production) | yes | yes (A and P) | yes | | yes |

| Study | Empirical study? | Relevant tasks? | Patient and control group? | Data interpretable in terms of accuracy and / or precision? | Sufficient statistical information? | Authors' response in case of data request | Inclusion? |
|---|---|---|---|---|---|---|---|
| Tysk (1983a) | yes (however, partially reported data from Tysk (1983b) | | | | | | |
| Tysk (1983b) | yes | yes (verbal time estimation; time production) | yes | yes (A) | yes | | yes |
| Ulferts, et al. (1999) | yes | yes (duration discri.; two interval) | yes (data from differently medicated groups of schizophrenic patients vs. control group) | yes (P) | yes | | yes |
| Van der Veen (2013) | yes | yes (time production) | yes | no (only percentage of correct responses is reported) | yes | data requested and received, but not analyzable in terms of precision and / or accuracy | no |
| Volz, et al. (2001) | yes | yes (duration discri. two interval) | yes | yes (P) | yes | | yes |
| Wahl and Sieg (1980) | yes | yes (verbal time estimation; time production) | yes | yes (A) | yes | | yes |
| Waters and Jablensky (2009) | yes | yes (duration discri. two interval) | yes (patients with first rank symptoms vs. patients without first rank symptoms vs. control group) | yes (P) | yes | | yes |

| Study | Empirical study? | Relevant tasks? | Patient and control group? | Data interpretable in terms of accuracy and / or precision? | Sufficient statistical information? | Authors' response in case of data request | Inclusion? |
|---|---|---|---|---|---|---|---|
| Yang, et al. (2004) | yes | yes (finger tapping) | no | | | | no |

*Note.* ToJ = Temporal-order judgment; JoS = Judgment of simultaneity. Studies excluded from the meta-analysis are indicated by grey ink.

### 2.2.7.2 *Table A.2*

Diagnostic criteria and demographic information on schizophrenic and control subjects as well as information on covered tasks, dependent measures, modality, and presented interval durations for each study included in the meta-analysis.

| Study | Diagnostics | Schizophrenics | | | | | Controls | | | | | Task | Accuracy measure | Precision measure | Modality | Interval duration (range) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Age | | Gender | | | Age | | Gender | | | | | | | |
| | | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | | | | | |
| Bolbecker, et al.(2014) | DSM-IV | | | 41 | 25 | 66 | | | 29 | 44 | 73 | Duration Discrimi. (temp. bisection) | BP | DL | auditory | <1 s; u-short |
| Braus (2002) | Brief Psychiatric Rating Scale | 34.9 | 9.1 | 21 | 14 | 35 | 31.4 | 8.1 | 13 | 13 | 26 | *ToJ* | | *SOA at 75% correct* | *auditory; visual* | |
| Broadhurst (1969) | not specified | 39.6 | 12.1 | 24 | 0 | 24 | 33.5 | 10.0 | 24 | 0 | 24 | Verbal Estim. (p) | Signed error | | not specified | 5 min; medium |

| Study | Diagnostics | Schizophrenics | | | | | Controls | | | | | Task | Accuracy measure | Precision measure | Modality | Interval duration (range) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Age | | Gender | | | Age | | Gender | | | | | | | |
| | | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | | | | | |
| Capa, et al. (2014) | DSM-IV | 37.2 | 9.2 | 14 | 6 | 20 | 33.4 | 11.4 | 14 | 6 | 20 | *ToJ* | | *SOA at 75% correct* | *visual* | |
| | | | | | | | | | | | | *JoS* | | *Inter stimulus interval at 50% 'sim.' –responses* | *visual* | |
| Carroll, et al. (2008) | DSM-IV | 39.7 | 11.0 | 15 | 8 | 23 | 41.1 | 12.1 | 6 | 16 | 22 | Duration Discrimi. (temp. bisection) | BP | WR [DL/BP] | auditory; visual | <1 s u-short |
| Carroll, et al. (2009a) | DSM-IV | 39.1 | 10.4 | 22 | 10 | 32 | 40.1 | 11.2 | 9 | 22 | 31 | Time Reprod. [Finger Tapping (contin.)] | Inter tapping interval (Signed error) | CV of inter tapping interval | auditory | <1 s; u-short |
| Carroll, et al. (2009b) | DSM-IV | 39.1 | 10.2 | 21 | 7 | 28 | 39.5 | 10.5 | 11 | 20 | 31 | Duration Discrimi. (temp. bisection) | BP | WR [DL/BP] | auditory; visual | <1 s u-short |
| Elvevag, et al. (2003) | DSM-IV | 32.8 | 8.4 | 15 | 4 | 19 | 29.4 | 11.5 | 6 | 17 | 23 | Duration Discrimi.; (temp. bisection) | BP | WR [DL/BP] | auditory; visual | <1 s; u-short 4.5 s; short |

| **Study** | **Diagnostics** | **Schizophrenics** | | | | | **Controls** | | | | | **Task** | **Accuracy measure** | **Precision measure** | **Modality** | **Interval duration (range)** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **Age** | | **Gender** | | | **Age** | | **Gender** | | | | | | | |
| | | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | | | | | |
| Foucher, et al. (2007) | DSM-IV | 33.0 | 9.0 | 21 | 9 | 30 | 32.0 | 11.0 | 22 | 11 | 33 | *JoS* | | *Inter stimulus interval at 50% 'sim.' –responses* | *auditory; visual* | |
| Johnson and Petzel (1971) | not specified | 48.4 | | 20 | 20 | 40 | 46.5 | | 20 | 20 | 40 | Time Production<br>Verbal Estim. (p) | Signed error<br>Signed error | | not specified<br>not specified | 30 s; medium<br>30 s; medium |
| Lalanne, et al. (2010) | DSM-IV | 35.7 | 6.3 | 9 | 9 | 18 | 34.3 | 6.4 | 9 | 9 | 18 | *JoS* | | *Inter stimulus interval at 50% 'sim.' –responses* | *visual* | |
| Lee, et al. (2009) | DSM-IV | 37.3 | 10.4 | 34 | 4 | 38 | 35.5 | 10.7 | 34 | 4 | 38 | Duration Discrimi. (temp. bisection) | BP | DL | auditory | ~1 s; short |
| Lhamon and Goldstone (1956) | not specified | 33 | | 34 | 3 | 37 | 27 | | 31 | 10 | 41 | Duration Discrimi. (temp. bisection) | BP | | auditory | 1 s; short |
| Lhamon and Goldstone (1973) Exp.1 | not specified | | | 160 | 0 | 160 | | | 103 | 57 | 160 | Duration Discrimi. (temp. bisection) | | 'Trans-mitted in-formation" | auditory; visual | 1 s; short |

| Study | Diagnostics | Schizophrenics | | | | | Controls | | | | | Task | Accuracy measure | Precision measure | Modality | Interval duration (range) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Age | | Gender | | | Age | | Gender | | | | | | | |
| | | *M* | *SD* | *N* Ma | *N* Fe | *N* | *M* | *SD* | *N* Ma | *N* Fe | *N* | | | | | |
| Lhamon and Goldstone (1973) Exp.2 | not specified | | | 17 | 13 | 30 | | | 17 | 13 | 30 | Duration Discrimi. (one interval reminder) | | 'Trans-mitted information" | auditory; visual | 1 s; short |
| Oyanadel, et al. (2014) | DSM-IV | 40.6 | 10.6 | 19 | 11 | 30 | 35.5 | 9.6 | 38 | 22 | 60 | Time Production | Signed error | | NA | 35 s; medium |
| | | | | | | | | | | | | Verbal Estim. (r) | Signed error | | NA | 35 s; medium |
| Penney, et al. (2005) | none (HR Schizos) | 25.5 | 2.1 | 12 | 5 | 17 | 25.9 | 2.2 | 20 | 14 | 34 | Duration Discrimi. (temp. bisection) | BP | DL | auditory; visual | 4.5 s; short |
| Rammsayer (1990) | DSM-III | 31.5 | 13.2 | 14 | 13 | 27 | 31.4 | 10.9 | 36 | 44 | 80 | Duration Discrimi. (two interval, adaptive) | | DL | auditory | <1 s; short |
| Roy, et al. (2012) | DSM-IV | 25.7 | 6.3 | 24 | 1 | 25 | 25.7 | 4.3 | 24 | 1 | 25 | Time Reprod. | | CV | auditory | 1.6 s; short |
| | | | | | | | | | | | | Verbal Estim. (r) | Signed error | | NA | ~40 min; large |
| Rutschmann (1973) | DSM-II | 21.2 | 2.4 | 7 | 0 | 7 | 23.8 | 0.8 | 9 | 0 | 9 | Time Reprod. | Signed error | | auditory | 1.3 s; short |
| | | | | | | | | | | | | Verbal Estim. (p) | Signed error | | auditory | 1.3 s; short |

| Study | Diagnostics | Schizophrenics | | | | | Controls | | | | | Task | Accuracy measure | Precision measure | Modality | Interval duration (range) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Age | | Gender | | | Age | | Gender | | | | | | | |
| | | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | | | | | |
| Schmidt, et al. (2011) | DSM-IV | 29.9 | 8.6 | 12 | 8 | 20 | 28.8 | 8.9 | 9 | 2 | 11 | *JoS* | | *Inter stimulus interval at 50% 'sim.' –responses* | *visual* | |
| Todd, et al. (2000) | ICD-10; DSM-IV | 28.3 | | 15 | 2 | 17 | 26.6 | | 18 | 2 | 20 | Duration Discrimi. (two interval, adaptive) | | DL | auditory | <1 s; u-short |
| Tracy, et al. (1998) | DSM-III | 47.3 | 10.1 | 9 | 10 | 19 | 26.6 | 11.0 | 14 | 29 | 43 | Verbal Estim. (p) | Abs. error and signed error | | not specified | 24 s; medium |
| | | | | | | | | | | | | Time Production | Abs. error and signed error | | NA | 24 s; medium |
| Turgeon, et al. (2012) | DSM-V | 39.2 | 9.3 | 14 | 6 | 20 | 39.2 | 14.0 | 12 | 8 | 20 | Time Production | Signed error | DL | NA | <1 s; u-short |
| | | | | | | | | | | | | Duration Discrimi. (deviant detection) | | | auditory | <1 s; u-short |

| Study | Diagnostics | Schizophrenics | | | | | Controls | | | | | Task | Accuracy measure | Precision measure | Modality | Interval duration (range) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Age | | Gender | | | Age | | Gender | | | | | | | |
| | | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | ***M*** | ***SD*** | ***N* Ma** | ***N* Fe** | ***N*** | | | | | |
| Tysk (1983a) | DSM-III | 34.8 | 11.1 | 30 | 20 | 50 | 37.6 | 10.6 | 26. | 34 | 60 | Time Production<br>Verbal Estim. (p)<br>Verbal Estim. (r) | Signed error<br>Signed error<br>Signed error | | not specified<br>not specified<br>NA | 20 s; medium<br>17.5 s; medium<br>7.5 min; medium |
| Ulferts, et al. (1999) | ICD-9 | 31.0 | 7.2 | 21 | 15 | 36 | 28.1 | 4.6 | 7 | 5 | 12 | Duration Discrimi. (two interval, adaptive) | | DL | auditory | <1 s; u-short |
| Volz, et al. (2001) | DSM-IV; ICD-10 | 31.7 | 12.1 | 9 | 0 | 9 | 25.3 | 3.6 | 15 | 0 | 15 | Duration Discrimi. (two interval, adaptive) | | DL | auditory | 1 s; short |
| Wahl and Sieg (1980) | not specified | | | 12 | 14 | 26 | | | 19 | 7 | 26 | Time Production<br>Verbal Estim. (p)<br>Verbal Estim. (r) | Signed error<br>Signed error<br>Signed error | | NA<br>not specified<br>NA | 30 s; medium<br>30 s; medium<br>~30 min; large |
| Waters and Jablensky (2009) | DSM-IV; ICD-10 | 34.8 | 9.3 | 35 | 0 | 35 | 44.0 | 9.9 | 16 | 0 | 16 | Duration Discrimi. (one interval reminder) | | *d′* | auditory | 1.2 s; short |

*Note.* Empty cells represent missing information in studies. Data in quotation marks refer to statements from the studies. Estim. = Estimation; (r) = retrospective; (p) = prospective; Reprod. = Reproduction; Discrimi = Discrimination; Ma = male; Fe = Female; *M* = Mean; *SD* = Standard deviation; *N* = sample size; DM = Dependent measure; HR Schizos = Subjects with high genetic risk for schizophrenia; Temp. = Temporal; BP = Bisection point (Point of subjective equality); DL = Difference limen; WR = Weber ratio; CV = Coefficient of variation; contin. = Continuation; abs. = absolute; SI = Standard interval; corr. resp. = Correct responses; *ToJ = Temporal-order judgment*; *JoS = Judgment of simultaneity*; NA = not applicable; *Italics* indicate tasks and measures referring to *temporal processing*; "Interval duration" not definable for temporal processing tasks.

### *2.2.7.3 Table A.3*

Averaged effect sizes (*g*), effect size variances (Var(*g*)), corresponding 95% confidence intervals ($CI_{lower}$ and $CI_{upper}$), and number of reported means (*J*) per combination of sample and *task*. *J* denotes the number of effects sizes on which the averaged *g* is based on.

| Study | Task | *J* | *g* | Var(*g*) | $CI_{lower}$ | $CI_{upper}$ |
|---|---|---|---|---|---|---|
| *Accuracy in time perception* | | | | | | |
| Rutschmann (1973) | verbal estimation | 7 | –0.369 | 0.258 | –1.365 | 0.627 |
| Tracy, et al. (1998) | verbal estimation | 2 | –0.224 | 0.150 | –0.983 | 0.535 |
| Oyanadel, et al. (2014) | verbal estimation | 1 | –0.070 | 0.050 | –0.508 | 0.368 |
| Wahl and Sieg (1980) | verbal estimation | 2 | 0.459 | 0.079 | –0.092 | 1.010 |
| Johnson and Petzel (1971) | verbal estimation | 1 | 0.478 | 0.051 | 0.035 | 0.921 |
| Roy, et al. (2012) | verbal estimation | 1 | 0.721 | 0.091 | 0.130 | 1.312 |
| Tysk (1983a) | verbal estimation | 4 | 0.726 | 0.039 | 0.339 | 1.113 |
| Broadhurst (1969) | verbal estimation | 2 | 0.965 | 0.093 | 0.367 | 1.563 |
| Oyanadel, et al. (2014) | time production | 2 | –0.276 | 0.050 | –0.714 | 0.162 |
| Turgeon, et al. (2012) | time production | 1 | 0.032 | 0.100 | –0.588 | 0.652 |
| Tracy, et al. (1998) | time production | 2 | 0.155 | 0.149 | –0.602 | 0.912 |
| Tysk (1983a) | time production | 3 | 0.849 | 0.040 | 0.457 | 1.241 |
| Johnson and Petzel (1971) | time production | 1 | 0.856 | 0.055 | 0.396 | 1.316 |
| Wahl and Sieg (1980) | time production | 1 | 1.540 | 0.100 | 0.920 | 2.160 |
| Rutschmann (1973) | time reproduction | 7 | –0.076 | 0.254 | –1.064 | 0.912 |
| Carroll, et al. (2009a) | time reproduction | 1 | 1.024 | 0.072 | 0.498 | 1.550 |
| Bolbecker, et al. (2014) | temporal bisection | 1 | –1.172 | 0.034 | –1.533 | –0.811 |
| Penney, et al. (2005) | temporal bisection | 2 | –0.966 | 0.097 | –1.576 | –0.356 |
| Elvevag, et al. (2003) | temporal bisection | 1 | –0.843 | 0.105 | –1.478 | –0.208 |
| Lee, et al. (2009) | temporal bisection | 2 | –0.453 | 0.054 | –0.908 | 0.002 |
| Carroll, et al. (2008) | temporal bisection | 2 | 0.246 | 0.090 | –0.342 | 0.834 |
| Carroll, et al. (2009b) | temporal bisection | 2 | 0.261 | 0.069 | –0.254 | 0.776 |
| Lhamon and Goldstone (1956) | temporal bisection | 2 | 0.721 | 0.055 | 0.261 | 1.181 |
| *Precision in time perception* | | | | | | |
| Bolbecker, et al. (2014) | temporal bisection | 1 | –5.952 | 0.156 | –6.730 | –5.178 |
| Penney, et al. (2005) | temporal bisection | 2 | –2.828 | 0.166 | –3.627 | –2.029 |
| Volz, et al. (2001) | two interval | 1 | –2.036 | 0.264 | –3.043 | –1.029 |
| Turgeon, et al. (2012) | temp. deviant detec. | 1 | –1.994 | 0.150 | –2.753 | –1.235 |
| Lhamon and Goldstone (1973E2) | one interv. reminder | 1 | –1.517 | 0.086 | –2.092 | –0.942 |
| Rammsayser 1990) | two interval | 1 | –1.406 | 0.059 | –1.882 | –0.930 |
| Lhamon and Goldstone (1973E1) | temporal bisection | 1 | –1.118 | 0.014 | –1.350 | –0.886 |
| Ulferts, et al. (1999) | two interval | 2 | –0.936 | 0.120 | –1.615 | –0.257 |
| Waters and Jablensky (2009) | one interv. reminder | 1 | –0.763 | 0.100 | –1.383 | –0.143 |
| Lee, et al. (2009) | temporal bisection | 2 | –0.697 | 0.056 | –1.161 | –0.233 |
| Todd, et al. (2000) | two interval | 1 | –0.675 | 0.115 | –1.340 | –0.010 |
| Elvevag, et al. (2003) | temporal bisection | 1 | –0.639 | 0.101 | –1.262 | –0.016 |
| Carroll, et al. (2009b) | temporal bisection | 2 | –0.631 | 0.071 | –1.153 | –0.109 |
| Carroll, et al. (2008) | temporal bisection | 2 | –0.312 | 0.090 | –0.900 | 0.276 |

| Study | Task | *J* | *g* | Var(*g*) | $CI_{lower}$ | $CI_{upper}$ |
|---|---|---|---|---|---|---|
| Carroll, et al. (2009a) | time reproduction | 1 | –1.238 | 0.076 | –1.778 | –0.698 |
| Roy, et al. 2012) | time reproduction | 3 | –0.394 | 0.087 | –0.972 | 0.184 |
| *Precision in temporal processing* | | | | | | |
| Schmidt, et al. (2011) | simultaneity judg. | 1 | –2.002 | 0.206 | –2.892 | –1.112 |
| Capa, et al. (2014) | simultaneity judg. | 1 | –0.834 | 0.109 | –1.481 | –0.187 |
| Foucher, et al. (2007) | simultaneity judg. | 3 | –0.812 | 0.069 | –1.327 | –0.297 |
| Lalanne, et al. (2010) | simultaneity judg. | 1 | –0.751 | 0.119 | –1.427 | –0.075 |
| Braus (2002) | temporal-order judg. | 2 | –1.330 | 0.082 | –1.891 | –0.769 |
| Capa, et al. (2014) | temporal-order judg. | 1 | –0.801 | 0.108 | –1.445 | –0.157 |

*Note.* In the analysis, continuation tapping has been regarded as time reproduction. Two interval, temporal bisection, one interval reminder tasks, and temporal deviant detection are grouped as duration discrimination tasks in the main analyses. Temporal bisection is the only duration discrimination task that provides information on accuracy of time perception (bisection point). Outlying data are indicated by grey ink.

### 2.2.7.4 Table A.4

Averaged effect sizes (*g*), effect size variances (Var(*g*)), corresponding 95% confidence intervals ($CI_{lower}$ and $CI_{upper}$), and number of reported means (*J*) per combination of sample and *interval range*. *J* denotes the number of effects sizes on which the averaged *g* is based on.

| Study | Interval range | J | g | Var(g) | $CI_{lower}$ | $CI_{upper}$ |
|---|---|---|---|---|---|---|
| *Accuracy in time perception* | | | | | | |
| Oyanadel and Buela-Casal (2014) | long | 1 | –0.070 | 0.050 | –0.509 | 0.368 |
| Wahl and Sieg (1980) | long | 1 | 0.200 | 0.077 | –0.345 | 0.745 |
| Roy, et al. (2012) | long | 1 | 0.721 | 0.091 | 0.131 | 1.312 |
| Oyanadel and Buela-Casal (2014) | medium | 2 | –0.276 | 0.050 | –0.714 | 0.162 |
| Tracy, et al. (1998) | medium | 4 | –0.034 | 0.148 | –0.788 | 0.720 |
| Johnson and Petzel (1971) | medium | 2 | 0.667 | 0.053 | 0.216 | 1.118 |
| Tysk (1983a) | medium | 6 | 0.781 | 0.039 | 0.394 | 1.168 |
| Broadhurst (1969) | medium | 2 | 0.965 | 0.093 | 0.367 | 1.563 |
| Wahl and Sieg (1980) | medium | 2 | 1.129 | 0.089 | 0.544 | 1.714 |
| Penney, et al. (2005) | short | 2 | –0.966 | 0.097 | –1.576 | –0.356 |
| Lee, et al. (2009) | short | 1 | –0.234 | 0.053 | –0.685 | 0.218 |
| Rutschmann (1973) | short | 10 | –0.193 | 0.255 | –1.183 | 0.797 |
| Carroll, et al. (2009b) | short | 1 | 0.243 | 0.068 | –0.270 | 0.756 |
| Lhamon and Goldstone (1956) | short | 2 | 0.721 | 0.055 | 0.261 | 1.182 |
| Tysk (1983a) | short | 1 | 0.768 | 0.039 | 0.379 | 1.157 |
| Bolbecker, et al. (2014) | ultra-short | 1 | –1.172 | 0.034 | –1.532 | –0.812 |
| Elvevag, et al. (2003) | ultra-short | 1 | –0.843 | 0.105 | –1.477 | –0.209 |
| Lee, et al. (2009) | ultra-short | 1 | –0.672 | 0.056 | –1.134 | –0.210 |
| Rutschmann (1973) | ultra-short | 4 | –0.296 | 0.257 | –1.289 | 0.698 |
| Turgeon, et al. (2012) | ultra-short | 1 | 0.032 | 0.100 | –0.587 | 0.652 |
| Carroll, et al. (2008) | ultra-short | 2 | 0.246 | 0.090 | –0.342 | 0.834 |
| Carroll, et al. (2009b) | ultra-short | 1 | 0.278 | 0.069 | –0.236 | 0.791 |
| Carroll, et al. (2009a) | ultra-short | 1 | 1.024 | 0.072 | 0.499 | 1.549 |
| *Precision in time perception* | | | | | | |
| Bolbecker, et al. (2014) | ultra-short | 1 | –5.952 | 0.156 | –6.727 | –5.178 |
| Carroll, et al. (2008) | ultra-short | 2 | –0.312 | 0.090 | –0.900 | 0.276 |
| Carroll, et al. (2009b) | short | 1 | –0.668 | 0.072 | –1.193 | –0.143 |
| Carroll, et al. (2009b) | ultra-short | 1 | –0.594 | 0.071 | –1.116 | –0.072 |
| Carroll, et al. (2009a) | ultra-short | 1 | –1.238 | 0.076 | –1.777 | –0.699 |
| Elvevag, et al. (2003) | ultra-short | 1 | –0.639 | 0.101 | –1.261 | –0.016 |
| Lee, et al. (2009) | short | 1 | –0.733 | 0.056 | –1.197 | –0.268 |
| Lee, et al. (2009) | ultra-short | 1 | –0.660 | 0.056 | –1.122 | –0.199 |
| Lhamon and Goldstone (1973Exp1) | short | 1 | –1.118 | 0.014 | –1.354 | –0.883 |
| Lhamon and Goldstone (1973Exp1) | short | 1 | –1.517 | 0.086 | –2.091 | –0.942 |
| Penney, et al. (2005) | short | 2 | –2.828 | 0.167 | –3.629 | –2.027 |
| Rammsayer (1990) | ultra-short | 1 | –1.406 | 0.059 | –1.881 | –0.931 |
| Roy, et al. (2012) | short | 2 | –0.466 | 0.088 | –1.046 | 0.113 |
| Roy, et al. (2012) | ultra-short | 1 | –0.250 | 0.086 | –0.824 | 0.324 |

| Study | Interval range | *J* | *g* | Var(*g*) | $CI_{lower}$ | $CI_{upper}$ |
|---|---|---|---|---|---|---|
| Todd, et al. (200) | ultra-short | 1 | –0.675 | 0.115 | –1.339 | –0.010 |
| Turgeon, et al. (2012) | ultra-short | 1 | –1.994 | 0.150 | –2.752 | –1.235 |
| Ulferts, et al. (1999) | short | 1 | –1.032 | 0.122 | –1.717 | –0.346 |
| Ulferts, et al. (1999) | ultra-short | 1 | –0.841 | 0.118 | –1.516 | –0.166 |
| Volz, et al. (2001) | short | 1 | –2.036 | 0.264 | –3.043 | –1.029 |
| Waters and Jablensky (2009) | short | 1 | –0.763 | 0.100 | –1.382 | –0.144 |

*Note.* In the analysis, continuation tapping has been regarded as time reproduction. Two interval, temporal bisection, one interval reminder tasks, and temporal deviant detection are grouped as duration discrimination tasks in the main analyses. Temporal bisection is the only duration discrimination task that provides information on accuracy of time perception (bisection point). Outlying data are indicated by grey ink.

### *2.2.7.5 Equations*

Eq.B.1

$$g = \frac{M_d - M_c}{s}$$

Eq.B.2

$$s = \sqrt{\frac{(n_d - 1)SD_d^2 + (n_c - 1)SD_c^2}{n_d + n_c - 2}}$$

Eq.B.3

$$g = \sqrt{F \frac{n_d + n_c}{n_d n_c} \frac{n_d + n_c}{n_d + n_c - 2}}$$

Eq.B.4

$$g = t\sqrt{\frac{n_d + n_c}{n_d n_c} \frac{n_d + n_c}{n_d + n_c - 2}}$$

Eq.B.5

$$Var(g) = \frac{n_d + n_c}{n_d n_c} + \frac{g^2}{2(n_d + n_c)}$$

## 2.3 How long did you look at me? The influence of gaze direction on perceived duration and temporal sensitivity[3]

### 2.3.1 Abstract

Faces that exhibit emotionally negative expressions in mutual gaze have been shown to induce a dilation of perceived duration. The influence of gaze by itself on duration judgments, however, has rarely been investigated. We argue for a social interaction hypothesis, according to which humans should be highly accurate and precise (sensitive) when processing the temporal dynamics of mutual gaze. In three experiments, we investigated whether the direction of observed gaze affects perceived duration and temporal sensitivity. In Experiment 1, subjects did indeed estimate the duration of direct gaze more accurately as compared to the duration of averted gaze. In Experiments 2 and 3, subjects had to categorize direct and averted gaze stimuli as being short or long in duration (temporal bisection). Experiment 2 found temporal sensitivity (but not mean duration judgments) to be improved in cases of mutual gaze. In Experiment 3, the effect of mutual gaze on prolonged subjective duration did replicate, however, it was rather small. Moreover, temporal precision was not improved in the case of naturalistic stimuli. In sum, effects of mutual gaze on duration judgments are rather weak, and cannot be attributed to arousal, as such ratings did not differ between direct and averted gaze stimuli.

### 2.3.2 Introduction

From every-day life, it is evident that the perception of time is highly subjective and can be influenced by several factors. One of the most common phenomena is the dilation of subjective time during unpleasant waiting periods. For example, if you have to wait for a delayed train while it is cold and windy, you will experience time as passing by extremely slowly. In situations of joyful activity, however, time seems to fly. We investigated whether mutual gaze affects subjective time and enhances the accuracy and precision of temporal judgments.

Cognitive models of time perception, as for example the prominent pacemaker-accumulator models, propose an internal clock that comprises a pacemaker emitting pulses and an accumulator collecting these pulses (Meck, 1996; Treisman, 1963). The more pulses

[3] This paper (pages 103 to 125) has been published in Perception together with Prof. Dr. Heiko Hecht: Thönes, S., & Hecht, H. (2016). How long did you look at me? The influence of gaze direction on perceived duration and temporal sensitivity. *Perception*, *45*(6), 612-630.

are accumulated, the longer the perceived duration of a given interval (e.g., Gibbon, 1977; Gibbon et al., 1984). Thus, changes in the accumulation as well as changes in the pacemaker can alter perceived duration. Higher levels of arousal (the level of excitement as induced for example by an emotion) are associated with an increased rate of pulse emission (Angrilli et al., 1997; Droit-Volet & Meck, 2007; Gil & Droit-Volet, 2012; Mella et al., 2011). Accordingly, if more pulses are emitted by the pacemaker within a specified time interval, more pulses can get accumulated and the interval is perceived to be longer. Alternately, paying attention to the time interval may alter the accumulation process and produce longer perceived duration as it causes the accumulator to miss fewer pulses from the pacemaker (Block & Zakay, 1996; S. W. Brown, 1997; Zakay & Block, 1996).

Over the last decades, a substantial body of research has investigated the effects of (negative) arousing stimuli on time perception by means of different facial expressions. Negative facial expressions are more arousing than neutral expressions (e.g., Dimberg & Ohman, 1996) and the prediction of subjective time dilation induced by arousing negative facial expressions has been confirmed frequently (Droit-Volet & Gil, 2009; Gan, Wang, Zhang, Li, & Luo, 2009; Gil & Droit-Volet, 2011; Smith, McIver, Di Nella, & Crease, 2010; Thayer & Schiff, 1975). However, the basic feature of gaze direction has rarely been investigated with regard to time perception, and the two published studies that have focused on gaze direction and time perception yielded inconclusive results using temporal bisection tasks (Doi & Shinohara, 2009; Jusyte, Schneidt, & Schönenberger, 2015). In a temporal bisection task (a special temporal discrimination procedure), the subject has to learn a short and a long anchor duration of, for example, one and two seconds, respectively. Subsequently, intermediate durations are presented and the subject is asked to categorize these as being either more similar to the short anchor duration or to the long anchor duration. If certain factors, such as stimulus properties, speed up or slow down the inner clock of a subject during the test phase, this should systematically affect the subject's temporal judgments. For instance, if the inner clock runs faster in reaction to an arousing facial expression, the amount of pulses that are emitted by the pacemaker and collected by the accumulator during a specific intermediate duration becomes larger. Consequently, an intermediate duration is perceived as being rather long and the subject is biased to categorize the duration as 'long'. The temporal bisection task has been used frequently in research on time perception in humans as well as in animals (e.g., Allan & Gibbon, 1991; G. D. A. Brown, McCormack, Smith, & Stewart, 2005; Droit-Volet & Wearden, 2002; Gautier & Droit-Volet, 2002; Meck, 1983; Siegel & Church, 1984; Wearden et al., 1998). Beside information about systematic over- and underestimation of duration, the subjects' behavior in a temporal bisection task can also be analyzed in terms of the variability of judgments across presentations of the same stimulus duration. The variability of estimates is usually expressed in terms of the duration difference limen, which is the duration difference between the two stimuli at which the subject is able to identify the longer /shorter interval with, for example, 75% correct responses. The smaller this duration difference, the more precise (sensitive) are the temporal judgments of the subject.

#### *2.3.2.1 Arousal hypothesis*

It is well established that arousal enhances the frequency of the pacemaker (e.g., Droit-Volet & Meck, 2007). Because direct gaze induces more arousal than averted gaze (Ellsworth & Langer, 1976; Helminen, Kaasinen, & Hietanen, 2011), direct (mutual) gaze should lead to a dilation of subjective time and prolonged duration judgments. Therefore, in a temporal bisection task, stimuli displaying direct gaze are assumed to be categorized as 'long' more often than stimuli that show averted gaze. This hypothesis has been confirmed by Doi and Shinohara (2009) using fotographs of faces that exhibited direct and averted gaze while expressing the emotions anger and happiness. The study by Jusyte et al. (2015), however, did not find this effect of gaze direction on temporal judgments in a bisection task. Their stimuli were faces showing direct or averted gaze, which exhibited either anger or a neutral affective state. If the arousal hypothesis is correct, and mutual gaze substantially raises arousal as do negative facial expressions, then mutual gaze should lead to an overestimation of duration. Arousal has not been found to have an effect on temporal sensitivity in several experiments on facial expressions and time perception (e.g., Gil & Droit-Volet, 2012). Therefore, if arousal is the relevant mechanism in this context, we would expect an overestimation of the duration of mutual gaze but no effect of gaze direction on temporal sensitivity.

#### *2.3.2.2 Social interaction hypothesis*

Beside possible effects of arousal as induced by direct and averted gaze, gaze is a highly important factor in social interaction. In this context, the direction of gaze as well as its duration are crucial aspects of implicit (nonverbal) communication. According to Boyarskaya, Sebastian, Bauermann, Hecht, and Tuescher (2015), gaze direction expresses communication disposition and attendance, synchronizes turn talking, regulates levels of emotionality, intimacy, affiliation and dominance, signals liking, attraction, credibility, and even mental health (Argyle & Dean, 1965; Argyle, Ingham, Alkema, & McCallin, 1972; Kendon, 1967; Kleinke, 1986). Due to these important functions of gaze, humans should be highly sensitive not only to gaze direction but also to gaze duration. Small differences in mutual gaze duration indicate whether communication behavior is socially appropriate. It has been shown that on average the preferred duration for mutual gaze is stable but may also depend on factors such as gender, age, and situational context (Harrison, Binetti, Coutrot, Mareschal, & Johnston, 2015). Given the fact that social interaction including gazing behavior is inherently dynamic, an accurate and precise representation of appropriate mutual gaze duration is of high social relevance as it is a requirement for recognizing and expressing adequate and successful social behavior (Harrison et al., 2015), such as turn-taking in conversations.

These arguments lead to a hypothesis of social interaction, according to which we would expect temporal judgments to mutual gaze stimuli to be particularly sensitive to small duration differences. In a recent conference paper, Cook, Ayhan, Lai, and Johnston (2011) provide some first evidence for such an effect of mutual gaze on temporal sensitivity and interpret this as an increase in temporal resolution, which represents a functionally relevant mechanism aiding the interpretation of social gaze cues.

In terms of the pacemaker-accumulator model, such a socially cued increase in temporal resolution can be thought of as an increase in clock speed. If the clock period is sufficiently smaller than the difference between a standard and a comparison interval that are to be discriminated in a temporal bisection task (e.g., clock period of 20 ms, interval difference 50 ms), then the subject should be able to give precise (sensitive) temporal judgments. If the clock period is too large (e.g., clock period of 80 ms, interval difference 50 ms), the subject's decisions become uncertain and more variable (temporal sensitivity drops). Besides an effect of clock rate on sensitivity, there should also be an effect on the mean estimates because more pulses are accumulated during a given temporal interval when the inner clock runs faster. Therefore, duration judgments should not only be more sensitive but also prolonged during mutual gaze as compared to averted gaze. Hence, if there is a dedicated increase in clock speed when processing socially relevant cues such as mutual gaze, we would expect an increase in temporal sensitivity and prolonged duration judgments in response to mutual as compared to averted gaze. In favor of the social interaction hypothesis, we expect the judgments of mutual gaze duration not only to be prolonged as compared to judgments of averted gaze duration but also to be more accurate, i.e., closer to the veridical duration of the interval to be estimated. Note that this is possible when the duration of averted gaze is underestimated.

### *2.3.2.3 The current study*

As the two previous studies on gaze and time perception (Doi & Shinohara, 2009; Jusyte et al., 2015) have provided mixed data using naturalistic stimuli, we used highly controlled morphed pictures of neutral faces in Experiments 1 and 2, as did Cook et al. (2011). In Experiment 3, we tested whether the effects obtained in Experiments 1 and 2 are also observable in a more ecological context when using pictures of a variety of real persons as stimuli. We also sought to induce higher arousal levels with the use of the naturalistic faces, as compared to the artificial stimuli used in Experiments 1 and 2.

In contrast to the previous studies, we manipulated gaze not by means of eye direction but by presenting the eyes aligned with the head. In our opinion, such alignment of head orientation and gaze direction more closely mimics everyday situations, in which people fully orient toward a communication partner.

In the first experiment, by means of a relative duration estimation task, we tested possible effects of gaze direction on the perception of time over a wide range of interval durations (1 s to 9 s). The purpose of this first experiment was to evaluate at which durations direct compared to averted gaze may have an influence on the perceived duration of stimulus presentation. In the second experiment, by means of a temporal bisection task, we explicitly focused on short durations below 2 seconds. Based on the results from the first experiment, this interval range appeared to be most relevant. In the third experiment, we applied the same task as in Experiment 2. However, we used naturalistic stimuli (photographs from different persons).

Based on the social interaction hypothesis, we assumed that subjects are more accurate in their temporal judgments of direct gaze (Experiment 1, 2, and 3) and more sensitive to small duration differences of direct gaze relative to averted gaze (Experiment 2 and 3). If arousal mediates these hypothesized effects, we would expect prolonged duration judgments to co-occur with differences of arousal as measurable by the Self-Assessment Manikin scales (SAM; Bradley & Lang, 1994).

### 2.3.3 Experiment 1

#### *2.3.3.1 Method*

*Sample*

A total of 24 students participated in the experiment in return for partial course credit. The data from three subjects were excluded from the analysis due to erratic responding in one case and extremely long operating time and poor comprehension in a second case. The remaining sample consisted of 22 subjects (four male) ranging in age between 19 and 33 years ($M$ = 24.00; $SD$ = 4.29). All subjects gave informed written consent according to the Declaration of Helsinki. All subjects had normal or corrected-to-normal vision.

*Apparatus*

Subjects were tested individually while seated in a room with dimmed light. All instructions and stimuli were presented by a computer equipped with a dual core E5700 3GHz processor and a NVIDIA Quadro FX1400 graphic card. The screen size (Nec MultiSync 90F) was 19" and the resolution was 1280 x 1024 pixels at a display rate of 89Hz and a color depth of 32 bit. We created the stimuli using MakeHuman software and presented the stimuli using Vizard 3. The subject's head was steadied by a chin rest at a viewing distance of 50 cm from the screen. All responses were given by using a computer mouse.

*Stimuli and procedure*

Each trial began with the presentation of a red fixation cross in the center of the screen for 750 ms. The subjects were instructed to fixate the cross and afterwards the eye region of the head stimulus throughout the whole trial. After the fixation cross had disappeared, a digitally morphed head was presented for 10.25 s. During the 10.25 s, we dynamically changed the head's orientation (yoked to gaze direction) after a variable period of time (*t)* to change between direct gaze and averted gaze. The vertical dimension of the head stimulus was 27.4° of visual angle (23 cm at a viewing distance of 50 cm). The horizontal dimensions were 18.7° of visual angle for the averted gaze stimuli (16 cm) and 15.1° for the direct gaze stimuli (13 cm). Between trials, $t$ varied between 1 s and 9 s in steps of 1 s. The dynamical shift in gaze direction always took 250 ms. Thus, in each trial, for a certain period of time, the head gazed at the subject and then, for a different period of time, gazed away from the subject, or vice versa. We counter-balanced the head's initial gaze direction (direct gaze vs. averted gaze) and the sign of the averted gaze (leftward as seen from the subject; +50° vs. rightward as seen from the subject; -50°). In the "direct gaze in the beginning" condition, the head gazed at the subject (0°) in the beginning of the trial and turned to the left (+ 50°) or to the right (-50°) after *t*. In the “averted gaze in the beginning" condition, the head was oriented to the left (+50°) or to the right (-50°) in the beginning of the trial, and then turned to the center (0°) after *t*. Figure 1 shows the time course of a single experimental trial.

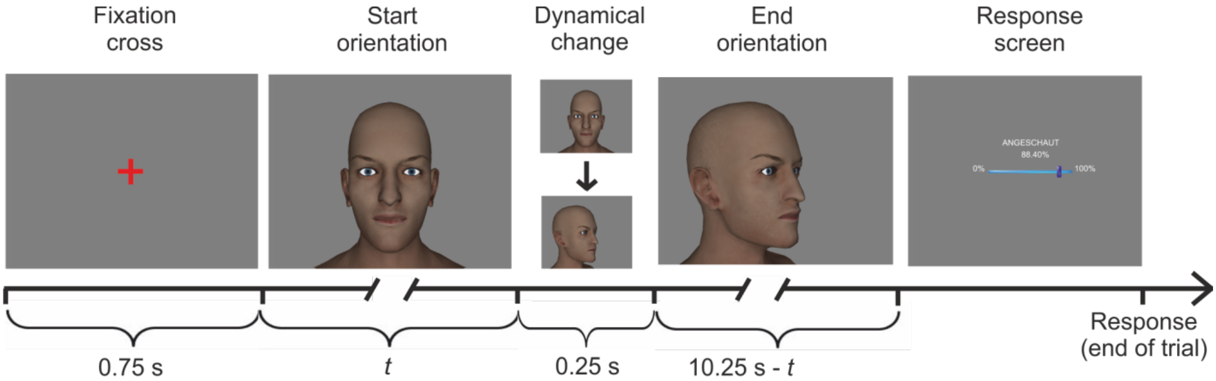


**Figure 1**

Experiment 1: Time course of the experimental procedure. The example shows the starting orientation of direct gaze, which is then averted to the right. Between trials, *t* varied between 1 and 9 s.

There were two experimental blocks. Each block consisted of 72 trials (9 *t* * 2 start orientations * 2 averted gaze orientations; each combination was presented twice per block) resulting in 144 trials per subject. The trials were presented in different random orders. The two experimental blocks differed only in the random order of trial presentation and in the instruction, that is, the head orientation to be timed. In the first block, the subject was instructed to attend to and to estimate the duration of direct gaze relative to the entire duration of head presentation. In the second block, the subject was instructed to attend to and to estimate the duration of averted gaze relative to the entire duration of head presentation. The block order was counter-balanced between subjects. Note that “start orientation” and “orientation of averted gaze” were control variables only. We had no hypotheses regarding systematic influences of “start orientation” and “orientation of averted gaze” on the subjects’ temporal judgments. Therefore, in the analyses, data were averaged across the different start orientations, and averaged across the end orientations of averted gaze resulting in 8 repetitions per subject with regard to the relevant variables “head orientation to be timed” and “target duration”. “Head orientation to be timed” (estimate direct vs. averted gaze) will be referred to as “timing task” in the results section.

After the head had disappeared from the screen, the subjects responded by indicating the duration of direct gaze and averted gaze, respectively, relative to the whole duration of head presentation by adjusting a slide bar on a digital slider. The slider was presented in the center of the screen after each trial. Above the slider, the word “looked at” (“angeschaut”) was presented in the first block and “not looked at” (“nicht angeschaut”) in the second block. Also, a percent value indicated the exact position of the slide bar. After the subject had responded, the next trial began immediately.

At the end of the experiment, by means of Self-Assessment Manikin scales (SAM; e.g., Bradley & Lang, 1994), the subjects indicated their affective reactions to the direct gaze and averted gaze stimuli. Here, each head was presented for 5 s. The presentation order of the three head orientations was counter-balanced between subjects. One after the other was evaluated in terms of valence, arousal, and dominance on 9-point Likert scales.

*2.3.3.2 Results*

In a first step, we transformed the subjects’ percent ratings into duration estimates of direct or averted gaze, respectively, in seconds. Therefore, we multiplied each percent rating with the whole trial duration (10.25 s). All statistical analyses were based on the relative error of the estimated target duration, that is, the difference between the estimated and actual target duration divided by the actual target duration,

$$error_{rel} = \frac{T_{est} - T_{obj}}{T_{obj}}.$$

Thus, positive values indicate overestimation.

Across subjects, we explored the data (3168 data points) by means of box plot analyses. Plotting the relative estimation error as a function of target duration revealed several outlying data points most likely caused by inadvertently switching the two tasks (estimating direct gaze in trials where averted gaze had to be estimated and vice versa). We used the outlier criterion proposed by (Tukey, 1977) and excluded all data points (8.9 % in total) exceeding Q1 − 1.5 * IQR and Q3 + 1.5 * IQR, where Q1 represents the lower quartile, Q3 represents the upper quartile, and IQR is the interquartile range (Q3 − Q1).

Figure 2 illustrates possible effects of timing task (gaze direction) on the relative estimation error in percent of target duration for each target duration (means and standard error of the mean). A tendency towards longer duration judgments for direct gaze stimuli as compared to averted gaze stimuli appeared especially at short target durations.

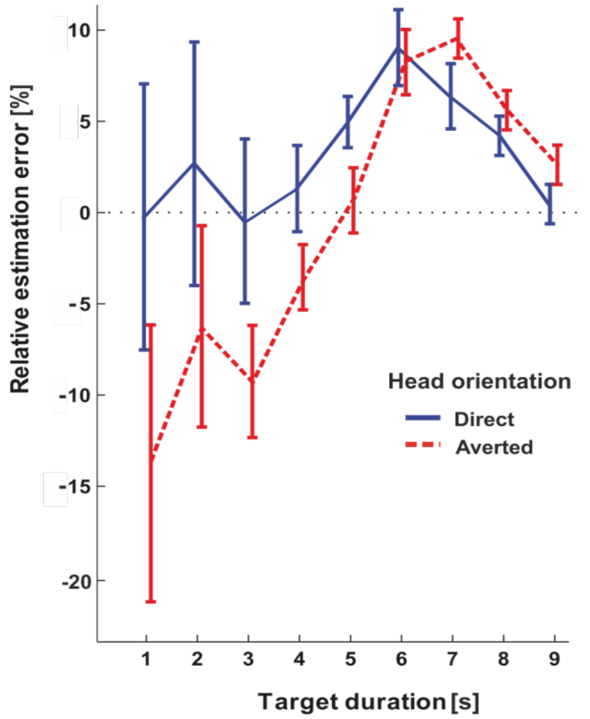


**Figure 2**

Experiment 1: Mean relative estimation errors (standard error of the mean in parentheses) in percent of target duration as a function of target duration and head orientation. Negative values indicate underestimation.

Possible effects of the control variables start orientation and averted gaze orientation on the relative estimation error were evaluated by means of a repeated-measures analysis of variance (rmANOVA). Start orientation had no significant effect on the relative estimation error, $F(1,21) = 2.919$; $p = .102$; $_{partial}\eta^2 = .122$ (direct: $M = 0.011$, $SD = 0.082$; averted: $M = 0.034$, $SD = 0.073$). Averted gaze orientation had a significant effect on the relative estimation error, $F(1,21) = 5.189$; $p = .033$; $_{partial}\eta^2 = .198$ (left: $M = 0.011$, $SD = 0.082$; right: $M = 0.034$ , $SD = 0.072$), with orientation of the avatar to the right side (from the subject's perspective) leading to larger duration estimates. The interaction between start orientation and averted gaze orientation also reached statistical significance, $F(1,21) = 5.504$; $p = .029$; $_{partial}\eta^2 = .208$, indicating overestimation of duration in trials where the avatar was averted to the right in the beginning.

In a second rmANOVA, we analyzed the effects of timing task (head orientation) and target duration on the relative estimation error. Huynh-Feldt corrected values are reported in cases where sphericity could not be assumed. Timing task had a significant effect on the relative estimation error, indicating temporal overestimation as well as larger accuracy when direct gaze duration was to be judged, as compared to averted gaze duration, $F(1,20) = 5.677$;

$p$ = .027; $_{partial}\eta^2$ = .221. Target duration, $F(2.47,49.47)$ = 3.005; $p$ = .048; ε = .309; $_{partial}\eta^2$ = .131, and the interaction between timing task and target duration, $F(3.95,78.92)$ = 2.545; $p$ = .047; ε = .493; $_{partial}\eta^2$ = .113, did also reach statistical significance, which indicates that effects of gaze direction on temporal judgments were most pronounced at short durations.

The influence of head orientation (direct vs. averted) on the SAM-ratings (arousal and valence) were analyzed with two paired samples *t*-tests. The descriptive differences between the two conditions of head orientation (arousal direct gaze: *M* = 3.667, *SD* = 1.853, arousal averted gaze: *M* = 3.286, *SD* = 1.463; valence direct gaze: *M* = 5.476, *SD* = 1.209, valence averted gaze: *M* = 5.286, *SD* = 1.209) did not reach statistical significance, arousal: $t(20)$ = 1.579; $p$ = .130, valence: $t(20)$ = 1.094; $p$ = .287. Moreover, there were no significant correlations between the arousal ratings of the stimuli and the relative estimation errors in the corresponding direct head orientation condition, $r$ = .225; $p$ = .327, and the averted head orientation condition, $r$ = .331; $p$ = .143, respectively.

#### *2.3.3.3 Discussion*

In the first experiment, we tested whether direct gaze stimuli as compared to averted gaze stimuli induce a dilation of subjective time leading to higher temporal accuracy over a wide range of target durations from 1 to 9 seconds. By means of a temporal estimation task, subjects had to judge the presentation duration of direct gaze and averted gaze stimuli relative to a fixed trial duration of 10.25 s. We found a significant effect of timing task (head orientation) on temporal judgments, with duration judgments of direct gaze stimuli being longer and more accurate (closer to the objective duration) relative to averted gaze stimuli. Based on the significant interaction between timing task and target duration, this effect was found to be most pronounced in the short duration range between 1 and 3 seconds even though between-subject variability was larger in this range too. As assessed by the SAM-scales, affective reactions to direct gaze and averted gaze stimuli did not differ significantly from each other. Also, as indicated by correlational analyses, higher levels of arousal were not systematically related to overestimation of duration (positive relative estimation errors). Thus, in favor of the social interaction hypothesis, the systematic effects of gaze direction on duration estimates do not seem to be mediated by arousal. Additionally, although irrelevant with respect to our hypotheses, we found a significant effect of the direction of the avatar's averted gaze (control variable) on the temporal judgments. Averted gaze to the right side (from the subject's perspective) produced larger temporal estimates than averted gaze to the left side, especially when averted gaze was exhibited at the beginning of the trial (significant interaction between averted gaze direction and start orientation).

### 2.3.4 Experiment 2

Based on the results from Experiment 1, we designed a second experiment using a more sensitive methodology that explicitly focused on the short duration range and additionally allowed for investigating temporal sensitivity by means of a temporal bisection task. Previous research on gaze / facial expressions and time perception has likewise focused on these rather short durations using temporal discrimination tasks. We assume that mutual gaze duration especially in the range of a few seconds represents an important social cue (social interaction hypothesis). Therefore, we expected the strongest effects at short durations and focused on these in the second experiment. Direct gaze and averted gaze stimuli as adapted from the first experiment had to be categorized as being either short or long. According to the social interaction hypothesis, we expected the temporal judgments of direct gaze stimuli to be prolonged, more accurate and more precise, which should be reflected in a left-shift of the psychometric function, with the point of subjective equality being closer to the standard duration (1200 ms), and a smaller difference limen in the direct gaze condition (steeper psychometric functions). According to the arousal hypothesis, potential differences in the arousal level would induce a left-shift in the psychometric functions in the direct gaze condition relative to the averted gaze condition, that is, an overestimation of duration of direct gaze stimuli compared to averted gaze stimuli accompanied by higher arousal ratings of direct gaze stimuli. Again, we measured the subjects' affective reactions to the stimuli by means of the SAM-scales in order to test whether direct gaze stimuli would be more arousing than averted gaze stimuli and whether higher levels of arousal are associated with temporal overestimation.

#### *2.3.4.1 Method*

*Sample*

26 students participated in the second experiment in return for partial course credit. The data from 6 subjects were excluded from the analysis due to extremely poor performance according to the criterion for outlier correction proposed by Tukey (1977). The remaining sample consisted of 20 subjects (4 male) ranging in age between 18 and 28 years (*M* = 22.05; *SD* = 2.52). As in Experiment 1, all subjects gave informed written consent and had normal or corrected-to-normal vision.

*Apparatus, stimuli and procedure*

The experiment was conducted in the same room using the same PC and screen as in Experiment 1. All device settings including the lighting of the room were identical to Experiment 1. Also, the stimuli were adapted from Experiment 1. However, in this experiment, stimuli were static and not rotating. All stimuli were presented using Python 2.7.

The subjects' task was to indicate after each trial whether the stimulus that had just been presented was either short or long in duration (one interval discrimination task). Each trial began with the presentation of a fixation cross for 2 s. Subsequently, a blank grey screen followed for 1s. Then, one of the three virtual heads illustrated in Figure 3 was presented.

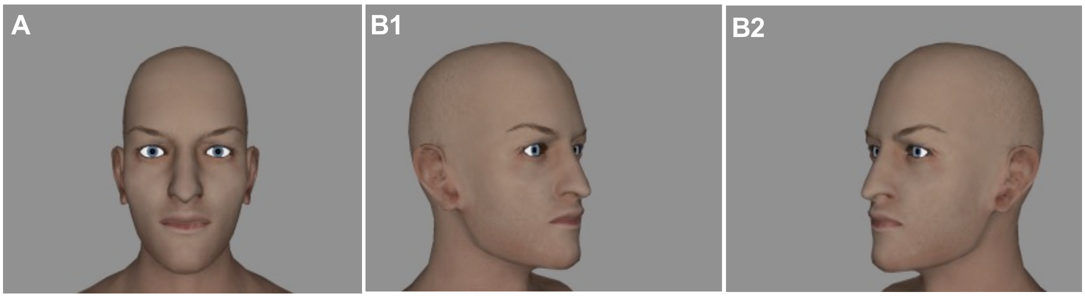


**Figure 3**

Experiment 2: Stimuli as presented in Experiment 2 and adapted from Experiment 1. Panel A represents the direct gaze condition (0°) and Panels B1 (rightward orientation; -50°) and B2 (leftward orientation; +50°) show the averted gaze conditions.

Beside head orientation (0° vs. +50° vs. -50°), we manipulated the presentation duration of the heads (980 vs. 1048 vs. 1121 vs. 1200 vs. 1284 vs. 1374 vs. 1470 ms). Thus, there were three short durations and three long durations while 1200 ms represented the intermediate duration. Starting from the intermediate duration, the shorter and longer durations represented steps of 7 % each. All possible combinations of conditions, 3 head orientations * 7 presentation durations, were presented 20 times resulting in a total of 420 trials per subject. The trials were presented in random orders.

After the head had disappeared, the blank grey screen followed for 1s. Then, the response screen appeared and the subject had to indicate whether the stimulus was short ("kurz") or long ("lang") in duration (relative to the preceding trials). All responses were given by pressing the left (short) or the right (long) button of a response box. There was no training block and the subjects did not know that there was an intermediate (standard) duration physically being neither short nor long. Prior to the experiment, the subjects had been informed that after a few trials they will get an impression of what is meant by "short" and “long”, respectively. Thus, each subject developed an implicit standard interval (the subjective intermediate duration) that provided a criterion for judging a stimulus’ duration as being either short (shorter than the implicit standard) or long (longer than the implicit standard).

After the temporal discrimination task, as in Experiment 1, the subjects indicated their affective reactions to the direct gaze and averted gaze stimuli by means of the SAM-scales.

### *2.3.4.2 Results*

For each subject, we plotted the proportion of “long” responses as a function of stimulus duration and head orientation. Data for the averted gaze (-50° and +50°) were averaged and compared to direct gaze. A repeated measures ANOVA indicated that the proportion of “long” responses was only insignificantly larger in the direct gaze condition ($M = 0.481$; $SD = 0.159$) as compared to the averted gaze condition ($M = 0.462$; $SD = 0.157$), $F(1,19) = 1.120$; $p = .303$; ${}_{partial}\eta^2 = .056$. There were significantly more long responses to longer stimulus durations than to shorter stimulus durations, $F(2.297,43.648) = 181.702$; $p < .001$; $\varepsilon = .383$; ${}_{partial}\eta^2 = .905$ (Huynh-Feldt corrected values are reported). Potential differences in the proportion of long responses between direct and averted gaze stimuli did not vary as a function of stimulus duration, $F(6,114) = 1.593$; $p = .155$; ${}_{partial}\eta^2 = .077$ (insignificant interaction between head orientation and stimulus duration). We fitted the psychometric model and determined the difference limen (DL) and the point of subjective equality (PSE; point of bisection) for each subject and head orientation (for an example, see Figure 4). Note that a shift of the psychometric function to the left, that is a smaller PSE, indicates overestimation of duration, and a steeper function (i.e., a smaller DL) indicates higher temporal discrimination sensitivity.

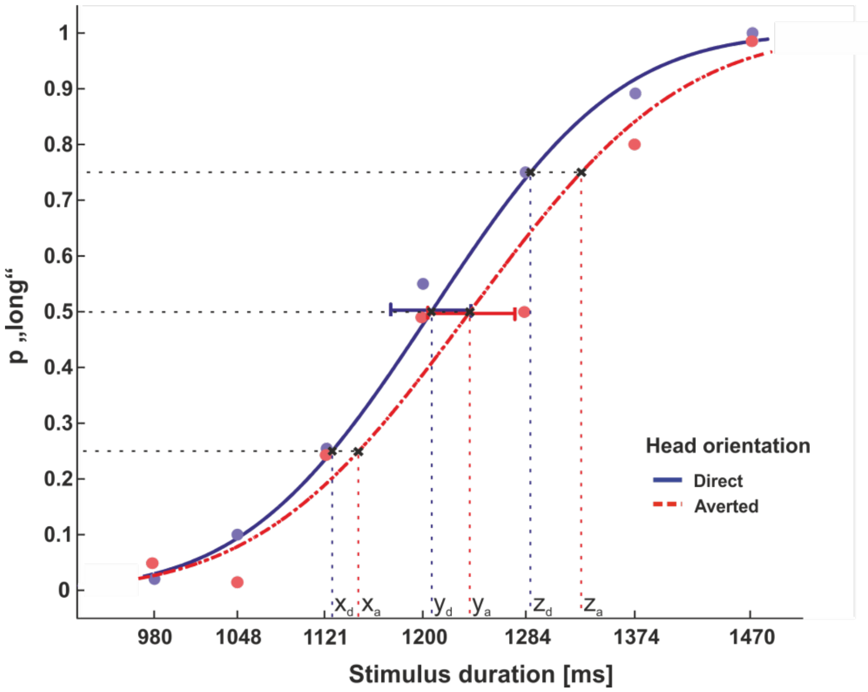


**Figure 4**

Experiment 2: Psychometric functions fitting the proportion of "long" responses as a function of stimulus duration and head orientation (example data from subject number 8). $y_a - y_d$ represents the effect of head orientation on the point of subjective equality (point of bisection). $(z_a - x_a)/2 - (z_d - x_d)/2$ represents the effect of head orientation on the difference limen. Error bars indicate the 95% confidence interval of the PSE.

The PSE and the DL were analyzed by means of two independent paired samples *t*-tests. For the influence of head orientation on PSE and DL, we report Cohen's $d_z$ (Cohen, 1988) as a measure of effect size in a dependent measures design. Here, $d_z$ is defined as $M_d$ / $SD_d$, where $M_d$ is the mean of the differences between the relative estimation error in the direct gaze condition and in the averted condition, and $SD_d$ is the standard deviation of the differences. The PSE was only insignificantly larger in the averted gaze condition ($M$ = 1251.44 ms; $SD$ = 124.54 ms) as compared to the direct gaze condition ($M$ = 1227.93 ms; $SD$ = 101.67 ms), $t(19) = 1.604$; $p = .125$; $d_z = 0.36$. The DL was significantly smaller in the direct gaze condition ($M$ = 138.88; $SD$ = 45.81 ms) as compared to the averted gaze condition ($M$ = 166.55 ms; $SD$ = 65.59 ms), $t(19) = 2.363$; $p = .010$; $d_z = 0.53$, indicating higher temporal sensitivity in trials of direct gaze.

As in Experiment 1, the influence of head orientation (direct vs. averted) on the SAM-ratings were analyzed by means of two paired samples *t*-tests. The descriptive differences between the two conditions of head orientation (arousal direct gaze: $M$ = 3.750, $SD$ = 1.618, arousal averted gaze: $M = 3.325$, $SD = 1.624$; valence direct gaze: $M = 4.450$, $SD = 1.191$, valence averted gaze: $M = 4.900$, $SD = 1.199$) did not reach statistical significance, arousal: $t(19) = 1.669$; $p = .111$, valence: $t(19) = 1.517$; $p = .146$. Additionally, by means of correlational analyses, we investigated whether higher arousal ratings were associated with smaller PSEs as predicted by the arousal hypothesis. However, the results did not indicate such a relationship (correlations between arousal rating and PSE for direct gaze: $r = .324$; $p = .164$, and averted gaze: $r = .137$; $p = .566$).

#### *2.3.4.3 Discussion*

By means of a duration discrimination task, we tested whether direct mutual gaze, as compared to averted gaze, leads to prolonged and more accurate duration estimates and to larger temporal sensitivity in the range of short durations below 2 seconds. According to the social interaction hypothesis, we expected an overestimation of duration of direct gaze stimuli relative to averted gaze stimuli, which should be reflected in a left shift of the psychometric functions, with the PSE being smaller and closer to the standard duration. This assumption was not supported by the data, as the smaller PSE and the larger proportion of long responses in the direct gaze condition remained statistically insignificant. According to the social interaction hypothesis, we also assumed subjects to be more sensitive to duration differences of direct gaze stimuli relative to averted gaze stimuli, as indicated by steeper psychometric functions. This part of the hypothesis was supported by an effect of gaze on temporal sensitivity (smaller difference limen when processing mutual gaze).

As there was only a weak trend towards higher ratings of arousal and lower ratings of valence of direct gaze stimuli as compared to averted gaze stimuli and the correlations between arousal ratings and the PSEs remained statistically insignificant, the notion that arousal is mediating the effects of gaze on time perception has not been supported.

### 2.3.5 Experiment 3

Experiments 1 and 2 indicated that temporal judgments are more accurate and more precise for direct as compared to averted gaze while the subjects' affective reactions did not differ between direct and averted gaze. These results are in favor of the social interaction hypothesis and do not support the notion that arousal is a relevant mechanism in the context of temporal processing of neutral gaze direction. Based on these results, using the same task procedure as in Experiment 2, we conducted a third experiment with naturalistic stimuli (photographs of different persons) in order to test whether the results are robust in a more ecological context, which is likely to induce higher arousal levels.

### *2.3.5.1 Method*

*Sample*

30 students participated in the third experiment in return for partial course credit. The data from 6 subjects were excluded from the analysis due to extremely poor performance according to the criterion for outlier correction proposed by Tukey (1977). The remaining sample consisted of 24 subjects (4 male) ranging in age between 18 and 44 years ($M = 24.58$; $SD = 7.46$). As in Experiments 1 and 2, all subjects gave informed written consent and had normal or corrected-to-normal vision.

*Apparatus, stimuli and procedure*

The experiment was conducted in the same room using the same PC and screen as before. All device settings including the lighting of the room were identical to those of Experiments 1 and 2. The task procedure was almost identical to Experiment 2. However, the stimuli were photographs of real persons instead of digitally morphed heads. There were 40 different photographs of 20 different persons (head only). For each person, one photograph displayed direct gaze and a second one averted gaze. Half of the photographs (20) were pictures of 10 celebrities taken from the internet. Between persons, these pictures differed with respect to several aspects (e.g., facial expression, clothing, degree of averted gaze). For each person (pair of pictures), we matched these variables between averted and direct gaze pictures. The other half of the photographs were of 10 colleagues from our laboratory. These pictures were taken using a Canon EOS 400D camera controlling for facial expression (all neutral), brightness, and degree of averted gaze (all 30°). For pictures of colleagues as well as celebrities, we balanced the sign of averted gaze (5 to the left, 5 to the right, each) and the gender of the person shown in the picture (5 female, 5 male). All photographs were matched in size (visual angle: about 27° in the vertical dimension and 17° in the horizontal dimension). The persons' heads were presented against a gray background (for an example, see Figure 5).

**Figure 5**

Experiment 3: Example for direct and averted gaze stimuli used in Experiment 3.Upper panel: colleague, lower panel: celebrity.

Each single stimulus was presented once during the experiment. Data were aggregated across stimuli of the same person category (celebrity vs. colleague), head orientation (direct vs. averted), and presentation duration (980 vs. 1048 vs. 1121 vs. 1200 vs. 1284 vs. 1374 vs. 1470 ms, as in Experiment 2), resulting into a 2 *person categories* * 2 *head orientations* * 7 *presentation durations* * 10 repetitions design. Thus, in total there were 280 trials per subject. Prior to the experimental phase, 14 trials were presented to the subject for the purpose of training. These trials were randomly chosen from the pool of experimental trials. As in Experiment 2, during each trial, the subjects' task was to indicate whether the stimulus that had just been presented was either short or long in duration, relative to the implicit standard duration.

After the temporal discrimination task, the subjects indicated their affective reactions to all direct gaze and averted gaze stimuli (40 stimuli in total) by means of the arousal-scale (SAM).

### *2.3.5.2 Results*

For each subject, we plotted the proportion of "long" responses as a function of stimulus duration, head orientation, and person category. A repeated measures ANOVA indicated that the proportion of "long" responses was significantly larger in the direct gaze condition ($M$ = 0.543; $SD$ = 0.177) as compared to the averted gaze condition ($M$ = 0.521; $SD$ = 0.175), $F(1,23) = 4.710$; $p = .041$; ${}_{partial}\eta^2 = .170$ (main effect of head orientation). There were significantly more long responses to longer stimulus durations than to shorter stimulus durations, $F(4.197,96.532) = 268.856$; $p < .001$; $\varepsilon = .700$; ${}_{partial}\eta^2 = .921$ (main effect of stimulus duration; Huynh-Feldt corrected values are reported in cases where sphericity cannot be assumed). Differences in the proportion of long responses between direct and averted gaze stimuli did not vary as a function of stimulus duration, $F(5.222,120.110) = 0.417$; $p = .844$; $\varepsilon = .870$; $_{partial}\eta^2 = .018$ (insignificant interaction between head orientation and stimulus duration). There was no effect of person category on the proportion of long responses, $F(1,23) = 0.251$; $p = .621$; $_{partial}\eta^2 = .011$ (insignificant main effect of person category). All remaining interactions did not reach statistical significance ($p > .05$).

We fitted the psychometric model and determined the difference limen (DL) and the point of subjective equality (PSE; point of bisection) for each subject, head orientation, and person category. The PSE and the DL were analyzed by means of a rmANOVA. The PSE was only insignificantly larger in the averted gaze condition (*M* = 1505.62 ms; *SD* = 89.01 ms) as compared to the direct gaze condition (*M* = 1492.67 ms; *SD* = 74.84 ms), $F(1,23) = 1.204$; $p = .284$; $_{partial}\eta^2 = .050$; $d_z = 0.22$. There was no effect of person category on the PSE (celebrity: *M* = 1494.15 ms; *SD* = 82.00 ms; colleague: *M* = 1504.14 ms; *SD* = 81.85 ms), $F(1,23) = 1.551$; $p = .225$; $_{partial}\eta^2 = .063$) and no interaction between head orientation and person category, $F(1,23) = 0.052$; $p = .821$; $_{partial}\eta^2 = .002$.

The DL did not vary between the direct gaze condition (*M* = 302.02; *SD* = 97.80 ms) and the averted gaze condition (*M* = 289.42 ms; *SD* = 79.76 ms), $F(1,23) = 0.960$; $p = .337$; $_{partial}\eta^2 = .040$, indicating no effect of gaze direction on temporal sensitivity. There was also no effect of person category on the DL (celebrity: *M* = 290.33 ms; *SD* = 77.88 ms; colleague: *M* = 301.10 ms; *SD* = 99.68 ms), $F(1,23) = 0.861$; $p = .363$; $_{partial}\eta^2 = .036$) and no interaction between head orientation and person category, $F(1,23) = 1.060$; $p = .314$; $_{partial}\eta^2 = .044$.

Potential influences of head orientation (direct vs. averted) and person category (celebrity vs. colleague) on the SAM-rating (arousal) were analyzed by means of a repeated measures ANOVA. Differences in arousal ratings to direct gaze stimuli (*M* = 3.35; *SD* = 1.22) and averted gaze stimuli (*M* = 3.49; *SD* = 1.20) were not significant, $F(1,23) = 0.582$; $p = .453$; $_{partial}\eta^2 = .025$. Pictures of celebrities (*M* = 3.81; *SD* = 1.40) were rated to be significantly more arousing than pictures of colleagues (*M* = 3.04; *SD* = 1.02), $F(1,23) = 13.145$; $p < .001$; $_{partial}\eta^2 = .364$. There was no interaction between head orientation and person category, $F(1,23) < .001$; $p > .999$; $_{partial}\eta^2 < .001$.

As in Experiments 1 and 2, by means of correlational analyses, we investigated whether higher arousal ratings were associated with smaller PSEs as predicted by the arousal hypothesis. Overall, the results did not indicate such a relationship (correlations between arousal rating and PSE across head orientations and person categories: $r = -.328$; $p = .118$).

#### *2.3.5.3 Discussion*

Similar to Experiment 2, we tested whether direct mutual gaze, as compared to averted gaze, leads to more accurate duration estimates and/or to improved temporal sensitivity in a temporal bisection task. In contrast to Experiment 2, we used naturalistic stimuli. According to the social interaction hypothesis, we expected an overestimation of duration (and at the same time an increase in accuracy) of direct gaze stimuli relative to averted gaze stimuli. And indeed, stimuli that exhibited direct gaze were judged longer more often than stimuli exhibiting averted gaze. However, in the psychophysical analysis, the PSE was neither smaller nor closer to the standard duration for direct gaze stimuli. The second assumption based on the social interaction hypothesis was that subjects process direct gaze with larger temporal sensitivity. This part of the hypothesis was clearly not supported by the data as the difference limen did not differ between trials of direct and averted gaze. The well-controlled pictures of colleagues did not differ from the less controlled pictures of celebrities. Both induced a comparable effect of direct gaze on duration judgments.

Arousal ratings did not differ between direct and averted gaze, nor were they correlated with the PSE. Moreover, the more arousing pictures of celebrities were not judged differently in duration than the less arousing pictures of colleagues. Thus, potential effects of gaze direction on temporal estimates cannot be explained by differences in the arousal level.

### 2.3.6 General discussion

Based on the fact that the direction as well as the duration of gaze have important functions in social interaction and communication (e.g., Boyarskaya et al., 2015; Kleinke, 1986), humans should be highly accurate and sensitive when estimating the duration of mutual gaze. According to this hypothesis of social interaction, we expected the temporal judgments of direct gaze stimuli to be prolonged, more accurate and more precise relative to judgments of averted gaze. Taken together, the three experiments provide evidence that direct gaze can change the subjective duration of gaze that is present on the order of several seconds. The results from the duration estimation task in the first experiment support the idea of prolonged and more accurate duration judgments of direct gaze stimuli as compared to averted gaze stimuli. The data from the temporal bisection task in the second experiment provided only weak evidence for such an effect of gaze direction on accuracy. However, subjects were more sensitive to differences in the presentation duration of direct gaze stimuli as compared to averted gaze stimuli (judgments were more precise). In the third experiment, we used the same time perception task (temporal bisection) as in the second experiment but different naturalistic stimuli. When simply analyzing the proportion of long responses to direct and averted gaze stimuli, we found the duration of direct gaze to be significantly overestimated. However, this effect was not reflected in smaller PSEs in the psychometric analysis, and there was no effect of gaze direction on the difference limen. Thus, the effect of larger temporal sensitivity when processing mutual gaze that has been observed in Experiment 2 was not replicated for the naturalistic stimuli.

Taken together, effects of gaze direction on duration judgments are task-dependent and rather weak. These results are in accordance with the previous studies on gaze direction and time perception. Whereas Doi and Shinohara (2009) provided evidence for prolonged duration judgments (more long responses) when faces exhibited direct gaze relative to averted gaze, the study by Jusyte et al. (2015) did not find such an effect when using a temporal bisection task. Cook et al. (2011) reported positive effects of mutual gaze on temporal sensitivity but no effects on accuracy.

We have also investigated whether potential effects of gaze direction on time perception are mediated by higher levels of arousal in situations of mutual gaze. According to the arousal hypothesis, direct gaze induces more arousal than averted gaze (Dimberg & Ohman, 1996), and hence a dilation of subjective time. Such an effect should primarily be reflected in an overestimation of direct gaze stimuli as compared to averted gaze stimuli, which was not consistently observable. Based on the arousal ratings, in all three experiments, the direct gaze stimuli were not evaluated as being significantly more arousing than the averted gaze stimuli. Additionally, there were no significant correlations between arousal ratings and mean duration estimates, that is, higher levels of arousal were not systematically related to temporal overestimation. Therefore, arousal does not seem to play an important role in the context of gaze direction and temporal processing. However, we cannot entirely rule out an arousal effect that was too small to detect with the resolving power of our experimental designs. We may have failed to detect a small correlation between arousal and duration judgments and are only able to rule out arousal as the main driving force.

As direct gaze has generally been shown to induce more arousal than averted gaze (Ellsworth & Langer, 1976; Helminen et al., 2011; Kleinke, 1986; Thayer & Schiff, 1975), the lack of differences in the subjects' affective reactions towards direct gaze and averted gaze stimuli in the first and second experiment might also be due to the neutral and artificial character of the stimuli used. For example, the mental state that an observer attributes to a gazing subject influences the processing of gaze properties, as for example gaze direction (Teufel et al., 2009). We sought to rule out this possible limitation by using naturalistic images of real persons in Experiment 3. These stimuli did not induce different levels of arousal but they did produce the effect of mutual gaze lengthening duration judgments.

It should be noted that the SAM scales may not provide a sufficiently reliable method for measuring arousal, especially with regard to the rather small sample sizes used for our experiments. Therefore, the conclusion that arousal does not play any important role in the context of our study should be viewed cautiously. In other contexts, perception of direct gaze has been found to be modulated by emotional expressions. For example, angry faces were more likely to be perceived as showing direct gaze than fearful faces (Ewbank, Jennings, & Calder, 2009; Rhodes, Addison, Jeffery, Ewbank, & Calder, 2012). As previous studies have supported the arousal hypothesis for direct gaze faces exhibiting angry expressions (e.g., Doi & Shinohara, 2009; Droit-Volet & Gil, 2009; Gil & Droit-Volet, 2011), it is possible that the enhanced perception of direct gaze of angry faces contributed to the stronger effect of direct gaze on perceived duration in these studies. Therefore, a substantial time dilation effect based on mutual gaze may only occur when direct gaze is combined with specific emotional expressions.

Though not of primary interest to our study, in the first experiment, we found an effect of averted gaze direction on the temporal judgments. Averted gaze to the right side (from the subject's perspective) led to generally larger temporal estimates than averted gaze to the left side. This effect can be explained in terms of the concept of a mental time line, according to which future-related concepts are represented on the right side while past-related concepts are represented on the left side of the visual field (Hartmann et al., 2014; Stocker, 2014). This relation holds for subjects from cultures where writing direction goes from the upper left to the lower right (Bergen & Lau, 2012). In the context of our first experiment, the orientation of the avatar (head with centered gaze) to the right may have triggered future-related concepts, prompting the subject to produce somewhat larger temporal estimates. At this point, such an effect remains speculative but it deserves investigation in future experiments.

In the second experiment, direct gaze was processed with larger temporal sensitivity. However, this effect was not replicable in the third experiment when pictures of real persons were used that unsystematically differed with respect to several variables between different persons. Could the visual properties of the stimuli in Experiment 1 and 2, such as overall luminance and contrast have differed systematically and sufficiently between conditions to be a potential confounding factor? Due to the side view in the averted gaze condition, the exposed sclera on either side of the iris was less visible in the averted than in the mutual gaze condition. This change in contrast is necessarily confounded with gaze direction and could have attracted more attention to the direct gaze stimulus. This may have contributed to the increase in temporal sensitivity in the mutual gaze condition. However, the subjects were explicitly instructed to always fixate and attend to the eye region of each stimulus. We believe that this instruction made the subjects less prone to potential bottom up effects of slightly brighter eye regions in the mutual gaze condition.

Another aspect of concern might be the fact that direct and averted gaze stimuli, especially in the first and second experiment, differed in the extent of deviation from bilateral symmetry, with the direct gaze stimulus being more symmetric than the averted gaze stimuli. It has recently been shown that symmetric stimuli are perceived as longer in duration than asymmetric stimuli (Bertamini, Ogden, Rampone, & Makin, 2013). Hence, the direct gaze stimulus may have induced temporal dilation due to the fact that it was more symmetric than the averted gaze stimuli. Moreover, bilateral symmetry may also have an effect on temporal sensitivity. As the (symmetric) direct gaze stimulus appears more regular, judgments of duration of this stimulus may also be more stable, that is, less variable and more sensitive. The influence of symmetry on temporal sensitivity may systematically be addressed in future research.

Systematical differences in terms of basic image properties, such as overall luminance and bilateral symmetry, between direct and averted gaze stimuli were less likely in the third experiment. Here, the set of stimuli was comprised of naturalistic pictures from different persons. With respect to several image properties, these stimuli unsystematically differed between and within trials of the same gaze orientation. What differed systematically was the orientation of gaze, either being direct or averted. If the social interaction hypothesis is correct and direct gaze reliably enhances temporal sensitivity and accuracy, this effect should clearly occur in an experiment using naturalistic stimuli that are close to everyday social encounters. As the third experiment did not produce an augmented effect of gaze direction on duration judgments, we cannot interpret the social interaction hypothesis in any quantitative fashion.

In conclusion and in accordance with previous studies, the present study provides only weak and task-dependent evidence for an influence of mutual gaze on subjective time. Across our three experiments, duration judgments of stimuli exhibiting mutual gaze were not consistently more sensitive than judgments to stimuli showing averted gaze. Direct gaze stimuli did produce longer subjective durations albeit with effect sizes that were rather small. These results question the social interaction hypothesis, according to which the need for adequate social behavior should produce highly accurate and sensitive duration judgments in the presence of mutual gaze.

### 2.3.7 Acknowledgements

We are grateful to Daniel Oberfeld, Klaus Landwehr and Bernhard Both for helpful comments and discussions on the manuscript and the statistical analyses. We would also like to thank Bernhard Both, Larissa Reis and Max Ramdohr for data collection and especially Agnes Münch for programming and technical assistance.

## 2.4 Counting does not improve but may compromise the accuracy of time production[4]

### 2.4.1 Abstract

It is common to use counting strategies in order to accurately produce short time intervals. However, evidence has been lacking. Does counting indeed improve the accuracy (deviation of mean produced duration from veridical duration) and precision (intra-individual variability of produced durations) of time productions in the range of seconds to minutes? In two experiments, we compared chronometric counting to intuitive timing (no counting) and to attentional control (simultaneously performing mental arithmetic). In Experiment 1, subjects had to produce time intervals of 60 s duration in a class-room setting. Counting did not improve the accuracy of time productions. To the contrary, subjects were most accurate when producing the 60-s time interval intuitively and overestimated duration when counting or doing mental arithmetic. Experiment 2 replicated these results in a laboratory setting for intervals of 60, 30, and 10 s. The advantage of intuitive judgment over counting was replicated for 60-s-intervals but disappeared at shorter intervals. As expected, mental arithmetic significantly impaired accuracy as well as precision and led to overproduction of duration. The results are discussed in the context of attentional processes and a potential word-length effect in counting.

### 2.4.2 Introduction

In many everyday situations, we have to judge short durations accurately, such as when boiling eggs or brewing tea. Does counting help when we have to produce temporal intervals in this range? Most people spontaneously start to count when they have to produce duration. The experimental studies that have reported benefits of a counting strategy on temporal judgments have used methods of *verbal estimation, time reproduction*, or *duration discrimination*, but never *time production tasks* (Getty, 1976; Grondin, Laflamme, & Mioni, 2015; Grondin, Meilleur-Wells, & Lachance, 1999; Grondin, Ouellet, & Roussel, 2004; Rattat & Droit-Volet, 2012; Wearden, Denovan, Fakhri, & Haworth, 1997). In a *time production task*, subjects have to indicate when a defined time interval has elapsed. We believe that such a task more closely mimics the above mentioned everyday situations. Therefore, we have

[4] A paper based on this manuscript (pages 126 to 139) has been published together with Prof. Dr. Heiko Hecht: Thönes, S. & Hecht, H. (in press). Counting does not improve the accuracy of long time productions. *Attention, Perception, & Psychophysics.*

investigated the potential effects of counting on *time productions* of 10-, 30-, and 60-second intervals while controlling for attentional processes in two experiments.

In previous studies on the effects of counting, very short intervals in the range of seconds have typically been used. For instance, Grondin et al. (1999) found that counting improves temporal accuracy for stimuli lasting up to several seconds in duration discrimination. Two studies focused on somewhat longer intervals up to 27 seconds, but reported no positive effects of counting on temporal accuracy in verbal time estimation, and time reproduction, respectively (Gilliland & Martin, 1940; Hinton & Rao, 2004). To our knowledge, only one study has used durations approaching one minute (Hicks & Allen, 1979a; reported negative effects of counting on accuracy of verbal estimates).

In a first field experiment, we therefore investigated whether counting leads to more accurate *productions* of a 60-second interval as compared to intuitive judgment without counting. In a third condition, we additionally asked the subjects to engage in mental arithmetic while producing the time interval of 60 s. Such a cognitively demanding task usually distracts attention from the timing task, which is well-known to lead to temporal overproduction (Block, 1990; Zakay & Block, 1996). The purpose of this task was to ensure that subjects followed instructions. Only when we replicate this effect can we interpret a potential null-effect of counting.

By means of a second experiment, we aimed at replicating the results from the first experiment in a lab situation and added two more target durations (10- and 30-s intervals). Moreover, the second experiment was comprised of several trial repetitions per condition, thus providing information not only about accuracy of time productions but also about their precision. While *accuracy* indexes the deviation of the *mean* produced time interval from the veridical value of the time interval (constant error), *precision* refers to the intra-individual *variability* of judgments across several trials (variable error). Studies on the effects of counting on precision of duration judgments indicate less variable judgments when subjects use counting strategies (Getty, 1976; Grondin et al., 1999; Grondin et al., 2004; Hinton & Rao, 2004; Killeen & Weiss, 1987; Rattat & Droit-Volet, 2012; Wearden & Lejeune, 2008). These exisiting studies, however, have focused on durations in the range up to few seconds only. We therefore investigated whether counting also improves the precision of duration judgments in the range up to 60 s.

### 2.4.3 Experiment 1

#### *2.4.3.1 Method*

The sample for the experiment was drawn from an undergraduate student population (approximately 80% female, mean age 22 years). The experiment was conducted at the beginning of a psychology lecture. The lecture room contained 110 seats. The students were informed about the task procedure and they were asked to decide whether or not they wanted to participate in the experiment. Two video cameras were used to record the simultaneous performance of 58 students who consented to participate.

The subjects' task was to produce time intervals of 60 s. In response to a start signal given by the experimenter, the subjects were instructed to keep track of time and to silently raise their hand holding up a card when they thought that 60 s had elapsed. After the cards had been given to the subjects, they were instructed to keep their eyes closed and to remain silent during the interval production task. Interval production was repeated twice resulting in three trials per subject in total. The subjects successively performed three different *Tasks* during the interval productions. In the condition intuitive timing, they were instructed to produce the time interval of 60 s without counting or any other potential strategy. In the condition counting, the subjects were instructed to count from 1 to 60 (one number per second) while producing the time interval. In the condition arithmetic (attentional control), the subjects had to count back from 1,000 in steps of 7 while producing the time interval. The cards the subjects had to raise in order to mark the end of each interval were either green, or yellow, or red. Based on the color of the cards, which were randomly distributed across the class (no clustering of groups), each subject was randomly assigned to one of three *Task order*-conditions. The experimental design followed a Latin square: a) subjects who received a green card were instructed to count during the first interval, to time intuitively during the second, and to perform the arithmetic task during the third; b) subjects holding a yellow card had to time intuitively first, followed by arithmetic and counting; c) a red card indicated the *Task order* arithmetic – counting – intuitive timing.

Each trial ended after all subjects had raised their hands. The experimenter then started the next trial, that is, he gave the start signal for the next 60 s to be produced.

*2.4.3.2 Results*

Based on the video recordings, the produced durations in seconds were determined for each subject in each trial / *Task* (intuitive timing, counting, arithmetic). The response coder was blind to the hypotheses of the experiment. For each subject, an interval was defined as being produced when the hand raise was fully executed, that is, when the upward movement of the arm had stopped. Data from two subjects were excluded from the analyses because in one trial their hands (cards) were not sufficiently visible. As a function of *Task* and *Task order*, the mean produced durations from the remaining 56 subjects are presented in Figure 1. Productions were most accurate (closest to the veridical duration of 60 s) when the subjects were instructed to produce the interval intuitively.

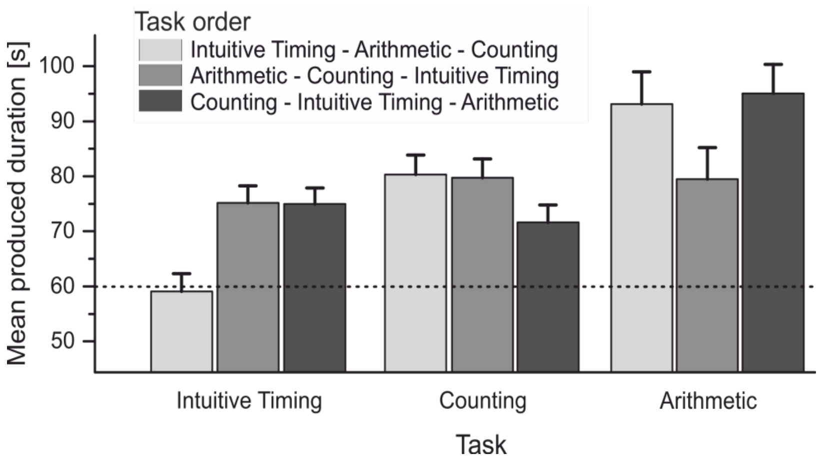


**Figure 1**

Mean produced duration in seconds as a function of *Task* and *Task order*. Error bars indicate standard errors of the mean. The dotted line indicates accurate interval production of 60 s.

The data were analyzed by means of an rmANOVA including the within-subjects factor *Task* and the between-subjects factor *Task order*. Huynh-Feldt corrected values are reported. For subsequent pairwise comparisons, we adjusted the alpha-level according to the Bonferroni Holm-procedure (Holm, 1979), and report Cohen's $d_z$ (Cohen, 1988) as a measure of effect size in a dependent measures design.

There was a significant effect of *Task* on produced duration, $F(1.745, 92.510) = 20.141$, $p < .001$, $\varepsilon = .873$, ${}_{partial}\eta^2 = .275$. As indicated by post-hoc *t*-tests for dependent samples, productions were significantly more accurate and shorter for the condition intuitive timing relative to counting, $t(55) = 2.561$, $p = .013$, $d_z = 0.34$, as well as compared to arithmetic, $t(55) = 5.208$, $p < .001$, $d_z = 0.70$. Moreover, productions were significantly more accurate and shorter for counting as compared to arithmetic, $t(55) = 3.460$, $p = .001$, $d_z = 0.46$.

There was no main effect of *Task order* on produced duration, $F(2, 53) = 0.345$, $p = .710$, ${}_{partial}\eta^2 = .013$. A significant interaction between *Task* and *Task order*, $F(3.491, 92.510) = 5.939$, $p = .001$, $\varepsilon = .873$, ${}_{partial}\eta^2 = .183$, indicated that the interval productions were particularly accurate ($M = 59.12$ s) and the effect of *Task* was most pronounced when intuitive timing was to be performed first.

Moreover, accuracy appeared to generally drop from trial to trial. Thus, we included the within-subjects factor *Trial* in a second rmANOVA analyzing whether time intervals were produced more accurately in the first trial as compared to the later trials. This was indeed the case as witnessed by a main effect of *Trial* on produced duration, $F(2, 110) = 8.751$, $p < .001$, ${}_{partial}\eta^2 = .137$. The productions in the first trial were closer on target ($M = 70.36$, $SD = 19.66$) than those in the second ($M = 82.02$, $SD = 17.56$) and third trial ($M = 84.20$, $SD = 20.95$). As these differences in accuracy between first and later trials may modulate the effects of *Task*, we additionally analyzed task-dependent differences in time productions of the first trial only (see Figure 2). A one-way ANOVA revealed a significant and large between-subjects effect of *Task* on produced duration, $F(2, 55) = 5.560$, $p = .006$, with more accurate interval productions by intuitively timing subjects as compared to counting subjects, $t(36) = 2.467$, $p = .019$, $d = 0.82$, and those who engaged in arithmetic, $t(33) = 3.176$, $p = .003$, $d = 1.12$ (unpaired samples *t*-tests). The counting condition did not significantly differ from the correct estimate of 60 s, $t(16) = 0.266$, $p = .794$, $d = 0.07$.

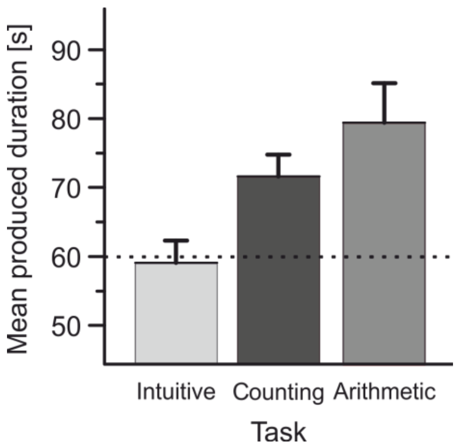


**Figure 2**

Mean produced duration in seconds as a function of task in the first trial. Error bars indicate standard errors of the mean. The dotted line indicates accurate interval production of 60 s.

#### *2.4.3.3 Discussion*

We tested whether chronometric counting improves the accuracy of time production of a 60-s time interval, in comparison to a no-counting condition (intuitive interval production), and a dual-task condition (mental arithmetic). Surprisingly, the temporal judgments in the intuitive condition were about 10 seconds shorter and closer to the target duration than temporal judgments in the counting condition. This effect was most pronounced and largest in size when the subjects were unbiased by previous interval productions. As expected, time productions were largest and most inaccurate when the subjects had engaged in the cognitively demanding arithmetic task.

Why did our subjects systematically overproduce the time interval when applying a counting strategy? This may be so because the mental production of larger numbers, for example sub-vocalizing "forty-seven", takes more time than the production of smaller numbers, such as "four" (Ellis, 1992). Such a word-length effect (Baddeley, Thomson, & Buchanan, 1975) may lead to slower counting in the range above 10 seconds, resulting in temporal overproduction.

Interestingly, the produced intervals became less accurate and larger by more than 10 seconds in the second and third trial as compared to the first trial. A related repetition effect has been described previously (Hicks & Allen, 1979a, 1979b; Ryan, 2011). This temporal overproduction (equivalent to underestimation in estimation tasks) in trials occurring later in the experiment may be explained by a decrease in the arousal level from the first to the later trials, which has caused a slower pacemaker rate of the internal clock (Gibbon, 1977; Gibbon et al., 1984; Treisman, 1963).

In a second experiment, we aimed at replicating the surprising advantage of intuitive timing over counting strategies in the production of time intervals in a lab situation and additionally added two more interval durations of 10 and 30 s. Moreover, by instructing subjects to count to 10 in different numerical ranges (1 to 10, 21 to 30, and 51 to 60), we tested whether the assumed word-length effect accounts for the temporal overproduction in the counting condition. If word length was the relevant factor, overproduction of duration should be most pronounced in the condition counting from 51 to 60, because this range contains the longest words to be vocalized.

### 2.4.4 Experiment 2

#### *2.4.4.1 Method*

*Sample*

A total of 24 students (16 female; mean age = 23.5, *SD* = 6.9) participated in the experiment in return for partial course credit. According to the criterion proposed by Tukey (1977), no outliers were detected. All subjects gave informed written consent according to the Declaration of Helsinki. All subjects had normal or corrected-to-normal vision and hearing. Based on the experimental design and the effect-size estimates obtained in Experiment 1, the sample size of 24 subjects was recommended by G*Power (Faul, Erdfelder, Lang, & Buchner, 2007).

*Apparatus*

Subjects were tested individually while seated in a room with dimmed light. Using the software Python 2.7, all instructions and stimuli were presented by a computer equipped with a dual core E5700 3GHz processor and an nVidea Quadro FX1400 graphics card. The screen size (Nec MultiSync 90F) was 19” and the resolution was 1280 x 1024 pixels at a display rate of 89 Hz. The auditory stimuli were presented via headphones (Ultrasone HFI-780). All responses were given by using the spacebar of the computer keyboard.

*Stimuli, task procedure, design*

The experiment was divided into two parts, and all factors were manipulated within subjects. In the first part, the subject had to produce intervals of 30 and 60 s duration. While producing a time interval, the subject was instructed to either count from 1 to 30/60, to perform mental arithmetic, or to time the interval intuitively without any counting strategy (as in Experiment 1). In the second part, the subject was instructed to produce time intervals of 10 seconds. During the interval production, the subject was instructed to either count from 1 to 10, to count from 21 to 30, to count from 51 to 60, or to time the interval intuitively. Prior to each part, the subject received detailed written instructions for the tasks in the upcoming part of the experiment. The veridical durations of the intervals were never revealed to the subject and no feedback was given throughout the experiment.

In the first part, each trial began with a short written instruction that indicated which interval duration had to be produced (30 vs. 60 s), and which specific task had to be performed (intuitive timing vs. counting from 1 to 30/60 vs. mental arithmetic: counting backwards from 1,000 in steps of 7). To proceed, the subject was instructed to press the response button. Subsequently, for 1.5 s, a white fixation cross appeared in the center of the black background screen. The fixation cross was followed by a sinus tone (1000 Hz, 50 ms) marking the beginning of the time interval to be produced by the subject. The subject was instructed to indicate the end of the time interval by pressing the response button and to keep the eyes closed during the production of the interval. Simultaneously to the button press, the tone was presented again, this time to mark the end of the interval. Subsequently, the next trial began with the presentation of the trial-specific instruction (e.g., counting; 30 s). Each trial was presented four times resulting in 24 (2 *Intervals* * 3 *Tasks* * 4) trials per subject in the first part.

In the second part, again, each trail began with the presentation of a short instruction that indicated which specific task had to be performed while producing the 10-s time interval (intuitive timing vs. counting from 1 to 10 vs. counting from 21 to 30 vs. counting from 51 to 60). The subject had to press the response button to proceed after having read the instructions. As in the first part, a fixation cross appeared for 1.5 s and was followed by the tone that marked the beginning of the interval. Again, the subject was instructed to indicate the end of the interval by pressing the response button and to keep eyes closed until the response. Simultaneously to the button press, the tone was presented again to mark the end of the interval. The next trial began with the presentation of the trial-specific instruction. Each trial was presented four times resulting in 16 (4 Tasks * 4) trials per subject in the second part.

Within each part, the trials were ordered randomly, separately for each subject. The whole experiment lasted approximately 40 minutes. The experimenter was blind to the hypotheses and to the results from the first experiment.

*2.4.4.2 Results*

*Part 1*

As a function of *Task* and *Interval*, we analyzed the mean produced duration (measure of accuracy). The descriptive data are presented in Figure 3, indicating no difference between counting and intuitive timing at 30 s duration, but an overproduction due to counting at 60 s. Arithmetic seems to cause strong overestimation irrespective of interval duration.

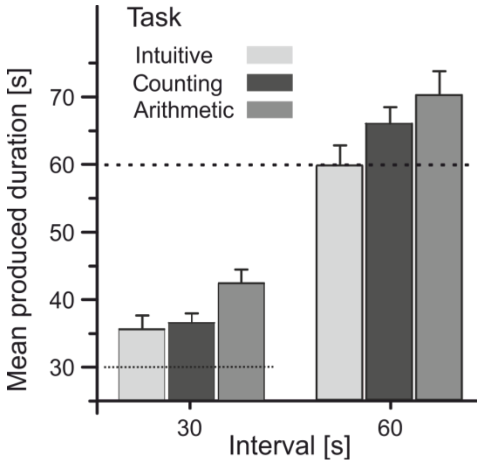


**Figure 3**

Mean produced duration in seconds as a function of *Task* and *Interval*. Error bars indicate standard errors of the mean. The dotted lines indicate accurate interval production of 30, and 60 s, respectively.

By means of an rmANOVA, we analyzed the effects of *Task* and *Interval* on mean produced duration statistically. Huynh-Feldt corrected values are reported. For pairwise comparisons, we additionally report $d_z$ as a measure of effect size in a dependent measures design (Cohen, 1988). For pairwise post hoc comparisons between the different conditions of *Task*, we adjusted the α-level according to the Bonferroni-Holm procedure (Holm, 1979).

There was a significant effect of *Interval*, $F(1, 23) = 458.387$, $p < .001$, $_{\text{partial}}\eta^2 = .952$. Produced duration was longer for the 60-s target interval ($M = 38.02$ s, $SD = 6.91$ s) as compared to the 30-s target interval ($M = 65.64$ s, $SD = 11.47$ s). More importantly, there was a significant effect of *Task*, $F(2, 46) = 8.300$, $p = .001$, $_{\text{partial}}\eta^2 = .265$, indicating overproduction of time intervals in the arithmetic condition ($M = 56.24$ s, $SD = 10.72$ s) compared to counting ($M = 51.45$ s, $SD = 8.49$ s), and overproduction in counting compared to intuitive timing ($M = 47.80$ s, $SD = 12.49$ s). A significant interaction between *Task* and *Interval,* $F(2, 46) = 4.591$, $p = .015$, $_{\text{partial}}\eta^2 = .166$, suggested that differences between counting and intuitive timing are specific to the 60 s target interval. We further analyzed the effect of *Task* and the interaction by means of paired samples *t*-tests. The pairwise comparisons confirmed that counting caused a significant overproduction of the 60-s target interval, compared to intuitive timing, $t(23) = 2.684$, $p = .013$, $d_z = 0.55$, with intuitive timing on average being closer to the veridical duration of 60 s. There was no such difference between counting and intuitive timing for the 30-s target duration, $t(23) = 0.263$, $p = .795$, $d_z = 0.05$. In comparison to counting, the 30 s target interval was significantly overproduced when subjects performed mental arithmetic, $t(23) = 3.445$, $p = .002$, $d_z = 0.70$. There was no significant difference between counting and arithmetic for the 60-s interval, $t(23) = 1.212$, $p = .238$, $d_z = 0.25$.

Additionally, as a second dependent variable, we analyzed the standard deviation of produced duration (a measure of precision). The descriptive data are presented in Figure 4.

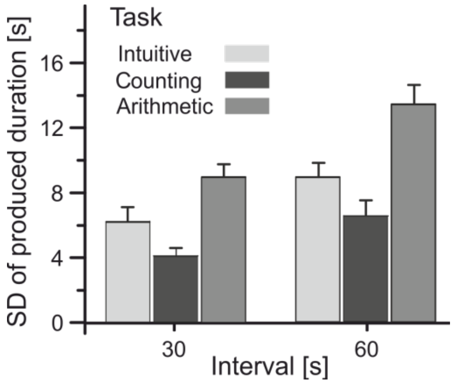


**Figure 4**

Standard deviation of produced duration in seconds as a function of *Task* and *Interval*. Error bars indicate standard errors of the mean.

In an rmAVOVA including the factors *Task* and *Interval*, there was a significant effect of *Interval*, $F(1, 23) = 22.569$, $p < .001$, $_{partial}\eta^2 = .495$, with production being more precise for the 30-s target interval ($M = 6.32$ s, $SD = 2.17$ s) as compared to the 60-s target interval ($M = 9.49$ s, $SD = 2.98$ s). There was also a significant effect of *Task*, $F(1.938, 44.585) = 21.117$, $p < .001$, $_{partial}\eta^2 = .479$, indicating less precise duration production in the arithmetic condition ($M = 10.92$ s, $SD = 3.56$ s) as compared to counting ($M = 5.34$ s, $SD = 2.87$ s) and intuitive timing ($M = 7.46$ s, $SD = 3.09$ s). There was no significant interaction between *Task* and *Interval*, $F(2, 46) = 0.532$, $p = .591$, $_{partial}\eta^2 = .023$. We further analyzed the effect of *Task* by means of paired samples *t*-tests. The pairwise comparisons confirmed that arithmetic impaired the precision of duration production of the 30-s target interval, $t(23) = 2.231$, $p = .036$, $d_z = 0.46$, and the 60-s target interval, $t(23) = 3.147$, $p = .005$, $d_z = 0.64$, in comparison to intuitive timing. Descriptive differences between counting and intuitive timing did not reach statistical significance, 30 s: $t(23) = 1.957$, $p = .063$, $d_z = 0.40$, 60 s: $t(23) = 1.900$, $p = .070$, $d_z = 0.39$, however, time productions tended to be more precise for counting at both intervals.

*Part 2*

As a function of *Task*, the descriptive data of mean produced duration (Panel A) and the standard deviation of produced duration (Panel B) are presented in Figure 5.

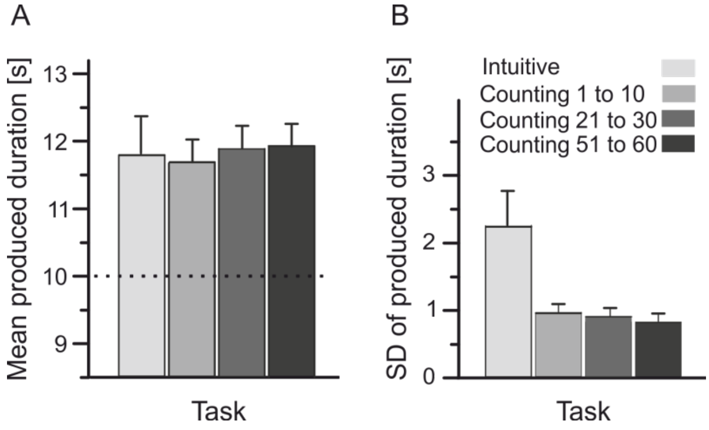


**Figure 5**

Mean produced duration (Panel A) and standard deviation of produced duration (Panel B) in seconds as a function of *Task*. Error bars indicate standard errors of the mean. The dotted line indicates accurate interval production of 10 s.

In a first rmANOVA, we tested a possible effect of word length on produced duration. Therefore, the factor Task included the three counting conditions (counting from 1 to 10 vs. counting from 21 to 30 vs. counting from 51 to 60). There was no significant effect of Task on produced duration, $F(1.647, 37.890) = 1.093$, $p = .335$, $_{\text{partial}}\eta^2 = .045$, indicating no effect of word length on produced duration.

In a second step, we investigated whether mean productions of the 10-s interval differed between normal counting (from 1 to 10) and intuitive timing. A paired-samples *t*-test did not indicate such an effect, $t(23) = 0.235$, $p = .816$, $d_z = 0.05$.

In an additional *t*-test, we compared the SD of produced 10 s duration (precision) between intuitive timing and counting from 1 to 10. Interval productions were significantly more precise for counting ($M = 1.00$ s, $SD = 0.70$ s) as compared to intuitive timing ($M = 2.23$ s, $SD = 2.41$ s), $t(23) = 2.592$, $p = .016$, $d_z = 0.53$.

*2.4.4.3 Discussion*

In Experiment 2, we replicated the result that the production of 60 s was most accurate when subjects timed the interval intuitively, whereas counting (similar to mental arithmetic) led to overproduction of time intervals. This result was restricted to the long interval, as there were no differences in mean productions between intuitive timing and counting at durations of 10 and 30 s. With regard to the variability of time production across trials, the subjects' productions were most variable/least precise when performing mental arithmetic. At 30 and 60 s duration, in comparison to intuitive timing, the precision of time production tended to be higher when subjects were instructed to count. For the short time interval of 10 s, temporal precision was clearly enhanced due to counting.

In order to test whether the temporal overproduction in the long interval-counting condition was caused by an effect of word length, we instructed our subjects to count to 10 in three different ways: from 1 to 10, from 21 to 30, and from 51 to 60. If word length was the relevant factor, overproduction of duration was expected to be strongest in the counting from 51 to 60-condition, as this condition contains the longest words to be vocalized. However, there was clearly no effect of counting strategy on mean interval productions. Based on this result, the overproduction of longer intervals cannot be explained in terms of a word length effect.

### 2.4.5 General Discussion

In two experiments, we tested whether chronometric counting improves the accuracy of duration judgments in the range between 10 and 60 seconds, in comparison to a no-counting condition (intuitive interval production), and a dual-task condition (mental arithmetic). At 10- and 30-s durations, mean interval productions did not differ between counting and intuitive timing. In the 60-s condition, however, and consistently across both experiments, the mean produced durations in the no-counting condition (intuitive interval production) where about 10 seconds shorter and closer to the target duration than the mean produced durations in the counting conditions. As expected, time productions were largest and most inaccurate when the subjects engaged in the cognitively demanding arithmetic task. This result is compatible with the well-established and robust effects of dual tasks on time perception (e.g., S. W. Brown, 1997; Champagne & Fortin, 2008; Rammsayer & Ulrich, 2011), thus indicating that subjects had carefully followed the instructions in our experiments. When a non-temporal task distracts attention from the timing process, temporal intervals are usually underestimated and overproduced (Block & Zakay, 1996; S. W. Brown, 2008; Zakay & Block, 1996). Note that within a pacemaker-accumulator model (Treisman, 1963), this would amount to the accumulator missing some beats for reasons of distraction.

In contrast to common belief and in contrast to the positive effects of counting on duration reproduction (Getty, 1976; Hinton & Rao, 2004) and discrimination (Grondin et al., 1999; Grondin et al., 2004; Wearden et al., 1997), our results clearly indicate negative effects of counting on the accuracy of time production of longer intervals. They are, however, consistent with findings by Hicks and Allen (1979a) who reported underestimation of duration in a verbal estimation task when subjects used a counting strategy.

Why did our subjects systematically overproduce the long time interval when applying a counting strategy? We had assumed that this may be the result of a word-length effect in the mental production of larger numbers. For example, sub-vocalizing "fifthy-one" takes more time than the production of smaller numbers, such as "one" (Ellis, 1992). Such a word-length effect (Baddeley et al., 1975) may lead to slower counting in the range above 10 seconds resulting in temporal overproduction. In the second experiment, we explicitly tested this hypothesis by instructing subjects to apply three different counting strategies, which were associated with the mental production of numbers of different length. We did not find evidence for an effect of word length on produced duration and therefore reject the word length-hypothesis. In both experiments, mean interval productions of 60 seconds were consistently comparable between the counting and the arithmetic condition (both conditions led to overproduction of duration). Accordingly, counting in the range of larger numbers could be viewed as a light dual-task condition that distracts some attention from the timing task, thus causing temporal overproduction. This explanations, however, is challenged by the result that the precision of time production was higher in the counting condition as compared to arithmetic. A distraction of attention would be expected to go along with an increase in variability (e.g., Grondin, Laflamme, & Gontier, 2014). Counting may simply not be a good strategy when it comes to the accurate production of longer intervals.

The precision of time production was enhanced when subjects were instructed to count. However, a clear advantage of counting on precision was restricted to the short interval duration of 10 s. Here, interval productions were more variable when the subjects were instructed to time the interval intuitively. These results are consistent with and extend the previous reports about positive effects of counting on the precision of short interval judgments that were limited to duration discrimination and reproduction tasks (Getty, 1976; Grondin et al., 1999; Grondin et al., 2004; Hinton & Rao, 2004; Rattat & Droit-Volet, 2012; Wearden & Lejeune, 2008).

Taken together, counting has differential effects on the accuracy and precision of temporal judgments and, importantly, these effects depend on the interval durations that are to be produced. Whereas counting does not improve mean duration judgments in the range up to

30 seconds and even compromises productions of 60-s intervals, the precision of duration judgments is enhanced by counting at short durations of 10 s. Mean duration judgments of 60 s are surprisingly accurate only when subjects intuitively produce duration. This latter result was highly consistent between our two experiments using differential experimental setups. Thus, do not bother to count when timing your breakfast egg.

### 2.4.6 Acknowledgements

We wish to thank Andreas Baranowski and Larissa Reis for their help with data collection and processing in Experiment 1. We are also grateful to Agnes Münch for programming and technical assistance and to Max Ramdohr for his help with data collection in Experiment 2.

## 2.5 Is mental time embodied interpersonally?[5]

### 2.5.1 Abstract

Recent evidence has shown that the spatialized mental representation of time is "embodied" – time is systematically expressed via the hands and the eyes. These findings suggest the existence of a manually reflected mental time line running horizontally from left (past) to right (future) and an ocularly reflected mental time line running from left/down (past) to right/up (future). We addressed the question whether mental time is also reflected interpersonally and investigated whether an avatar's head orientation (left, straight, right) would facilitate a subject's temporal processing in relation to the horizontal mental time line. Our results provide clear evidence for manually reflected mental time running from left to right, even for temporal auditory words that are free of potential visual (reading direction) confounds. The data, however, do not indicate an activation of the horizontal mental time line by interpersonal cues, such as static head orientation and body orientation. Additionally, our results suggest the existence of an eye contact (gaze) effect in processing mental time.

### 2.5.2 Introduction

Recent evidence has shown that the spatialized mental representation of time is "embodied" – that it is systematically expressed via the hands and the eyes (Bonato et al., 2012; Stocker et al., 2015). For the *hands*, it has been shown (in cultures that read from left to right) that the left hand reacts faster to the past and the right hand faster to the future, suggesting the existence of a *manually reflected mental time line* running horizontally from left (past) to right (future) (Ouellet et al., 2010; Santiago et al., 2007). For the *eyes,* it has recently been shown that Western observers look more leftward/downward when processing the personal (episodic) past and more rightward/upward when processing the personal (episodic) future (Hartmann et al., 2014). Furthermore, observers look more upward when processing time words relating to the future than when processing time words relating to the past (Stocker et al., 2015). These results suggest that an *ocularly reflected mental time line* is running from left/down (past) to right/up (future). Despite these findings, various questions in relation to the degree of embodiment of time remain open and uninvestigated. In the current paper, we have addressed the question if mental time is reflected interpersonally and

[5] This manuscript (pages 140 to 152) has been submitted for publication together with Dr. Kurt Stocker, Prof. Dr. Peter Brugger, and Prof. Dr. Heiko Hecht.

investigated whether an avatar's head (gaze) orientation (left, straight, right) would facilitate temporal processing of subjects in relation to the horizontal mental time line. Does observing *another* person's (an avatar's) head (gaze) orientation influence our own sense of time?

The *manual* reflection of a mental time line has mainly been shown for the horizontal (left-to-right) dimension (faster reaction to the past with a left key press and to the future with a right key press) (Ouellet et al., 2010; Santiago et al., 2007) as well as for the sagittal (back-to-front) dimension (e.g., faster reaction to past events by moving a handle backward and to future events by moving a handle forward) (Eikmeier, Alex-Ruf, Maienborn, & Ulrich, 2015). However, most of the left/right manual studies have used verbal stimuli in written format (written words or sentences that refer to the past or the future), which leaves room for a potential confound: Presenting the stimulus in written form naturally involves reading the word or sentence from left to right. Now, it is well established that the flow of time is related to the direction of the writing system. In Western cultures, time runs from left to right in the horizontal domain, but this left-to-right time line can easily be reversed to a right-to-left-time line after only five minutes of mirror reading (Casasanto & Bottini, 2014). Thus, with manual studies of mental time that involve written stimuli, we never know if the effect only occurs because the decision is influenced by the current reading direction, or if the effect is more deeply rooted and occurs without immediate prior exposure to written language. Fortunately, a few of these manual studies that have found evidence for the left-to-right mental time line did use auditory stimuli (listening to recorded spoken words or sentences that refer to the past or the future) (Kong & You, 2011; Ouellet et al., 2010). Since listening to stimuli does not involve reading, these auditory-stimuli studies have to some degree established that the manually reflected left-to-right flow of time can also occur in the absence of reading. However, for German (the language of our subjects) it remains still unknown, whether manual responses would reveal a consistent left-to-right mapping of time in the absence of prior immediate exposure to reading (with auditory stimuli only a back/front effect has been reported thus far; Eikmeier, Hoppe, & Ulrich, 2015). For German, it is a particularly interesting question if the manually reflected left-to-right time line can be evoked in the absence of reading, as recent research has shown that in German the sagittal (back-to-front) time axis is more basic than the horizontal (left-to-right) axis (Eikmeier, Alex-Ruf, et al., 2015). For this reason, we employed a classic left/right key-press paradigm together with auditory time stimuli (spoken words such as "gestern (yesterday)", "morgen (tomorrow)", and others; cf. below). One aim of our study was thus to replicate the left/right-hand encoding of time in a so far uninvestigated field (German auditory processing of time).

The findings that the left hand reacts faster to past stimuli and the right hand to future stimuli also relates to findings of the mental number line, where small numbers are represented to the left and larger numbers to right – an effect that is famously called the SNARC (spatial-numerical association of response codes) effect (Dehaene, Bossini, & Giraux, 1993). Our first hypothesis (hypothesis in relation to manually reflected mental time) concerned what has been labeled a STEARC effect (spatial-temporal association of response codes; Ishihara, Keller, Rossetti, & Prinz, 2008), a spatial mapping not of number magnitude, but of temporal flow. Specifically, we predicted that such an effect would still be observed when subjects processed temporal concepts auditorily (that is, without possible confounds of immediate prior cognitive processing of reading direction). Thus, responses to auditorily-presented past-related words should be faster when the response has to be given by the left hand, and responses to future-related words should be faster by the right hand (compatible conditions). Likewise, responses to auditorily-presented past-related words produced by the right hand as well as responses to future-related words by the left hand should be slower (incompatible conditions).

The main focus of our study was, however, to simultaneously add more embodied components to this classic manual temporal task, so that not only the already well-researched embodied dimension *manual reaction* could be investigated, but also the potentially embodied dimension of interpersonal cues, such as *head/eye orientation* of another subject. Irrespective of potential interpersonal effects, the *ocular* (eye gaze) reflection of the mental time line has been shown in three ways in Western cultures. First, our eyes look along a time line from left/down (past) to right/up (future) while processing episodic memory or episodic future thinking (Hartmann et al., 2014). Second, our eyes look along a time line from "less up" (past) to "more up" (future) during online auditory processing of temporal-relation words like *before-that* (German: "vorher") and *after-that* (German: "nachher") (Stocker et al., 2015). Third, our eyes look along a time line from left (earlier items) to right (later items) during serial recall tasks (Rinaldi, Brugger, Bockisch, Bertolini, & Girelli, 2015). However, it has never been investigated whether interpersonal cues, such as head or gaze orientation of another human being, can trigger temporal embodiment effects.

We investigated the question of the possible existence of interpersonally reflected mental time by using an avatar (interpersonal situation), and asked whether the avatar's head orientation (left, straight, right) would facilitate the subject's temporal processing in relation to a horizontal (left-right) mental time line. Thus, our second hypothesis (hypothesis in relation to interpersonally reflected mental time) is that head orientation of the avatar to the left facilitates past processing, whereas head orientation of the avatar to the right facilitates future processing. Importantly, in line with a crucial differentiation between egocentric and "alter-ego-centered" perspectives in reduplicative bodily experiences (Brugger, 2002), we will distinguish whether the avatar faces the observer ("front-view avatar") or is viewed from behind ("back-view avatar"). Accordingly, with regard to the horizontal dimension, the perspective of the subject matches the perspective of the avatar when the avatar is viewed from behind (e.g., left/past from the avatar is left/past from the subject). The perspectives do not match when the avatar faces the subject (e.g., left/past from the avatar is right/future from the subject). Therefore, if viewing the avatar is perceived as a social situation, an effect of the avatar's head orientation on temporal word processing should be stronger when the avatar is viewed from behind.

### 2.5.3 Material and methods

#### *2.5.3.1 Sample*

A total of 20 students participated in the experiment in return for partial course credit. The data from one subject were excluded from the analysis due to poor task performance according to the criterion for outliers as proposed by Tukey (1977). The remaining sample consisted of 19 subjects (six male) with a mean age of 23.37 years ($SD$ = 7.43 years). All subjects gave informed written consent according to the Declaration of Helsinki. All subjects had normal or corrected-to-normal vision.

#### *2.5.3.2 Apparatus*

Subjects were tested individually while seated in a room with dimmed lights. All instructions and stimuli were presented by a computer equipped with a dual core E5700 3GHz processor and an NVIDIA Quadro FX1400 graphics card. The screen size (Nec MultiSync 90F) was 19" and the resolution was 1280 x 1024 pixels at a display refresh rate of 89 Hz and a color depth of 32 bit. We created the visual stimuli (human avatars) using MakeHuman software and presented them using Python 2.7. The subject's head was steadied by a chin rest at a viewing distance of 50 cm from the screen. All responses were given by using a response box with two buttons (10 cm distance between buttons), which were arranged horizontally (left vs. right) relative to the subject. The response box was placed centered in relation to the screen.

#### 2.5.3.3 Stimuli and procedure

Each trial began with the presentation of a red fixation cross, which was presented centrally in the upper part of the screen against a grey background. After 0.75 s, the fixation cross disappeared and the grey screen remained for 0.25 s. Subsequently, an avatar was presented. The vertical dimension of the avatar was approximately 22° of visual angle (20 cm at a viewing distance of 50 cm). The horizontal extent of the avatar was approximately 8° of visual angle (7 cm). The avatar's head was presented at the same position where the fixation cross had been presented. Relative to the subject's perspective, the avatar's head was oriented either to the left, to the right, or straight ahead. Additionally, the avatar's body was oriented either directly toward the subject (front visible) or averted 180º from the subject (back visible). Eye orientation of the avatar was always left constant as sagittally ahead. Overall, the combination of body and head orientations resulted in 6 different avatar orientations as presented in Figure 1, which varied randomly across trials. The subject's task was to name the avatar's orientation verbally as quickly as possible. Possible responses were "front_0", "front_90", "front_270", "back_90", "back_180", and "back_270". "0" indicated orientation toward the subject (eye gaze to the subject), "90" orientation to the left, "180" orientation into the screen (eye gaze in the same direction as the subject), and "270" orientation to the right, from the perspective of the subject. The answers were given in German. The only purpose of this task was to ensure that the subject processed the orientation of the avatar thoroughly. The responses were monitored by the examiner.

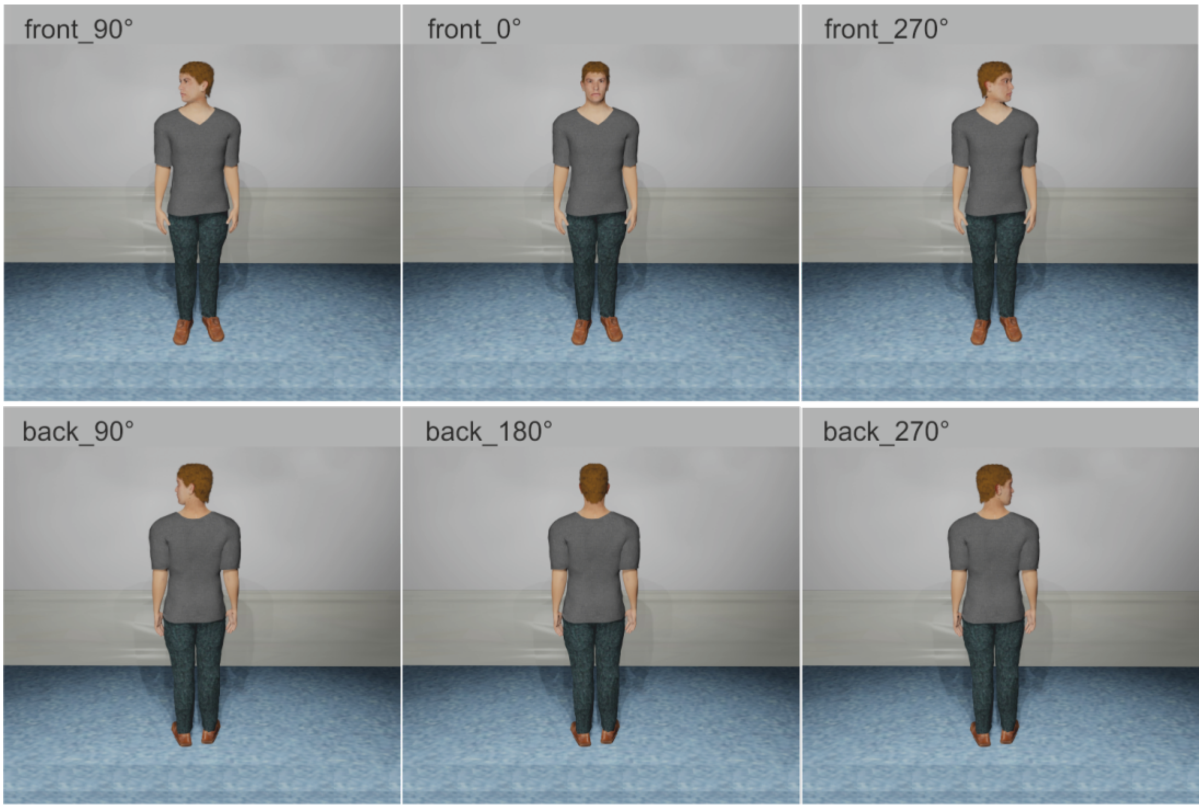


**Figure 1**

Visual stimuli as presented (without written labels) in the experiment.

3 s after the onset of the avatar presentation, one of four different target words was presented binaurally via Ultrasone HFI-780 headphones. Possible target words were “vorgestern” (the day before yesterday; p2), “gestern” (yesterday; p1), “morgen” (tomorrow; f1), and “übermorgen” (the day after tomorrow; f2). The four target words were recorded with a hypothesis-blind German male speaker in a phonetics laboratory, and were subsequently matched with respect to volume. The subject’s task was to indicate whether the target word was past-related (“vorgestern”, “gestern”) or future-related (“morgen”, “übermorgen”) by pressing one of the two buttons as quickly and correctly as possible. There were two experimental blocks. In one block, past-related words had to be indicated by pressing the left button with the index finger of the left hand, and future-related words had to be indicated by pressing the right button with the index finger of the right hand. This stimulus response-mapping was compatible to the assumed mental representation of past- vs. future-related concepts (compatible SR-mapping). In a second block, future-related words had to be indicated by pressing the left button with the left index finger, whereas past-related words had to be indicated by pressing the right button with the right index finger. This reversed stimulus-response mapping was incompatible with the assumed mental representation of past- vs.

future-related concepts (incompatible SR-mapping). The order of blocks (compatible vs. incompatible SR-mapping) was counter-balanced between subjects. Immediately after the subject had responded to the target word, the avatar disappeared from the screen. A grey screen was presented for 1 s before the next trial began with the presentation of the fixation cross. Responses > 3000 ms after the presentation of the target word were discarded from the analysis, and the respective trial was repeated at the end of the current block.

At the beginning of each block, we presented 12 training trials to familiarize the subject with the task. The training trials were drawn randomly from the pool of experimental trials. As in the experimental blocks, a trial was repeated if the subject responded incorrectly. For each trial, we recorded accuracy and reaction time.

#### *2.5.3.4 Design*

Each combination of Word (p2 vs. p1 vs. f1 vs. f2), Body-orientation (front vs. back), and Head-orientation (left vs. straight vs. right) was presented eight times, resulting in 192 trials (block 1). All trials were repeated (block 2) with reversed SR-mapping (compatible vs. incompatible). Within each block, the trials were ordered randomly, separately for each subject. Block order was counterbalanced between subjects.

### 2.5.4 Results

Subjects responded too slowly in 0.6 % of all trails. Accordingly, there were only few trial repetitions per subject due to slow responding ($M = 2.30$; range: 0 to 17). On average, subjects responded incorrectly in 1.5 % of all trials (range: 0 to 4.1 %). As incorrect responses were extremely rare, we analyzed the reaction times only from trials where the subjects responded correctly and fast enough (within 3 s after the onset of the target word). For the analyses, we averaged data across past words (p1 and p2), and future words (f1 and f2), respectively.

We analyzed the data by means of an rmANOVA including the within-subjects factors SR-mapping (compatible vs. incompatible), Word (past vs. future), Body-orientation of the avatar (front vs. back), and Head-orientation of the avatar (left vs. straight vs. right, from the perspective of the subject). Huynh-Feldt corrected values are reported. For pairwise comparisons, we additionally report Cohen's $d_z$ as a measure of effect size in the dependent measures design. According to Cohen (1988), $d_z$ is defined as $M_d / SD_d$, where $M_d$ is the mean of the differences, and $SD_d$ is the standard deviation of the differences. Alternatively, for factors with two factor-levels in an rmANOVA, $d_z$ can be calculated by dividing the square root of the *F*-value by the square root of the sample size. For pairwise post hoc comparisons, we adjusted the α-level according to the Bonferroni-Holm procedure (Holm, 1979).

#### 2.5.4.1 Spatial-temporal compatibility effect

As illustrated in Figure 2, there was a significant effect of SR-mapping (content of target word – side of response) on reaction time, $F(1,18) = 14.40$; $p = .001$; $_{partial}\eta^2 = .445$; $d_z = 0.87$, indicating delayed responses in trials with incompatible SR-mapping ($M = 1.133$ s; $SD = 0.231$ s) relative to trials with compatible SR-mapping ($M = 0.995$ s; $SD = 0.209$ s).

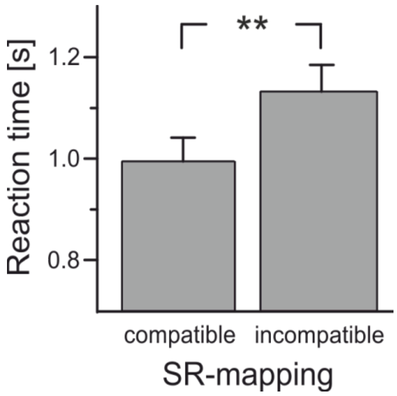


**Figure 2**

Reaction time as a function of compatible vs. incompatible SR-mapping (content of target word – side of response). Error bars indicate standard errors of the mean. ** indicates $p < .01$.

#### 2.5.4.2 Compatibility effects of avatar orientation and processing of temporal word content

There was no interaction between Word and Head-orientation, $F(1.734,31.214) = 0.845$; $p = .425$; $\varepsilon = .867$; $_{partial}\eta^2 = .045$, indicating no compatibility effect between avatar head orientation and word processing (e.g., facilitation of future-related concepts by an avatar head-orientation to the right). As there was also no significant interaction between Word, Body-orientation, and Head-orientation, $F(1.975,35.542) = 0.015$; $p = .984$; $\varepsilon = .987$; $_{partial}\eta^2 = .001$, a potential compatibility effect between avatar head-orientation and word processing did not depend on body orientation (front vs. back) of the avatar. Thus, there was no effect of perspective matching.

#### 2.5.4.3 Effects of gaze

Irrespective of the word to be processed, there was a significant effect of Body-orientation, $F(1,18) = 5.759$; $p = .027$; $_{partial}\eta^2 = .242$; $d_z = 0.55$, revealing a generally faster processing of words in trials with the front-view avatar (*M* = 1.056; *SD* = 0.209) as compared to trials with the back-view avatar (*M* = 1.072; *SD* = 0.200). Inspection of the descriptive data (see Figure 3, Panel A) indicated that responses were particularly fast in trials where the front-view avatar was looking straight, thus gazing at the subject. Although the interaction between Body-orientation and Head-orientation was only marginally significant, $F(2,36) = 2.538$; $p = .093$; ${}_{partial}\eta^2 = .124$, we assumed that the special condition of *direct avatar gaze to the subject* accounts for the effect of Body-orientation on reaction time. In order to test this assumption, we defined Gaze-orientation (direct vs. averted) as a new factor. Note that the two factor-levels of Gaze-orientation accrue from the factor-level combinations Body-orientation *front* and Head-orientation *straight*, and Body-orientation *back* and Head-orientation *straight*, respectively. A paired-samples *t*-test, revealed a significant effect of Gaze-orientation on reaction time, $t(18) = 2.919$; $p = .009$; $d_z = 0.67$, confirming that responses in trials signifying direct avatar gaze to the subject (*M* = 1.046 s; *SD* = 0.214 s) were significantly faster than in trials with avatar gaze oriented to the front (*M* = 1.079 s; *SD* = 0.212 s). In an additional analysis, averted gaze was defined as the aggregate of all combinations of Body-orientation and Head-orientation except of *straight-front* (which was again defined as direct gaze). A paired-samples *t*-test confirmed that responses in trials signifying direct gaze were again significantly faster than in trials with averted gaze (*M* = 1.073 s; *SD* = 0.205 s), $t(18) = 2.792$; $p = .012$; $d_z = 0.64$ (see Figure 3, Panel B). In a third analysis, we compared averted and direct gaze only for the front-avatar condition, where Head-orientation to the left and Head-orientation to the right were defined as averted gaze (*M* = 1.062 s; *SD* = 0.210 s). Here, the effect of Gaze was smaller and did not reach statistical significance, $t(18) = 1.399$; $p = .179$; $d_z = 0.32$.

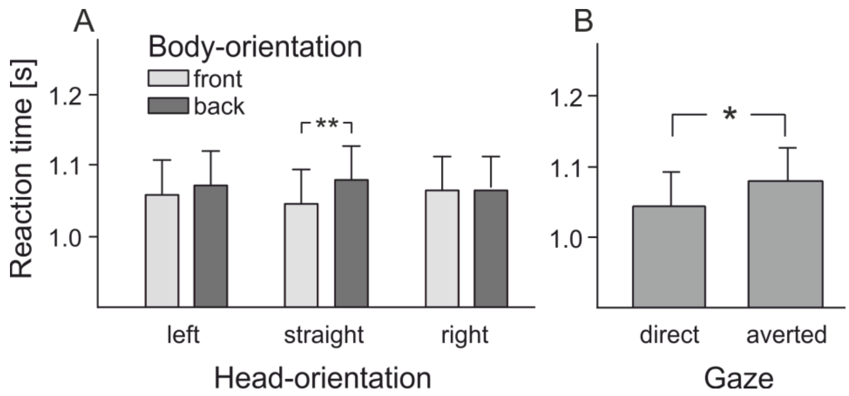


**Figure 3**

Panel A: Reaction time as a function of Body-orientation and Head-orientation. Note that the combination Body-orientation *front* and Head-orientation *straight* signifies direct gaze back to the subject, while the combination Body-orientation *back* and Head-orientation *straight* signifies gaze to the front in the same direction as the subject. Panel B: Reaction time as a function of direct vs. averted Gaze. Note that only the combination of Body-orientation *front* and Head-orientation *straight* signifies direct gaze; all other combinations of Body-orientation and Head-orientation signify averted gaze. Error bars indicate standard errors of the mean. * indicates $p < .05$. ** indicates $p < .01$.

#### 2.5.4.4 *Effects of word and further results*

In the main analysis, the effect of Word on reaction time was significant, $F(1,18) = 36.250$; $p < .001$; ${}_{\text{partial}}\eta^2 = .668$, $d_z = 1.38$, indicating faster processing of future words ($M$ = 1.037 s; $SD$ = 0.201 s) compared to past words ($M$ = 1.091 s; $SD$ = 0.214 s). We further investigated this effect and looked into the reaction-time data for each of the four words separately. As indicated by the descriptive data (see Figure 4), the effect of word (past vs. future) seemed to be driven by delayed reactions in response to p2 (“vorgestern”; the day before yesterday). In a subsequent analysis, we repeated the main rmANOVA but the factor levels of Word (past vs. future) were based solely on p1 (“gestern”, yesterday), and f1 (“morgen”; tomorrow), respectively. Here, the main effect of Word disappeared, $F(1,18) = 0.058$; $p = .813$; ${}_{\text{partial}}\eta^2 = .003$. All interactions including the factor Word remained non-significant, as in the main analysis.

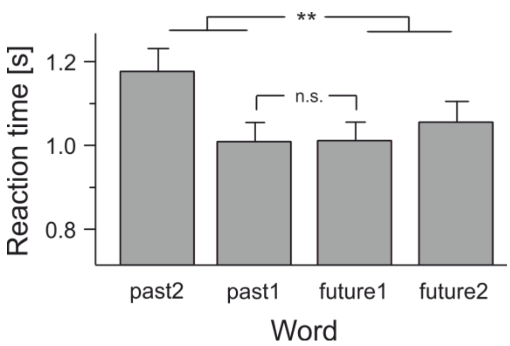


**Figure 4**

Reaction time as a function of Word (past2: “vorgestern” – the day before yesterday; past1: “gestern” – yesterday; future1: “morgen” – tomorrow; future2: “übermorgen” – the day after tomorrow). Error bars indicate standard errors of the mean. ** indicates $p < .01$. n.s.: non-significant.

In the main analysis, the interaction between Body-orientation, Head-orientation, and SR-mapping was also statistically significant, $F(2,36) = 5.674$; $p = .007$; ${}_{\text{partial}}\eta^2 = .240$, again pointing to an effect of gaze that has been substantiated by the analyses reported above. Importantly, this effect was not modulated by the content of the word to be processed (insignificant interaction between Body-orientation, Head-orientation, SR-mapping, and

Word), again indicating no facilitation of, for example, past-related concepts by an avatar head orientation to the left under specific circumstances (e.g., only in trials with compatible SR-mapping). The main effect of Head-orientation and all remaining interactions were statistically insignificant.

### 2.5.5 Discussion

#### *2.5.5.1 Manually reflected mental time*

Our first hypothesis (in relation to manually reflected mental time) – that there would be a spatial-temporal association of response codes (STEARC; Ishihara et al., 2008) when processing temporal concepts auditorily (that is, without possible reading direction confounds) – was clearly confirmed. According to the notion of a manually embodied horizontal mental time line, we assumed that compatibility between the spatial representation of a temporal concept and the side of the motor response facilitates action, whereas incompatibility was assumed to impede responses. In line with previous studies (e.g., Ouellet et al., 2010; Santiago et al., 2007), our results strongly support this hypothesis. Reaction times were small when past-related auditory words had to be indicated by a left hand response and future-related auditory words by a right hand response. Reaction times increased when we reversed this mapping. This compatibility effect was large in size and clearly supports the notion of a strong spatial representation of time along the horizontal axis.

#### *2.5.5.2 Mental time is not interpersonally reflected in head orientation*

Our second hypothesis (in relation to interpersonally reflected mental time) was that an avatar's head-orientation to the left should facilitate past processing, as the avatar is "pointing" towards the past of the horizontal mental time line, while an avatar's head-orientation to the right should facilitate future processing. We assumed that this effect would be pronounced when the avatar's perspective (back-view avatar), and thus its horizontal mental time line, matches the perspective/the mental time line of the subject. This hypothesis is clearly refuted. Our results provided no indications for compatibility effects between the avatar's head- and body-orientation and the (spatial) processing of temporal words. We do not know if head orientation is able to assume a role when it is perceived within the context of a pointing gesture and thus becomes part of the task. In the present setup, where head orientation had no consequence for the response action, mental time was not interpersonally affected.

#### *2.5.5.3 Effects of gaze orientation on word processing time*

Beside the results in relation to our hypotheses, we found a more basic effect of gaze direction on reaction time. When the avatar looked toward and faced the subject, reaction times to the words shortened (irrespective of word content). One interpretation might be that the increased arousal induced by direct gaze has sped up the reactions (e.g., Helminen et al., 2011). The significant effect of avatar gaze direction (direct vs. averted) supports this notion. It is well established that direct gaze is a powerful stimulus for human beings. This effect, sometimes referred to as the "eye contact effect" (Senju & Johnson, 2009), reportedly speeds up face detection (Senju, Hasegawa, & Tojo, 2005) and gender discrimination (Macrae, Hood, Milne, Rowe, & Mason, 2002). However, the effects of direct gaze are less well known for more *abstract* domains of cognitive processing, such as temporal cognition.

Our main analysis also indicated that subjects reacted faster in response to the future words "morgen" (tomorrow) and "übermorgen" (the day after tomorrow) as compared to the past words "gestern" (yesterday) and "vorgestern" (the day before yesterday). However, this effect was mainly driven by delayed responses to "vorgestern" and completely disappeared when comparing "morgen" and "gestern". Accordingly, the results do not provide evidence for a generally faster processing of future relative to past words. Possibly, effects of word length (Baddeley et al., 1975; Ellis, 1992) and frequency, and therefore accessibility, have contributed to the delayed reactions in response to "vorgestern" and "übermorgen" as compared to "gestern", and "morgen". A possible explanation why categorizing „vorgestern" as past takes even longer than categorizing „übermorgen" as future could be that the morpheme "vor" of "vorgestern" might cause a slight temporal confusion as this spatial morpheme (meaning *in front of*) often has a metaphorical future meaning in German (for instance "*vor*aus denken" means "to think forward", "to think about the future"). Accordingly, the temporal ambiguity of "vor-" may have resulted in particularly delayed responses to the past word "vorgestern".

Beside the main effect of Word, it should be noted, however, that the analysis including only the words "morgen" and "gestern" yielded similar results as the main analysis including all four time words. Therefore, possible differences in the processing of the four time words did not affect the assumed but absent compatibility effects between avatar head orientation and word content.

#### *2.5.5.4 Conclusion*

In conclusion, we have established a robust STEARC effect for the processing of auditory stimuli (temporal adverbials), while controlling for potential confounds of reading direction inherent in any written-language stimulus presentation. Against our second hypothesis, processing of past/future-related words was not facilitated by a leftward/rightward orientation of an avatar's head, not even when subject and avatar were oriented in a compatible way (subject behind avatar both facing the screen). Instead, direct gaze of an avatar sped up the processing of time words across all conditions. Thus, our results do not only support the notion of mental time as being embodied, they also indicate that interpersonal cues need to be factored into the equation, albeit in a more straightforward manner than previously thought.

### 2.5.6 Acknowledgments

We wish to thank Agnes Münch for programming and technical assistance, and Larissa Reis for help with the data collection. We also thank Andrea Bizzeti (Phonogram Archive of the University of Zürich) for recording the audio word stimuli.

# 3. General Discussion

The five parts as presented in the Main section represent recent theoretical and empirical work on research questions in the field of subjective time. While focusing on specific aspects and influencing factors, such as time perception in depression, the results reveal some fundamental characteristics of time perception, temporal processing, and spatial representation of time and extend the current research on subjective time on a conceptual level.

First, the meta-analysis on time perception in depressive patients (part 2.1) provides clear evidence that judgments of passage and judgments of duration represent independent aspects of time perception. While time is passing slower for depressive patients than for healthy control subjects, duration judgments, such as verbal estimates, time productions, time reproductions, and duration discrimination, do not systematically differ between patients and control subjects. Accordingly, factors that influence the subjective passage of time do not necessarily affect the perception of duration. Therefore, passage and duration need to be conceptualized as disentangled temporal dimensions, and the application of the pacemaker-accumulator model should be restricted to the mental representation of duration and not extend to the perception of time passage. This view is also supported in a recent study by Wearden (2015) who reported that duration judgments and time passage ratings are uncorrelated. Interestingly, however, a positive correlation between judgments of long duration in the range of several minutes and time passage ratings has recently been observed (Droit-Volet, Trahanias, & Maniadakis, 2017). These results suggest some kind of overlap in the mental representation of longer duration and time passage but a clear dissociation between the concepts of passage and duration in the range of seconds. And even with regard to duration judgment tasks, especially the meta-analysis on time perception and temporal processing in schizophrenic patients shows, as expected, that different tasks require (and measure) different aspects of temporal information processing (Gil & Droit-Volet, 2011). For example, while schizophrenic patients overestimated duration in verbal time estimation and time production tasks relative to healthy controls, they tended to underestimate the duration of comparison intervals in temporal bisection (duration discrimination) tasks. Beside the basic mental representation of duration, the demands of memory (and motor) processes differ between verbal time estimation, time production, and duration discrimination tasks. In order to estimate or produce a given time interval, a subject has to represent duration samples from long term memory, whereas duration discrimination is rather working memory demanding. In the light of this notion, the observed differences in performance between tasks are not only explainable but they are also to be expected.

Second, as the passage of time is represented spatially, the concept of passing time may be more closely related to the concept of a mental time line than to the concept of duration. As a result of part 2.5 and as shown by previous empirical research (e.g., Bonato et al., 2012; Ishihara et al., 2008), the spatiotemporal association of response codes reliably shows that the passage of subjective time is reflected in a mental time line running spatially from left (past) to right (future). This representation of past on the left and future on the right may also be interpreted in terms of future time moving from the right, passing by as present, and becoming past when moving further to the left.[6] The relationship between the passage of time and its spatialized mental representation should be addressed by future research. For instance, a slower passage of time as observed in patients with depression may be accompanied by an expansion of the subject's spatiotemporal representation, which could be reflected in a pronounced STEARC effect in patients. A similar issue may be the relationship

[6] In this regard, please refer to Stocker (2012), who provides an elaborate discussion on how spatiotemporal relations are reflected in (metaphorical) language.

between duration judgments and temporal processing abilities. Even though the two concepts represent different aspects of subjective time (the representation of duration is not necessary for detecting temporal order), the results from our meta-analysis on time perception and temporal processing in patients with schizophrenia indicate similar impairment in both domains (enhanced variability of temporal judgments in patients). This result points to a connection between temporal processing and the representation of duration, which needs to be investigated in more detail by future research. Especially in this regard, neuroscientific research may help to understand the underlying processes and (shared) mechanisms of time perception and temporal processing.

Third, the meta-analytical review on time perception and temporal processing in schizophrenic patients (part 2.2) shows that measures of accuracy and measures of precision reflect independent aspects of subjective time. Whereas the precision of temporal judgments is clearly reduced in schizophrenic patients, mean duration judgments (accuracy) do not differ between schizophrenic patients and healthy control subjects. Similarly, we have obtained different results for measures of accuracy and precision in the experiments on potential effects of gaze direction on duration judgments (part 2.3) as well as in the study on effects of counting on time production (part 2.4). Whereas the precision of time production was enhanced when subjects used a counting strategy, the accuracy of temporal judgments dropped. Therefore, factors that have an influence on the variability of temporal judgments do not necessarily affect their accuracy in the same way. Although this had been known before (e.g., Lhamon & Goldstone, 1956), previous research on time perception and temporal processing did often not consider accuracy and precision as independent dimensions of performance in temporal tasks. Future research should take these considerations seriously. With regard to the pacemaker-accumulator model of time perception, the aspect of temporal precision could be illustrated easily. Beside the mean speed of the pacemaker, which is indicated by the density of emitted pulses, the variability of clock speed (pulse-to-pulse variability) could be illustrated by varying distances between the single pulses (see Figure 1).

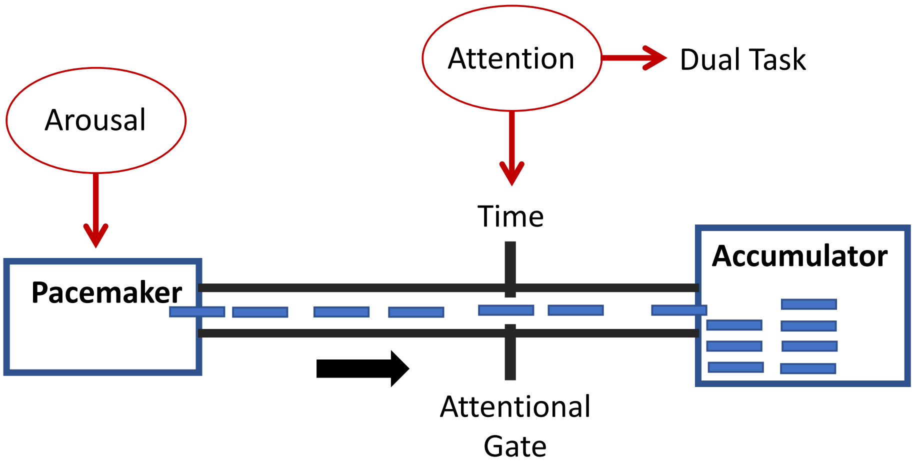


**Figure 1.** Modified model of the internal clock. The precision (variability) of duration representation is illustrated by means of varying pulse-to-pulse distances.

The experiments on the effects of chronometric counting on duration judgments (part 2.4) provided surprising results. When it comes to longer time intervals in the range of 30 and 60 s, intuitive duration judgments without any timing strategy lead to more accurate interval productions than timing with a counting strategy. While studies on time perception usually focus on specific factors that lead to over- or underestimation of duration, the nature of intuitive timing in the absence of specific stimuli may be an interesting line of future research.

As shown by the present work and by research on subjective time in general, time perception and temporal processing are influenced by numerous factors, including, for example, emotional content of presented tones (Noulhiane, Mella, Samson, Ragot, & Pouthas, 2007) or faces (Droit-Volet & Gil, 2009), body temperature (Rammsayer, 1997; Wearden & Pentonvoak, 1995), level of dopaminergic activity in specific (time perception-related) parts of the brain (Rammsayer, 2009), age (Bisson, Tobin, & Grondin, 2012), attentional distraction

(Block & Zakay, 1996), etc. This list could be continued for pages. In order to systemize the large body of literature on subjective time, it is important to aggregate and reduce the different influencing factors to more basic and established (psychological) concepts. In this regard, the most important basic concepts are arousal and attention, as presented in the General introduction section (part 1.3.2.2). According to the notion that these two concepts are central for time perception and temporal processing, a given task or environment either affects the subject's arousal level, which accelerates or decelerates the pacemaker of the internal clock, or it affects the amount of attention paid to the time, which, in turn, improves or impedes the accumulation of pulses. However, the influences of any additional factors that are not reducible to arousal and attention need to be factored into the equation of a comprehensive model of time perception and temporal processing. Motion may be one such factor, as moving objects have been shown to induce a dilation of perceived duration that is not simply explainable in terms of arousal or attentional processes (Au, Ono, & Watanabe, 2012; Yamamoto & Miura, 2016).

To date, psychology is far from presenting a unified theory of subjective time. In contrast to the study of objective (physical) time, which has been formulized successfully for example within the scope of general relativity (e.g., Bennett et al., 2010), the study of subjective time still lacks a coherent conceptual system and authors do not always agree on the meaning and relations of specific concepts, such as passage of time and duration. In this regard, a broad conceptual system, as presented in Figure 1 of part 1.1, can provide the frame for empirical studies that investigate the relationship between different aspects of subjective time, such as passage and duration. Maybe, these concepts need to be treated fully independent in the sense of different dimensions of subjective time. And maybe, subjective time turns out to be a mere collective term that solely connects rather independent temporal concepts on a linguistic level.

Moreover, the different aspects of subjective time, such as duration judgments and the temporal processing ability, are influenced by numerous factors that interact in complex ways, but most research focuses on rather isolated factors in single experiments without aspirations to synthesize results and bring forward theory. It is important to leave behind this realm of research on isolated time perception phenomena. Instead, future research needs to synthesize the results from the large body of literature on the different aspects of subjective time within a broader frame. This should be done by means of meta-analytical and theoretical work. Based on such work, new and theoretically driven hypotheses will be formulated, which, again, need to be tested in empirical studies in a systematic way. For now, it remains a challenging project in the study of subjective time to successfully reduce the various influencing factors to more basic dimensions, such as arousal and attention, and to agree on basic concepts in order to come closer to a broad and comprehensive theory of subjective time.

# 4. REFERENCES


Allan, L. G., & Gibbon, J. (1991). Human bisection at the geometric mean. *Learning and Motivation, 22*(1-2), 39-58. doi: 10.1016/0023-9690(91)90016-2

Allman, M. J., & Meck, W. H. (2012). Pathophysiological distortions in time perception and timed performance. *Brain, 135*, 656-677. doi: 10.1093/brain/awr210

Andreasen, N. C., Nopoulos, P., O'Leary, D. S., Miller, D. D., Wassink, T., & Flaum, L. (1999). Defining the phenotype of schizophrenia: Cognitive dysmetria and its neural mechanisms. *Biological Psychiatry, 46*(7), 908-920. doi: 10.1016/s0006-3223(99)00152-3

Andreasen, N. C., & Pierson, R. (2008). The role of the cerebellum in schizophrenia. *Biological Psychiatry, 64*(2), 81-88. doi: 10.1016/j.biopsych.2008.01.003

Angrilli, A., Cherubini, P., Pavese, A., & Manfredini, S. (1997). The influence of affective factors on time perception. *Perception & Psychophysics, 59*(6), 972-982. doi: 10.3758/bf03205512

Argyle, M., & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry, 28*(3), 289-304. doi: 10.2307/2786027

Argyle, M., Ingham, R., Alkema, F., & McCallin, M. (1972). The different functions of gaze. *Semiotica, 7*, 19-32.

Au, R. K. C., Ono, F., & Watanabe, K. (2012). Time dilation induced by object motion is based on spatiotopic but not retinotopic positions. *Frontiers in Psychology, 3*. doi: 10.3389/fpsyg.2012.00058

Augustinus, A. (ca. 400/ 2000). *Confessiones XI: Quid est ergo tempus? [Confessiones: What is time?]*. Hamburg: Meiner.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of verbal learning and verbal behavior, 14*(6), 575-589.

Bangert, A. S., Reuter-Lorenz, P. A., & Seidler, R. D. (2011). Dissecting the clock: Understanding the mechanisms of timing across tasks and temporal intervals. *Acta Psychologica, 136*(1), 20-34. doi: 10.1016/j.actpsy.2010.09.006

Bech, P. (1975). Depression: Influence on time estimation and time experience. *Acta Psychiatrica Scandinavica, 51*(1), 42-50. doi: 10.1111/j.1600-0447.1975.tb00211.x

Beck, A. T., Erbaugh, J., Ward, C. H., Mock, J., & Mendelsohn, M. (1961). An Inventory for Measuring Depression. *Archives of General Psychiatry, 4*(6), 561-&.

Belsley, D. A., Kuh, E., & Welsch, R. A. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. Hoboken, N.J.: Wiley.

Benazzi, F., & Akiskal, H. S. (2003). The dual factor structure of self-rated MDQ hypomania: energized-activity versus irritable-thought racing. *Journal of Affective Disorders, 73*(1-2), 59-64. doi: 10.1016/s0165-0327(02)00333-6

Bennett, J. O., Donahue, M., Schneider, N., & Voit, M. (2010). *Astronomie: die kosmische Perspektive [Astronomy: the cosmic perspective]*. München: Pearson.

Bergen, B. K., & Lau, T. T. C. (2012). Writing direction affects how people map space onto time. *Frontiers in psychology, 3*.

Bertamini, M., Ogden, R., Rampone, G., & Makin, A. (2013). Symmetry is forever. *Perception 42 ECVP Abstract Supplement, 42*, 234.

Biermann, T., Kreil, S., Groemer, T. W., Maihofner, C., Richter-Schmiedinger, T., Kornhuber, J., & Sperling, W. (2011). Time perception in patients with major depressive disorder during vagus nerve stimulation. *Pharmacopsychiatry, 44*(5), 179-182. doi: 10.1055/s-0031-1280815

Bisson, N., Tobin, S., & Grondin, S. (2012). Prospective and retrospective time estimates of children: A comparison based on ecological tasks. *PLOS One, 7*(3), e33049. doi: 10.1371/journal.pone.0033049

Blewett, A. E. (1992). Abnormal subjective time experience in depression. *British Journal of Psychiatry, 161*, 195-200.

Block, R. A. (1990). Models of Psychological Time. In R. A. Block (Ed.), *Cognitive models of psychological time* (pp. 1-35). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Block, R. A., & Zakay, D. (1996). Models of psychological time revisited. In H. Helfrich (Ed.), *Time and Mind* (pp. 171-195).

Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review, 4*(2), 184-197. doi: 10.3758/bf03209393

Boehnke, S. E., & Phillips, D. P. (2005). The relation between auditory temporal interval processing and sequential stream segregation examined with stimulus laterality differences. *Perception and Psychophysics, 67*(6), 1088-1101.

Bojanovsky, J. (1965). Time perception in endogenous and psychogenic depressions. *Activitas Nervosa Superior, 7*(2), 196.

Bojanovsky, J., & Tölle, R. (1973). Influence of antidepressive therapy on disturbed time estimation of depressives. *Psychiatria Clinica, 6*(6), 321-329.

Bolbecker, A. R., Hong, S. L., Kent, J. S., Forsyth, J. K., Klaunig, M. J., Lazar, E. K., . . . Hetrick, W. P. (2011). Paced finger-tapping abnormalities in bipolar disorder indicate timing dysfunction. *Bipolar Disorders, 13*(1), 99-110. doi: 10.1111/j.1399-5618.2011.00895.x

Bolbecker, A. R., Westfall, D. R., Howell, J. M., Lackner, R. J., Carroll, C. A., O'Donnell, B. F., & Hetrick, W. P. (2014). Increased timing variability in schizophrenia and bipolar disorder. *Plos One, 9*(5). doi: 10.1371/journal.pone.0097964

Bonato, M., Zorzi, M., & Umilta, C. (2012). When time is space: Evidence for a mental time line. *Neuroscience and Biobehavioral Reviews, 36*(10), 2257-2273. doi: 10.1016/j.neubiorev.2012.08.007

Bonnot, O., de Montalembert, M., Kermarrec, S., Botbol, M., Walter, M., & Coulon, N. (2011). Are impairments of time perception in schizophrenia a neglected phenomenon? *Journal of Physiology-Paris, 105*(4-6), 164-169. doi: 10.1016/j.jphysparis.2011.07.006

Bonnot, O., & Georgieff, N. (2000). Cognitive process of time perception. Interest in studies on schizophrenia. *Annales Medico-Psychologiques, 158*(6), 475-482.

Bourdet, C., Brochard, R., Rouillon, F., & Drake, C. (2003). Auditory temporal processing in schizophrenia: high level rather than low level deficits? *Cogn Neuropsychiatry, 8*(2), 89-106.

Boyarskaya, E., Sebastian, A., Bauermann, T., Hecht, H., & Tuescher, O. (2015). The Mona Lisa Effect: Neural Correlates of Centered and Off-Centered Gaze. *Human Brain Mapping, 36*(2), 619-632. doi: 10.1002/hbm.22651

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment mannequin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59. doi: 10.1016/0005-7916(94)90063-9

Braus, D. F. (2002). Temporal perception and organisation, neuronal synchronisation and schizophrenia. *Fortschritte Der Neurologie Psychiatrie, 70*(11), 591-600. doi: 10.1055/s-2002-35172

Broadhurst, A. (1969). Time estimation related to personality cognitive speed and schizophrenia. *Life Sciences 8*(2P2), 69-78. doi: 10.1016/0024-3205(69)90119-2

Brown, G. D. A., McCormack, T., Smith, M., & Stewart, N. (2005). Identification and bisection of temporal durations and tone frequencies: Common models for temporal and nontemporal stimuli. *Journal of Experimental Psychology-Human Perception and Performance, 31*(5), 919-938. doi: 10.1037/0096-1523.31.5.919

Brown, S. W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. *Perception and Psychophysics, 59*(7), 1118-1140.

Brown, S. W. (2008). Time and attention: Review of the literature. In S. Grondin (Ed.), *Psychology of time* (pp. 111-138). Bingley: Emerald.

Brugger, P. (2002). Reflective mirrors: perspective-taking in autoscopic phenomena. *Cognitive Neuropsychiatry, 7*(3), 179-194. doi: 10.1080/13546800244000076

Bschor, T., Ising, M., Bauer, M., Lewitzka, U., Skerstupeit, M., Muller-Oerlinghausen, B., & Baethge, C. (2004). Time experience and time judgment in major depression, mania and healthy subjects. A controlled study of 93 subjects. *Acta Psychiatrica Scandinavica, 109*(3), 222-229.

Capa, R. L., Duval, C. Z., Blaison, D., & Giersch, A. (2014). Patients with schizophrenia selectively impaired in temporal order judgments. *Schizophrenia Research, 156*(1), 51-55. doi: 10.1016/j.schres.2014.04.001

Carlson, V. R., & Feinberg, I. (1968). Individual variations in time judgment and the concept of an internal clock. *Journal of Experimental Psychology, 77*(4), 631-640. doi: 10.1037/h0026048

Carroll, C. A., Boggs, J., O'Donnell, B. F., Shekhar, A., & Hetrick, W. P. (2008). Temporal processing dysfunction in schizophrenia. *Brain and Cognition, 67*(2), 150-161. doi: 10.1016/j.bandc.2007.12.005

Carroll, C. A., O'Donnell, B. F., Shekhar, A., & Hetrick, W. P. (2009a). Timing dysfunctions in schizophrenia as measured by a repetitive finger tapping task. *Brain and Cognition, 71*(3), 345-353. doi: 10.1016/j.bandc.2009.05.009

Carroll, C. A., O'Donnell, B. F., Shekhar, A., & Hetrick, W. P. (2009b). Timing dysfunctions in schizophrenia span from millisecond to several-second durations. *Brain and Cognition, 70*(2), 181-190. doi: 10.1016/j.bandc.2009.02.001

Casasanto, D., & Bottini, R. (2014). Mirror Reading Can Reverse the Flow of Time. *Journal of Experimental Psychology-General, 143*(2), 473-479. doi: 10.1037/a0033297

Champagne, J., & Fortin, C. (2008). Attention sharing during timing: Modulation by processing demands of an expected stimulus. *Perception and Psychophysics, 70*(4), 630-639. doi: 10.3758/pp.70.4.630

Cheng, R. K., Ali, Y. M., & Meck, W. H. (2007). Ketamine "unlocks" the reduced clock-speed effects of cocaine following extended training: Evidence for dopamine-glutamate interactions in timing and time perception. *Neurobiology of Learning and Memory, 88*(2), 149-159. doi: 10.1016/j.nlm.2007.04.005

Ciullo, V., Spalletta, G., Caltagirone, C., Jorge, R. E., & Piras, F. (2015). Explicit Time Deficit in Schizophrenia: Systematic Review and Meta-Analysis Indicate it Is Primary and Not Domain Specific. *Schizophrenia Bulletin*, sbv104.

Clausen, J. (1950). An evaluation of experimental methods of time judgment. *Journal of Experimental Psychology, 40*(6), 756-761. doi: 10.1037/h0056354

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J.: L. Erlbaum Associates.

Cook, R., Ayhan, I., Lai, A., & Johnston, A. (2011). Glance, glimpse or stare? The discrimination of gaze duration. *Journal of Vision, 11*(11), 1225-1225. doi: 10.1167/11.11.1225

Coull, J. T., Cheng, R. K., & Meck, W. H. (2011). Neuroanatomical and Neurochemical Substrates of Timing. *Neuropsychopharmacology, 36*(1), 3-25. doi: 10.1038/npp.2010.113

Davalos, D. B., Kisley, M. A., & Freedman, R. (2005). Behavioral and electrophysiological indices of temporal processing dysfunction in schizophrenia. *Journal of Neuropsychiatry and Clinical Neurosciences, 17*(4), 517-525. doi: 10.1176/appi.neuropsych.17.4.517

Davalos, D. B., Kisley, M. A., Polk, S. D., & Ross, R. G. (2003). Mismatch negativity in detection of interval duration deviation in schizophrenia. *Neuroreport, 14*(9), 1283-1286. doi: 10.1097/01.wnr.0000081874.45938.98

Davalos, D. B., Kisley, M. A., & Ross, R. G. (2002). Deficits in auditory and visual temporal perception in schizophrenia. *Cognitive neuropsychiatry, 7*(4), 273-282. doi: 10.1080/13546800143000230

Davalos, D. B., Kisley, M. A., & Ross, R. G. (2003). Effects of interval duration on temporal processing in schizophrenia. *Brain and Cognition, 52*(3), 295-301. doi: 10.1016/s0278-2626(03)00157-x

Davalos, D. B., Rojas, D. C., & Tregellas, J. R. (2011). Temporal processing in schizophrenia: Effects of task-difficulty on behavioral discrimination and neuronal responses. *Schizophrenia Research, 127*(1-3), 123-130. doi: 10.1016/j.schres.2010.06.020

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology-General, 122*(3), 371-396. doi: 10.1037/0096-3445.122.3.371

Delevoye-Turrell, Y., Wilquin, H., & Giersch, A. (2012). A ticking clock for the production of sequential actions: Where does the problem lie in schizophrenia? *Schizophrenia Research, 135*(1-3), 51-54. doi: 10.1016/j.schres.2011.12.020

Densen, M. E. (1977). Time perception and schizophrenia. *Perceptual and Motor Skills, 44*(2), 436-438.

Dilling, C. A., & Rabin, A. I. (1967). Temporal experience in depressive states and schizophrenia. *Journal of Consulting Psychology, 31*(6), 604-608. doi: 10.1037/h0025160

Dimberg, U., & Ohman, A. (1996). Behold the wrath: Psychophysiological responses to facial stimuli. *Motivation and Emotion, 20*(2), 149-182. doi: 10.1007/bf02253869

Doi, H., & Shinohara, K. (2009). The perceived duration of emotional face is influenced by the gaze direction. *Neuroscience Letters, 457*(2), 97-100. doi: 10.1016/j.neulet.2009.04.004

Droit-Volet, S. (2013). Time perception, emotions and mood disorders. *Journal of Physiology-Paris, 107*(4), 255-264. doi: 10.1016/j.jphysparis.2013.03.005

Droit-Volet, S., & Gil, S. (2009). The time-emotion paradox. *Philosophical Transactions of the Royal Society B-Biological Sciences, 364*(1525), 1943-1953. doi: 10.1098/rstb.2009.0013

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences, 11*(12), 504-513. doi: 10.1016/j.tics.2007.09.008

Droit-Volet, S., Trahanias, P., & Maniadakis, M. (2017). Passage of time judgments in everyday life are not related to duration judgments except for long durations of several minutes. *Acta psychologica, 173*, 116-121. doi: 10.1016/j.actpsy.2016.12.010

Droit-Volet, S., & Wearden, J. (2002). Speeding up an internal clock in children? Effects of visual flicker on subjective duration. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology, 55*(3), 193-211. doi: 10.1080/02724990143000252

Dubois, F. S. (1954). The sense of time and its relation to psychiatric illness. *American Journal of Psychiatry, 111*(1), 46-51.

Ebert, D., & Lammers, C. H. (1997). The central dopaminergic system and depression. *Nervenarzt, 68*(7), 545-555. doi: 10.1007/s001150050159

Eikmeier, V., Alex-Ruf, S., Maienborn, C., & Ulrich, R. (2015). How Strongly Linked Are Mental Time and Space Along the Left-Right Axis? *Journal of Experimental Psychology-Learning Memory and Cognition, 41*(6), 1878-1883. doi: 10.1037/xlm0000129

Eikmeier, V., Hoppe, D., & Ulrich, R. (2015). Response mode does not modulate the space-time congruency effect: Evidence for a space-time mapping at a conceptual level. *Acta Psychologica, 156*, 162-167. doi: 10.1016/j.actpsy.2014.10.008

Ellis, N. (1992). Linguistic relativity revisited: The bilingual word-length effect in working memory during counting, remembering numbers, and mental calculation. *Advances in psychology, 83*, 137-155.

Ellsworth, P. C., & Langer, E. J. (1976). Staring and approach: Interpretation of stare as a nonspecific activator. *Journal of Personality and Social Psychology, 33*(1), 117-122. doi: 10.1037//0022-3514.33.1.117

Elvevag, B., Brown, G. D. A., McCormack, T., Vousden, J. I., & Goldberg, T. E. (2004). Identification, of tone duration, line length, and letter position: An experimental approach to timing and working memory deficits in schizophrenia. *Journal of Abnormal Psychology, 113*(4), 509-521. doi: 10.1037/0021-843x.113.4.509

Elvevag, B., McCormack, T., Gilbert, A., Brown, G. D. A., Weinberger, D. R., & Goldberg, T. E. (2003). Duration judgements in patients with schizophrenia. *Psychological Medicine, 33*(7), 1249-1261. doi: 10.1017/s0033291703008122

Ewbank, M. P., Jennings, C., & Calder, A. J. (2009). Why are you angry with me? Facial expressions of threat influence perception of gaze direction. *Journal of vision, 9*(12), 16.11-17. doi: 10.1167/9.12.16

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi: 10.3758/bf03193146

Fechner, G. T. (1860). *Elemente der Psychophysik [Elements of psychophysics]*. Leipzig: Breitkopf und Härtel.

Foucher, J. R., Lacambre, M., Pham, B. T., Giersch, A., & Elliott, M. A. (2007). Low time resolution in schizophrenia Lengthened windows of simultaneity for visual, auditory and bimodal stimuli. *Schizophrenia Research, 97*(1-3), 118-127. doi: 10.1016/j.schres.2007.08.013

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The Psychology of Music*. San Diego: Academic Press, inc. (London) Ltd.

Franck, N., Posada, A. S., Pichon, S., & Haggard, P. (2005). Altered subjective time of events in schizophrenia. *Journal of Nervous and Mental Disease, 193*(5), 350-353. doi: 10.1097/01.nmd.0000161699.76032.09

Gallagher, S. (2012). Time, emotion, and depression. *Emotion Review, 4*(2), 127-132. doi: 10.1177/1754073911430142

Gan, T., Wang, N. Y., Zhang, Z. J., Li, H., & Luo, Y. J. (2009). Emotional influences on time perception: evidence from event-related potentials. *Neuroreport, 20*(9), 839-843. doi: 10.1097/WNR.0b013e32832be7dc

Gautier, T., & Droit-Volet, S. (2002). Attentional distraction and time perception in children. *International Journal of Psychology, 37*(1), 27-34. doi: 10.1080/00207590143000324

Geller, B., Zimerman, B., Williams, M., DelBello, M., Frazier, J., & Beringer, L. (2002). Phenomenology of prepubertal and early adolescent bipolar disorder: Examples of elated mood, grandiose behaviors, decreased need for sleep, racing thoughts and hypersexuality. *Journal of Child and Adolescent Psychopharmacology, 12*(1), 3-9. doi: 10.1089/10445460252943524

Getty, D. J. (1976). Counting processes in human timing. *Perception & Psychophysics, 20*(3), 191-197. doi: 10.3758/bf03198600

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*(3), 279-325.

Gibbon, J. (1992). Ubiquity of scalar timing with a Poisson clock. *Journal of Mathematical Psychology, 36*(2), 283-293.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences, 423*(May), 52-77.

Gil-Da-Costa, R., Stoner, G. R., Fung, R., & Albright, T. D. (2013). Nonhuman primate model of schizophrenia using a noninvasive EEG method. *Proceedings of the National Academy of Sciences of the United States of America, 110*(38), 15425-15430. doi: 10.1073/pnas.1312264110

Gil, S., & Droit-Volet, S. (2009). Time perception, depression and sadness. *Behavioural Processes, 80*(2), 169-176. doi: 10.1016/j.beproc.2008.11.012

Gil, S., & Droit-Volet, S. (2011). "Time flies in the presence of angry faces" ... depending on the temporal task used! *Acta Psychologica, 136*(3), 354-362. doi: 10.1016/j.actpsy2010.12.010

Gil, S., & Droit-Volet, S. (2012). Emotional time distortions: The fundamental role of arousal. *Cognition & Emotion, 26*(5), 847-862. doi: 10.1080/02699931.2011.625401

Gilliland, A. R., & Martin, R. (1940). Some factors in estimating short time intervals. *Journal of Experimental Psychology, 27*(3), 243.

Goldstone, S., & Lhamon, W. T. (1974). Studies of auditory-visual differences in human time judgment: 1. Sounds are judged longer than lights. *Perceptual and Motor Skills, 39*(1), 63-82.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Grinker, J., Glucksman, M. L., Hirsch, J., & Viseltear, G. (1973). Time perception as a function of weight reduction - differentiation based on age at onset of obesity. *Psychosomatic Medicine, 35*(2), 104-111.

Grondin, S. (2008). *Psychology of time*. Bingley: Emerald.

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention Perception & Psychophysics, 72*(3), 561-582. doi: 10.3758/app.72.3.561

Grondin, S. (2012). Violation of the scalar property for time perception between 1 and 2 seconds: Evidence from interval discrimination, reproduction, and categorization. *Journal of Experimental Psychology-Human Perception and Performance, 38*(4), 880-890. doi: 10.1037/a0027188

Grondin, S., Laflamme, V., & Gontier, E. (2014). Effect on perceived duration and sensitivity to time when observing disgusted faces and disgusting mutilation pictures. *Attention Perception & Psychophysics, 76*(6), 1522-1534. doi: 10.3758/s13414-014-0682-7

Grondin, S., Laflamme, V., & Mioni, G. (2015). Do not count too slowly: evidence for a temporal limitation in short-term memory. *Psychonomic bulletin & review, 22*(3), 863-868.

Grondin, S., & McAuley, J. D. (2009). Duration discrimination in crossmodal sequences. *Perception, 38*(10), 1542-1559. doi: 10.1068/p6359

Grondin, S., Meilleur-Wells, G., & Lachance, R. (1999). When to start explicit counting in a time-intervals discrimination task: A critical point in the timing process of humans. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 993-1004.

Grondin, S., Ouellet, B., & Roussel, M. E. (2004). Benefits and limits of explicit counting for discriminating temporal intervals. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 58*(1), 1-12. doi: 10.1037/h0087436

Grondin, S., Pouthas, V., Samson, S., & Roy, M. (2006). Mechanisms and disorders connected to the adaptation to the time. *Canadian Psychology-Psychologie Canadienne, 47*(3), 170-183. doi: 10.1037/cp2006007

Hamilton, M. (1960). A Rating Scale for Depression. *Journal of Neurology Neurosurgery and Psychiatry, 23*(1), 56-62. doi: Doi 10.1136/Jnnp.23.1.56

Harrison, C., Binetti, N., Coutrot, A., Mareschal, I., & Johnston, A. (2015). Individual differences in preference for mutual gaze duration. . *Journal of Vision, - [Abstract]*.

Hartmann, M., Martarelli, C. S., Mast, F. W., & Stocker, K. (2014). Eye movements during mental time travel follow a diagonal line. *Consciousness and cognition, 30*, 201-209.

Hawkins, W. L., French, L. C., Crawford, B. D., & Enzle, M. E. (1988). Depressed affect and time perception. *Journal of Abnormal Psychology, 97*(3), 275-280. doi: 10.1037/0021-843x.97.3.275

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*, 39-65.

Helminen, T. M., Kaasinen, S. M., & Hietanen, J. K. (2011). Eye contact and arousal: The effects of stimulus duration. *Biological Psychology, 88*(1), 124-130. doi: 10.1016/j.biopsycho.2011.07.002

Hicks, R. E., & Allen, D. A. (1979a). Counting eliminates the repetition effect in judgments of temporal duration. *Acta Psychologica, 43*(5), 361-366. doi: 10.1016/0001-6918(79)90030-1

Hicks, R. E., & Allen, D. A. (1979b). Repetition effect in judgments of temporal duration across minutes, days, and months. *American Journal of Psychology, 92*(2), 323-333. doi: 10.2307/1421927

Hinton, S. C., & Rao, S. M. (2004). "One-thousand one... one-thousand two... ": Chronometric counting violates the scalar property in interval timing. *Psychonomic Bulletin & Review, 11*(1), 24-30. doi: 10.3758/bf03206456

Hirsh, I. J. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America, 31*(6), 759-767. doi: 10.1121/1.1907782

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika, 75*(4), 800-802.

Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*(2), 169-183. doi: 10.1037/a0018555

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65-70.

Ishihara, M., Keller, P. E., Rossetti, Y., & Prinz, W. (2008). Horizontal spatial representations of time: Evidence for the STEARC effect. *Cortex, 44*(4), 454-461. doi: 10.1016/j.cortex.2007.08.010

Javitt, D. C., Steinschneider, M., Schroeder, C. E., & Arezzo, J. C. (1996). Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: Implications for schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America, 93*(21), 11962-11967. doi: 10.1073/pnas.93.21.11962

Jirsa, R., Libiger, J., Mohr, P., Radil, T., & Indra, M. (1996). Rhythmic finger-tapping task and fast segmentation of neural processing in schizophrenics. *Biological Psychiatry, 40*(12), 1301-1304. doi: 10.1016/s0006-3223(96)00249-1

Johnson, J. E., & Petzel, T. P. (1971). Temporal orientation and time estimation in chronic schizophrenics. *Journal of Clinical Psychology, 27*(2), 194-&. doi: 10.1002/1097-4679(197104)27:2<194::aid-jclp2270270210>3.0.co;2-f

Jones, C. R. G., & Jahanshahi, M. (2009). The Substantia Nigra, the Basal Ganglia, Dopamine and Temporal Processing. *Journal of Neural Transmission-Supplement*(73), 161-171. doi: 10.1007/978-3-211-92660-4_13

Jusyte, A., Schneidt, A., & Schönenberger, M. (2015). Temporal estimation of threatening stimuli in social anxiety disorder: Investigation of the effect of state anxiety and fearfulness. *Journal of Behavior Therapy and Experimental Psychiatry, 47*, 25 - 33.

Kapur, S., & Mann, J. J. (1992). Role of the dopaminergic system in depression. *Biological Psychiatry, 32*(1), 1-17. doi: 10.1016/0006-3223(92)90137-o

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica, 26*(1), 22-&. doi: 10.1016/0001-6918(67)90005-4

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics, 53*(3), 983-997.

Killeen, P. R., & Weiss, N. A. (1987). Optimal timing and the weber function. *Psychological Review, 94*(4), 455-468. doi: 10.1037//0033-295x.94.4.455

Kitamura, T., & Kumar, R. (1982). Time passes slowly for patients with depressive state. *Acta Psychiatrica Scandinavica, 65*(6), 415-420. doi: 10.1111/j.1600-0447.1982.tb00865.x

Kitamura, T., & Kumar, R. (1983). Time-estimation and time production in depressive patients. *Acta Psychiatrica Scandinavica, 68*(1), 15-21.

Kleinke, C. L. (1986). Gaze and exe-contact - a research review. *Psychological Bulletin, 100*(1), 78-100. doi: 10.1037//0033-2909.100.1.78

Kong, F., & You, X. (2011). Space-time compatibility effects in the auditory modality. *Experimental psychology, 59*(2). doi: 10.1027/1618-3169/a000129

Kornbrot, D. E., Msetfi, R. M., & Grimwood, M. J. (2013). Time perception and depressive realism: Judgment type, psychophysical functions and bias. *Plos One, 8*(8). doi: 10.1371/journal.pone.0071585

Kuhs, H., Hermann, W., Kammer, K., & Tolle, R. (1989). The daily course of the symptomatology and the impaired time-estimation in endogenous-depression (melancholia). *Journal of Affective Disorders, 17*(3), 285-290.

Lalanne, L., van Assche, M., & Giersch, A. (2010). When predictive mechanisms go wrong: disordered visual synchrony thresholds in schizophrenia. *Schizophrenia bulletin*, sbq107.

Lapid, E., Ulrich, R., & Rammsayer, T. (2008). On estimating the difference limen in duration discrimination tasks: A comparison of the 2AFC and the reminder task. *Perception and Psychophysics, 70*(2), 291-305. doi: 10.3758/pp.70.2.291

Lee, K. H., Bhaker, R. S., Mysore, A., Parks, R. W., Birkett, P. B. L., & Woodruff, P. W. R. (2009). Time perception and its neuropsychological correlates in patients with schizophrenia and in healthy volunteers. *Psychiatry Research, 166*(2-3), 174-183. doi: 10.1016/j.psychres.2008.03.004

Lee, K. H., Dixon, J. K., Spence, S. A., & Woodruff, P. W. R. (2006). Time perception dysfunction in psychometric schizotypy. *Personality and Individual Differences, 40*(7), 1363-1373. doi: 10.1016/j.paid.2005.10.021

Lehmann, H. E. (1967). Time and psychopathology. *Annals of the New York Academy of Sciences, 138*(A2), 798-&. doi: 10.1111/j.1749-6632.1967.tb55023.x

Lemke, M. R., Koethe, N. H., & Schleidt, M. (1999). Timing of movements in depressed patients and healthy controls. *Journal of Affective Disorders, 56*(2-3), 209-214. doi: 10.1016/s0165-0327(99)00034-8

Leonard, B. E. (2014). Impact of inflammation on neurotransmitter changes in major depression: An insight into the action of antidepressants. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 48*, 261-267. doi: 10.1016/j.pnpbp.2013.10.018

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*(2), Suppl 2:467-477.

Lewis, P. A., & Miall, R. C. (2003). Brain activation patterns during measurement of sub- and supra-second intervals. *Neuropsychologia, 41*(12), 1583-1592. doi: 10.1016/s0028-3932(03)00118-0

Lhamon, W. T., & Goldstone, S. (1956). The time sense: Estimation of one-second durations by schizophrenic patients. *Archives of Neurology and Psychiatry, 76*(DEC), 625-629.

Lhamon, W. T., & Goldstone, S. (1973). Temporal information processing in schizophrenia. *Archives of General Psychiatry, 28*(1), 44-51.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models* (2nd ed.). Cary, N.C.: SAS Institute, Inc.

Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., & Mason, M. F. (2002). Are you looking at me? Eye gaze and person perception. *Psychological Science, 13*(5), 460-464.

Mahlberg, R., Kienast, T., Bschor, T., & Adli, M. (2008). Evaluation of time memory in acutely depressed patients, manic patients, and healthy controls using a time reproduction task. *European Psychiatry, 23*(6), 430-433. doi: 10.1016/j.eurpsy.2007.07.001

Malapani, C., Deweer, B., & Gibbon, J. (2002). Separating storage from retrieval dysfunction of temporal memory in Parkinson's disease. *Journal of Cognitive Neuroscience, 14*(2), 311-322. doi: 10.1162/089892902317236920

Marazziti, D., Consoli, G., Picchetti, M., Carlini, M., & Faravelli, L. (2010). Cognitive impairment in major depression. *European Journal of Pharmacology, 626*(1), 83-86. doi: 10.1016/j.ejphar.2009.08.046

Martin, B., Giersch, A., Huron, C., & van Wassenhove, V. (2013). Temporal event structure and timing in schizophrenia: Preserved binding in a longer "now". *Neuropsychologia, 51*(2), 358-371. doi: 10.1016/j.neuropsychologia.2012.07.002

Martin, B., Wittmann, M., Franck, N., Cermolacce, M., Berna, F., & Giersch, A. (2014). Temporal structure of consciousness and minimal self in schizophrenia. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.01175

Martinez-Cascales, J. I., de la Fuente, J. M., Santiago, J., & Santiago, J. (2013). When schizophrenia improves time perception. *Journal of Cognitive Neuroscience*, 262-262.

Matthews, W., & Meck, W. (2016). Temporal Cognition: Connecting Subjective Time to Perception, Attention, and Memory. *Psychological Bulletin, 142*(8), 865-907.

Meck, W. H. (1983). Selective adjustment of the speed of internal clock and memory processes. *Journal of Experimental Psychology-Animal Behavior Processes, 9*(2), 171-201. doi: 10.1037/0097-7403.9.2.171

Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Cognitive Brain Research, 3*(3-4), 227-242. doi: 10.1016/0926-6410(96)00009-2

Meck, W. H. (2005). Neuropsychology of timing and time perception. *Brain and Cognition, 58*(1), 1-8. doi: 10.1016/j.bandc.2004.09.004

Mella, N., Conty, L., & Pouthas, V. (2011). The role of physiological arousal in time perception: Psychophysiological evidence from an emotion regulation paradigm. *Brain and Cognition, 75*(2), 182-187. doi: 10.1016/j.bandc.2010.11.012

Mezey, A. G., & Cohen, S. I. (1961). Effect of depressive illness on time judgment and time experience. *Journal of Neurology Neurosurgery and Psychiatry, 24*(3), 269-270. doi: 10.1136/jnnp.24.3.269

Mioni, G., Stablum, F., & Grondin, S. (in prep.). Time perception in anxious and depressed patients.

Mishara, A. L., & Gallistel, C. R. (2005). Deficits in time perception in patients with schizophrenia are due to dysfunctional memory not a clock module. *Schizophrenia Bulletin, 31*(2), 370-370.

Msetfi, R. M., Murphy, R. A., & Kornbrot, D. E. (2012). The effect of mild depression on time discrimination. *Quarterly Journal of Experimental Psychology, 65*(4), 632-645. doi: 10.1080/17470218.2011.608908

Mundt, C., Richter, P., van Hees, H., & Stumpf, T. (1998). Zeiterleben und Zeitschätzung depressiver Patienten [Time-experience and time-estimation in depressive patients]. *Nervenarzt, 69*(1), 38-45.

Münzel, K., Gendner, G., Steinberg, R., & Raith, L. (1988). Time estimation of depressive patients: The influence of interval content. *European Archives of Psychiatry and Clinical Neuroscience, 237*(3), 171-178.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*(12), 2544-2590.

Nachmias, J. (2006). The role of virtual standards in visual discrimination. *Vision Research, 46*(15), 2456-2464. doi: 10.1016/j.visres.2006.01.029

Nenadic, I., Gaser, C., Volz, H. P., Rammsayer, T., Hager, F., & Sauer, H. (2000). Altered cerebral activation in schizophrenia during individually adjusted task difficulty - Evidence from a fMRI study on time perception. *Schizophrenia Research, 41*(1), 134-134. doi: Doi 10.1016/S0920-9964(00)90624-8

Newton, I. (1687/ 1972). *Philosophiae Naturalis Principia Mathematica [The mathematical principles of natural philosophy]*. Cambridge, MA: Harvard University Press.

Nichols, H. S., & Park, S. (2011). Time perception and its relation to working memory in schizophrenia. *Schizophrenia Bulletin, 37*, 248-249.

Normand, S. L. T. (1999). Meta-analysis: Formulating, evaluating, combining, and reporting. *Statistics in Medicine, 18*(3), 321-359.

Nosachev, G. N. (1991). Time perception and surviving by depressive patients with manic-depressive psychosis and attack-like schizophrenia. *Zhurnal Nevropatologii I Psikhiatrii Imeni S S Korsakova, 91*(1), 114-117.

Nosachev, G. N. (1992). Perception and experience of time by patients with depression in manic-depressive psychosis and attack-like schizophrenia. *Neurosci Behav Physiol, 22*(2), 175-178.

Noulhiane, M., Mella, N., Samson, S., Ragot, R., & Pouthas, V. (2007). How emotional auditory stimuli modulate time perception. *Emotion, 7*(4), 697-704. doi: 10.1037/1528-3542.7.4.697

Oberfeld, D. (2014). An objective measure of auditory stream segregation based on molecular psychophysics. *Attention, Perception, & Psychophysics, 76*(3), 829-851. doi: 10.3758/s13414-013-0613-z

Oberfeld, D., Thönes, S., Palayoor, B. J., & Hecht, H. (2014). Depression does not affect time perception and time-to-contact estimation. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.00810

Orme, J. E. (1966). Time estimation and nosology of schizophrenia. *British Journal of Psychiatry, 112*(482), 37-&. doi: 10.1192/bjp.112.482.37

Ouellet, M., Santiago, J., Funes, M. J., & Lupianez, J. (2010). Thinking About the Future Moves Attention to the Right. *Journal of Experimental Psychology-Human Perception and Performance, 36*(1), 17-24. doi: 10.1037/a0017176

Overall, J. E., & Gorham, D. R. (1962). The brief psychiatric rating scale. *Psychological reports, 10*(3), 799-812.

Oyanadel, C., & Buela-Casal, G. (2014). Time perception and psychopathology: Influence of time perspective on quality of life of severe mental illness. *Actas Espanolas De Psiquiatria, 42*(3), 99-107.

Papageorgiou, C., Karanasiou, I. S., Kapsali, F., Stachtea, X., Kyprianou, M., Tsianaka, E. I., . . . Papadimitriou, G. N. (2013). Temporal processing dysfunction in schizophrenia as measured by time interval discrimination and tempo reproduction tasks. *Progress in Neuro-*

*Psychopharmacology & Biological Psychiatry, 40*, 173-179. doi: 10.1016/j.pnpbp.2012.07.017

Parsons, B. D., Gandhi, S., Aurbach, E. L., Williams, N., Williams, M., Wassef, A., & Eagleman, D. M. (2013). Lengthened temporal integration in schizophrenia. *Neuropsychologia, 51*(2), 372-376. doi: 10.1016/j.neuropsychologia.2012.11.008

Pearl, D., & Berg, P. S. D. (1963). Time perception and conflict arousal in schizophrenia. *Journal of Abnormal Psychology, 66*(4), 332-&. doi: 10.1037/h0046325

Penney, T. B., Meck, W. H., Roberts, S. A., Gibbon, J., & Erlenmeyer-Kimling, L. (2005). Interval-timing deficits in individuals at high risk for schizophrenia. *Brain and Cognition, 58*(1), 109-118. doi: 10.1016/j.bandc.2004.09.012

Peterburs, J., Nitsch, A. M., Miltner, W. H. R., & Straube, T. (2013). Impaired Representation of Time in Schizophrenia Is Linked to Positive Symptoms and Cognitive Demand. *Plos One, 8*(6). doi: 10.1371/journal.pone.0067615

Plomp, R. (1964). Rate of decay of auditory sensation. *Journal of the Acoustical Society of America, 36*(2), 277-282.

Rabin, A. I. (1957). Time-estimation of schizophrenics and non-psychotics. *Journal of Clinical Psychology, 13*(1), 88-90. doi: 10.1002/1097-4679(195701)13:1<88::aid-jclp2270130125>3.0.co;2-q

Rammsayer, T. (1990). Temporal discrimination in schizophrenic and affective disorders: Evidence for a dopamine-dependent internal clock. *International Journal of Neuroscience, 53*(2-4), 111-120.

Rammsayer, T. (1993). On Dopaminergic Modulation of Temporal Information-Processing. *Biological Psychology, 36*(3), 209-222. doi: Doi 10.1016/0301-0511(93)90018-4

Rammsayer, T. (1997). Effects of body core temperature and brain dopamine activity on timing processes in humans. *Biological Psychology, 46*(2), 169-192. doi: Doi 10.1016/S0301-0511(97)05251-4

Rammsayer, T. (2009). Effects of Pharmacologically Induced Dopamine-Receptor Stimulation on Human Temporal Information Processing. *Neuroquantology, 7*(1), 103-113.

Rammsayer, T., & Ulrich, R. (2011). Elaborative rehearsal of nontemporal information interferes with temporal processing of durations in the range of seconds but not milliseconds. *Acta Psychologica, 137*(1), 127-133. doi: DOI 10.1016/j.actpsy.2011.03.010

Ratcliffe, M. (2012). Varieties of temporal experience in depression. *Journal of Medicine and Philosophy, 37*(2), 114-138. doi: 10.1093/jmp/jhs010

Rattat, A. C., & Droit-Volet, S. (2012). What is the best and easiest method of preventing counting in different temporal tasks? *Behavior Research Methods, 44*(1), 67-80. doi: 10.3758/s13428-011-0135-3

Regan, D., & Gray, R. (2000). Visually guided collision avoidance and collision achievement. *Trends in Cognitive Sciences, 4*(3), 99-107. doi: 10.1016/s1364-6613(99)01442-4

Repp, B. H. (2000). Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychological Research-Psychologische Forschung, 63*(2), 106-128.

Rhodes, G., Addison, B., Jeffery, L., Ewbank, M., & Calder, A. J. (2012). Facial Expressions of Threat Influence Perceived Gaze Direction in 8 Year-Olds. *Plos One, 7*(11). doi: 10.1371/journal.pone.0049317

Richter, P., & Benzenhofer, U. (1985). Time-estimation and chronopathology in endogenous depression. *Acta Psychiatrica Scandinavica, 72*(3), 246-253.

Rinaldi, L., Brugger, P., Bockisch, C. J., Bertolini, G., & Girelli, L. (2015). Keeping an eye on serial order: Ocular movements bind space and time. *Cognition, 142*, 291-298. doi: 10.1016/j.cognition.2015.05.022

Rosnow, R. L., & Rosenthal, R. (1996). Computing contrasts, effect sizes, and counternulls on other people's published data: general procedures for research consumers. *Psychological Methods, 1*(4), 331-340. doi: 10.1037/1082-989x.1.4.331

Roy, M., Grondin, S., & Roy, M. A. (2012). Time perception disorders are related to working memory impairment in schizophrenia. *Psychiatry Research, 200*(2-3), 159-166. doi: 10.1016/j.psychres.2012.06.008

Rutschmann, J. (1973). Time judgments by magnitude estimation and magnitude production and anxiety - problem of comparison between normals and certain schizophrenic patients. *Journal of Psychology, 85*(2), 187-223.

Ryan, L. J. (2011). Temporal context affects duration reproduction. *Journal of Cognitive Psychology, 23*(1), 157-170. doi: 10.1080/20445911.2011.477812

Santiago, J., Lupianez, J., Perez, E., & Funes, M. J. (2007). Time (also) flies from left to right. *Psychonomic Bulletin & Review, 14*(3), 512-516. doi: 10.3758/bf03194099

Schmidt, H., McFarland, J., Ahmed, M., McDonald, C., & Elliott, M. A. (2011). Low-level temporal coding impairments in psychosis: Preliminary findings and recommendations for further studies. *Journal of Abnormal Psychology, 120*(2), 476-482. doi: 10.1037/a0023387

Senju, A., Hasegawa, T., & Tojo, Y. (2005). Does perceived direct gaze boost detection in adults and children with and without autism? The stare-in-the-crowd effect revisited. *Visual Cognition, 12*(8), 1474-1496. doi: 10.1080/13506280444000797

Senju, A., & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends in cognitive sciences, 13*(3), 127-134.

Sevigny, M. C., Everett, J., & Grondin, S. (2003). Depression, attention, and time estimation. *Brain and Cognition, 53*(2), 351-353. doi: 10.1016/s0278-2626(03)00141-6

Siegel, S. F., & Church, R. M. (1984). The decision rule in temporal bisection. *Annals of the New York Academy of Sciences, 423*(MAY), 643-645. doi: 10.1111/j.1749-6632.1984.tb23481.x

Smith, S. D., McIver, T. A., Di Nella, M. S., & Crease, M. L. (2010). The effect of emotional valence and arousal on time perception: Evidence from a temporal-bisection task. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 64*(4), 314-314.

Stocker, K. (2012). The Time Machine in Our Mind. *Cognitive Science, 36*(3), 385-420. doi: 10.1111/j.1551-6709.2011.01225.x

Stocker, K. (2014). Mental perspective in multiple-event memory and foresight. *Journal of Cognitive Science, 15*(1), 57-95.

Stocker, K., Hartmann, M., Martarelli, C. S., & Mast, F. W. (2015). Eye movements reveal mental looking through time. *Cognitive Science*, 1-23. doi: 10.1111/cogs.12301

Straus, E. W. (1947). Disorders of personal time in depressive states. *Southern Medical Journal, 40*(3), 254-259.

Sutton, A. J., & Higgins, J. P. T. (2008). Recent developments in meta-analysis. *Statistics in Medicine, 27*(5), 625-650. doi: 10.1002/sim.2934

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods, 5*(1), 13-30.

Teixeira, S., Machado, S., Paes, F., Velasques, B., Silva, J. G., Sanfim, A. L., . . . Arias-Carrion, O. (2013). Time Perception Distortion in Neuropsychiatric and Neurological Disorders. *Cns & Neurological Disorders-Drug Targets, 12*(5), 567-582.

Teufel, C., Alexis, D. M., Todd, H., Lawrance-Owen, A. J., Clayton, N. S., & Davis, G. (2009). Social cognition modulates the sensory coding of observed gaze direction. *Current Biology, 19*(15), 1274-1277.

Thayer, S., & Schiff, W. (1975). Exe-contact, facial expression, and experience of time. *Journal of Social Psychology, 95*(1), 117-124.

Thoenes, S., & Oberfeld, D. (2015). Time perception in depression: A meta-analysis. *Journal of Affective Disorders, 175*, 359-372. doi: 10.1016/j.jad.2014.12.057

Todd, J., Michie, P. T., Budd, T. W., Rock, D., & Jablensky, A. V. (2000). Auditory sensory memory in schizophrenia: inadequate trace formation? *Psychiatry Research, 96*(2), 99-115. doi: S0165-1781(00)00205-5 [pii]

Todd, J., Michie, P. T., & Jablensky, A. (2003). Association between reduced duration mismatch negativity (MMN) and raised temporal discrimination thresholds in schizophrenia. *Clinical Neurophysiology, 114*(11), 2061-2070. doi: 10.1016/s1388-2457(03)00246-3

Tracy, J. I., Monaco, C., McMichael, H., Tyson, K., Chambliss, C., Christensen, H. L., & Celenza, M. A. (1998). Information-processing characteristics of explicit time estimation by patients with schizophrenia and normal controls. *Perceptual and Motor Skills, 86*(2), 515-526.

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock". *Psychological Monographs, 77*(13), 1-31.

Treutwein, B., & Strasburger, H. (1999). Fitting the psychometric function. *Perception & Psychophysics, 61*(1), 87-106.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading: AddisonWesley.

Turgeon, M., Giersch, A., Delevoye-Turrell, Y., & Wing, A. M. (2012). Impaired predictive timing with spared time interval production in individual with schizophrenia. *Psychiatry Research, 197*(1-2), 13-18. doi: 10.1016/j.psychres.2012.03.003

Tysk, L. (1983a). Estimation of time and the subclassification of schizophrenic disorders. *Perceptual and Motor Skills, 57*(3), 911-918.

Tysk, L. (1983b). Time-estimation by healthy-subjects and schizophrenic-patients - a methodological study. *Perceptual and Motor Skills, 56*(3), 983-988.

Tysk, L. (1984). Time perception and affective disorders. *Perceptual and Motor Skills, 58*(2), 455-464.

Ulferts, J., Meyer-Lindenberg, A., & Gallhofer, B. (1999). Time discrimination: a comparison between healthy controls, unmedicated schizophrenics, zotepine-treated schizophrenics and schizophrenics treated with conventional neuroleptics. *Neuropsychiatrie, 13*(3), 133-138.

Umbricht, D., & Krljes, S. (2005). Mismatch negativity in schizophrenia: a meta-analysis. *Schizophrenia Research, 76*(1), 1-23. doi: 10.1016/j.schres.2004.12.002

van der Veen, F. M., Roder, C. H., & Smits, M. (2013). Feedback processing in schizophrenia: Effects of affective value and remedial action. *Psychiatry Research-Neuroimaging, 213*(2), 108-114. doi: 10.1016/j.pscychresns.2013.02.004

van Houwelingen, H. C., Arends, L. R., & Stijnen, T. (2002). Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine, 21*(4), 589-624. doi: 10.1002/sim.1040

Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods, 1*, 112-125.

Volz, H. P., Nenadic, I., Gaser, C., Rammsayer, T., Hager, F., & Sauer, H. (2001). Time estimation in schizophrenia: an fMRI study at adjusted levels of difficulty. *Neuroreport, 12*(2), 313-316. doi: Doi 10.1097/00001756-200102120-00026

Vorberg, D., & Wing, A. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of perception and action* (Vol. 2, pp. 181-262).

Wahl, O. F., & Sieg, D. (1980). Time-estimation among schizophrenics. *Perceptual and Motor Skills, 50*(2), 535-541.

Waters, F., & Jablensky, A. (2009). Time discrimination deficits in schizophrenia patients with first-rank (passivity) symptoms. *Psychiatry Research, 167*(1-2), 12-20. doi: 10.1016/j.psychres.2008.04.004

Watt, J. D. (1991). Effect of boredom proneness on time perception. *Psychological Reports, 69*(1), 323-327. doi: 10.2466/pr0.69.5.323-327

Wearden, J. H. (2008). Slowing down an internal clock: Implications for accounts of performance on four timing tasks. *Quarterly Journal of Experimental Psychology, 61*(2), 263-274. doi: 10.1080/17470210601154610

Wearden, J. H. (2015). Passage of time judgements. *Consciousness and Cognition, 38*, 165-171. doi: 10.1016/j.concog.2015.06.005

Wearden, J. H., Denovan, L., Fakhri, M., & Haworth, R. (1997). Scalar timing in temporal generalization in humans with longer stimulus durations. *Journal of Experimental Psychology-Animal Behavior Processes, 23*(4), 502-511. doi: 10.1037//0097-7403.23.4.502

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology, 51*(2), 97-120.

Wearden, J. H., & Lejeune, H. (2008). Scalar properties in human timing: Conformity and violations. *Quarterly Journal of Experimental Psychology, 61*(4), 569-587. doi: 10.1080/17470210701282576

Wearden, J. H., & Pentonvoak, I. S. (1995). Feeling the heat: Body temperature and the rate of subjective time, revisited. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology, 48*(2), 129-141.

Werner, F. M., & Covenas, R. (2010). Classical Neurotransmitters and Neuropeptides Involved in Major Depression: a Review. *International Journal of Neuroscience, 120*(7), 455-470. doi: 10.3109/00207454.2010.483651

Wiener, M., Lee, Y. S., Lohoff, F. W., & Coslett, H. B. (2014). Individual differences in the morphometry and activation of time perception networks are influenced by dopamine genotype. *Neuroimage, 89*, 10-22. doi: 10.1016/j.neuroimage.2013.11.019

Wolfinger, R. D. (1993). Covariance structure selection in general mixed models. *Communications in Statistics-Simulation and Computation, 22*(4), 1079-1106.

Wyrick, R. A., & Wyrick, L. C. (1977). Time experience during depression. *Archives of General Psychiatry, 34*(12), 1441-1443.

Yadid, G., & Friedman, A. (2008). Dynamics of the dopaminergic system as a key component to the understanding of depression. *Serotonin-Dopamine Interaction: Experimental Evidence and Therapeutic Relevance, 172*, 265-286. doi: 10.1016/s0079-6123(08)00913-8

Yamamoto, K., & Miura, K. (2016). Effect of motion coherence on time perception relates to perceived speed. *Vision Research, 123*, 56-62. doi: 10.1016/j.visres.2015.11.004

Yang, Y. K., Yeh, T. L., Chiu, N. T., Lee, I. H., Chen, P. S., Lee, L. C., & Jefffies, K. J. (2004). Association between cognitive performance and striatal dopamine binding is higher in timing and motor tasks in patients with schizophrenia. *Psychiatry Research-Neuroimaging, 131*(3), 209-216. doi: 10.1016/j.pscychresns.2003.07.002

Zakay, D., & Block, R. A. (1996). The role of attention in time estimation processes. In M. A. Pastor & J. Artieda (Eds.), *Time, internal clocks, and movement* (pp. 129-141). Amsterdam; New York: Elsevier.